overSEAS 2013

This thesis was submitted by its author to the School of English and American Studies, Eötvös Loránd University, in partial fulfilment of the requirements for the degree of Master of Arts. It was found to be among the best theses submitted in 2013, therefore it was decorated with the School's Outstanding Thesis Award. As such it is published in the form it was submitted in **overSEAS 2013** (http://seas3.elte.hu/overseas/2013.html)

DIPLOMAMUNKA
MA THESIS

Benkő Ágnes

Anglisztika MA

Angol elméleti nyelvészet szakirány

2013

CERTIFICATE OF RESEARCH

By my signature below, I certify that my ELTE MA thesis, entitled *The role of markedness in phonological theory* is entirely the result of my own work, and that no degree has previously been conferred upon me for this work. In my thesis I have cited all the sources (printed, electronic or oral) I have used faithfully and have always indicated their origin. The electronic version of my thesis (in PDF format) is a true representation (identical copy) of this printed version.

If this pledge is found to be false, I realize that I will be subject to penalties up to and including the forfeiture of the degree earned by my thesis.

Date: ..... Signed: .....

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
Bölcsészettudományi Kar

DIPLOMAMUNKA
MA THESIS

A jelöltség szerepe a fonológiában
The role of markedness in phonological theory

Témavezető:
Dr. Törkenczy Miklós
Egyetemi tanár

Készítette:
Benkő Ágnes
Anglisztika MA
Angol elméleti nyelvészet szakirány

2013

Abstract

This thesis aims to examine the role of markedness in phonology. Markedness is a multifaceted issue that has been widely used in phonology to explain natural (typologically frequent or expected) and unnatural (typologically infrequent or unexpected) phenomena. It has been connected with different units, such as features, segments, classes of segments, rules, syllables, inventories. Markedness has been incorporated into various theories (e.g., standard generative phonology, Optimality Theory); as a result, different concepts of markedness have developed, which involve different assumptions about universality, explanatory power, etc.

The first part of the thesis is a critical review of the general issues arising in connection with markedness. Several criteria have been proposed (e.g., frequency, language acquisition, phonological processes, phonetic factors) with the help of which marked and unmarked elements can be identified. These criteria, however, do not necessarily coincide with the source of markedness. An examination of the source of markedness reveals that markedness is influenced by phonetic as well as phonological factors (among other possible sources).

In the second part of the thesis, two phenomena are presented, which involve processes regarded as determiners of markedness. Through the example of Hungarian vowel harmony and English vowel reduction, we examine phonetic principles in more detail. Both of these processes display neutralisation of quality contrast. Vowel harmony involves the suspension of backness and rounding contrast, while vowel reduction is claimed to be an F2 neutralisation process. The site of neutralisation may be claimed to be a weak position from a phonetic and a phonological point of view as well, where maintaining contrasts is difficult. We argue that articulatory and perceptual principles influence the markedness of vowel harmony and vowel reduction patterns.

Acknowledgements

I would like to say a big thank you to my supervisor, *Miklós Törkenczy*, for his unremitting patience, encouragement and help over the years. His insightful comments on the seemingly endless versions of this thesis were of enormous help, without which this work would not have been possible.

Contents

1 Introduction

In this thesis, we will examine the different senses and uses of markedness in phonological theory, discuss commonly arising issues (section 2.1) with a focus on the scope of markedness, explanatory power, universality, theory-dependency, implications, and criteria. We will explore different possibilities to handle markedness phenomena in certain formalised and non-formalised accounts (section 3). We will investigate the possible sources of phonological markedness (section 4), dwelling on phonetic and phonological factors and their relevance in language typology (section 5) and phonological theories (section 6). The next part of this thesis introduces the possible motivations behind two phonological patterns (vowel harmony (section 7) and vowel reduction (section 8)) often considered to be related to markedness. Finally, we will discuss the connection of the two phenomena and their relation to markedness (section 9).

2 Markedness in phonology

The notion of markedness pervades linguistic theory: it appears in morphology, syntax and semantics as well as in phonology. Though markedness seems to play an important role in linguistics, there is no unified approach to this issue. First, it is difficult to capture what markedness actually is. The concept seems to revolve around the notion of asymmetry, but there is no widely accepted definition (Battistella 1996). Markedness is often used as a cover term for anything that is unusual, unnatural, rare or unexpected. As the scope of markedness seems to be too wide, there have been attempts to decompose markedness. Haspelmath (2006), for example, distinguishes twelve senses and six roles of markedness in linguistics.

Phonological markedness is used to describe and explain a variety of ‘natural’ or ‘unnatural’ phenomena as well as to make predictions about the possible inventories of languages, language change, the stages of language acquisition, etc. The concept of phonological markedness is still very vague, which can hardly be defined in a single, unified way. Numerous expressions are used to describe marked and unmarked elements, of which a non-exhaustive list is provided in (1) (based on Rice 1999, 2007, Hume 2011).

(1)

marked	unmarked
less natural	more natural
less normal	more normal
less optimal	more optimal
active	inactive
not expected	expected
more complex	less complex
not basic	basic
more specific	more general
less stable	more stable
less common	more common
less frequent	more frequent
language-specific	universal
implies unmarked feature	implied by marked feature
unlikely to be epenthetic	likely to be epenthetic
trigger of assimilation	target of assimilation
remains in coalescence	lost in coalescence

retained in deletion	lost in deletion
subject to neutralisation	result of neutralisation
later in language acquisition	earlier in language acquisition
early loss in language deficit	late loss in language deficit
harder to articulate	easier to articulate
perceptually more salient	perceptually less salient
perceptually less salient	perceptually more salient
less phonetically variable	more phonetically variable

There are several issues illustrated in the list in (1). One of the fundamental problems with these terms is that they are used to explain the vague notion of markedness, but most of them are too broad expressions themselves. While it seems unavoidable to talk about basic, simple, or natural segments, inventories, etc., the exact meaning (and context) of these terms is often left unspecified. Second, markedness seems to be composed of several interrelated factors, but the above list is way too long and diverse for every item to count in the same way. Some of the above items may not play a role in markedness phenomena at all, others may turn out to be more or less important. Widely different issues, such as phonetic grounding, language development, or frequency are raised, and they are often conflicting, as in the case of perceptual salience (perceptually more salient elements have been described as both marked and unmarked). Therefore, the close inspection and assessment of the above items as to whether they are the source, criteria, or consequence of markedness appears to be necessary.

2.1 Issues

In what follows, we will touch on a number of problems that arise in connection with phonological markedness, such as the scope of markedness, explanatory power, universality, theory-dependency, implications and criteria of markedness. These issues are closely intertwined, which often causes confusion and potentially leads to a misconception of markedness.

2.1.1 The scope of markedness

One of the many problems that arises in connection with markedness is what unit can be marked or unmarked. Different frameworks, accounts or authors take different stances on this question. Hence, features, segments, classes of sounds, inventories, syllables, and processes/ rules have all been described with the help of markedness.

It is often assumed that features carry markedness information. For instance, in Chomsky and Halle's framework (1968), the units to which markedness values are assigned are features. Nevertheless, the markedness of features is often dependent on context or other co-occurring features. While voicing in stops may be considered marked for several reasons, the voicing of vowels seems to be a universal and natural process. Kean also suggests that markedness theory is "a theory of the distinctive features which characterize the segments of languages" (1975: 2), claiming that unmarked features are the likely specifications, while marked ones are the unlikely specifications.

Some accounts consider the segment to be the relevant unit of markedness. For example, for the Prague school, markedness is a property of individual segments (Brasington 1982). In such a segmental markedness approach, schwa may be assumed

to be the unmarked segment among the vowels in English, as it is the default epenthetic vowel and it is the most likely to delete.

It is also possible to compare the markedness of classes of segments. Voiceless obstruents may be considered unmarked on the basis of word-final devoicing in several languages, such as Dutch, German, Russian or Turkish (Rice 1999). Comparing places of articulation, coronal segments are assumed to be unmarked, for example, on the basis of assimilation processes in Korean.

The markedness of an inventory depends on the segments that it contains. While a simple three-vowel system usually contains unmarked elements (e.g., /a i u/), a more complex system, such as the Hungarian vowel inventory, may be considered marked because it contains marked vowels, such as /ø/ or /y/. The markedness of inventories may also depend on the distribution of vowels in space; accordingly, evenly distributed inventories may be considered less marked.

Considering syllables, it has been widely acknowledged that the CV syllable is the unmarked syllable type, while closed syllables constitute a more marked category. Generally, it is assumed that the more consonants an onset or coda contains, the more marked the syllable is. The unmarked status of the CV syllable is based on cross-linguistic generalisations and implicational relations (e.g., Kaye & Lowenstamm 1981) as well as on phonetic considerations (e.g., Wright 2004).

It is also possible to talk about marked and unmarked clusters; for example, Steriade (1997) mentions that certain onset clusters in German are marked in the sense that they are impossible onset clusters, while others are unmarked, that is, they are possible onset clusters. It is important to point out, however, that it is not the syllabic position that is marked/unmarked, but the actual segments in that particular position. Kiparsky notes that there is a “robust phonological generalization that marked feature

values tend to be suppressed in certain prosodic positions" (2008: 40), and he demonstrates this claim by presenting examples of neutralisation in coda position. Many approaches agree that position or context is significant in determining the markedness value of a segment, but opinions differ on whether it is the syllabic position that should be considered or the phonetic properties of adjacent sounds.

Processes or rules may also be assessed with respect to markedness. For instance, voicing in intervocalic environments may be assumed to be unmarked on the basis of phonetic considerations. Chomsky and Halle provide examples for more natural and expected (2a), and less natural (2b) rules (1968: 401).

(2)

(a) i → u

k, g → č, ǰ / __ i, e[1]

(b) i → ɨ

p, b → t, d / __ i, e

It is possible to evaluate the markedness of segments and classes of segments solely on the basis of the intrinsic characteristics of segments (usually expressed in binary features). The markedness of features also plays a role in determining the markedness of rules, nevertheless, an additional factor, context, also has to be taken into consideration. The markedness of inventories also seems to be only partially dependent on the features of segments. While the featural specification of /e/ is the same in (3a) and (3b), dispersion theoretic considerations (the distribution of segments in vowel space) suggest that (3b) is a more marked inventory (if we assume that redundant features are specified too).

[1] We have kept the original transcription symbols used by Chomsky and Halle (1968).

(3)

(a) /i/ /u/

/e/ /o/

/a/

(b) /i/ /u/

/e/

/a/

In the case of syllable markedness, different factors may be at play. Regardless of the content, the markedness of a syllable may still be determined; for example, a CV syllable is considered less marked than a CVC syllable. What seems to be assessed here is a sequence of syllabic and non-syllabic slots and their position. It is also possible to take the content of segments into consideration, in which case their inherent properties (sonority or perceptual cues) and their position with respect to each other is evaluated.

2.1.2 Explanatory power

A further question is whether the notion of markedness can be utilised to explain certain phenomena, or it is merely the description of cross-linguistic (or language-internal) observations. As explanations are thought to provide more insight, most accounts strive to incorporate markedness in a way to offer an explanation for the observed phenomena. Nevertheless, the explanatory power of some approaches is often considered dubious.

Opinions diverge on explanatory power since the concept of satisfactory explanation varies from framework to framework. Generative approaches see the explanation of phenomena in highly abstract universal principles, which may not be sufficient, for instance, for the advocates of a phonetic approach. Liljencrants and Lindblom (1972) criticise the approach adopted in *The Sound Pattern of English* as it is too formal, and even though marking conventions were introduced to express the intrinsic content of features, they appear to be based on observations only. Similarly,

Ohala mentions that marking conventions are circular, as the marked/unmarked configurations of features are designed to explain why certain sounds or features are more common than others, but they simply "come from our accumulated experience" with the different systems of languages (1990: 159).

As the concept of markedness is often used without clarifying what it is based on, some might consider it a theoretically primitive concept (Battistella 1996). Others are of the opinion that the notion of markedness has no explanatory power if one assumes that markedness does not follow from anything (e.g., Bybee 2001, Hume 2004). Labelling an element as marked or unmarked does not provide deeper insight into the workings of language in itself.

Brasington notes that markedness is "merely an observational notion and that our attention should in fact be directed towards establishing the *constants* which conspire to regulate the occurrence of particular phonetic features in particular environments" (1982: 85). Determining precisely what these constants are is a crucial but complicated task, yet without this theories of markedness cannot offer real explanations to phenomena.

2.1.3 Universality

A well-known problem is the universal versus language-specific nature of markedness. If a given sound is marked in one language, do we have reason to suppose that it is marked in other (or all of the) languages as well? Markedness generalisations often hold for the vast majority of languages, but in a small number of cases unexpected phenomena have been found. For example, there exist languages with a word-initial geminate /hː/ or with initial CCCCCC clusters, but otherwise these phenomena are very rare (Blevins 2006).

The universal approach is very appealing, especially for the advocates of generative phonology, as it can provide neat, categorical generalisations about what is possible in human languages, and it can make useful predictions. On the other hand, it rules out several marginally existing phenomena (such as word-initial CCCCCC clusters noted above).

Universality is also related to the issue of whether such markedness constraints function actively in the minds of the speakers, that is, whether they are included in Universal Grammar. Some authors are of the opinion that markedness is an intrinsic part of Universal Grammar (see e.g., Kiparsky 2008). Alternatively, one may assume that markedness constraints are not part of speakers' linguistic knowledge and are only the remnants of certain diachronic changes or the mechanical results of physical constraints governing articulation and perception (Blevins 2006).

While the universal approach is theoretically more suited for capturing the nature or the general properties of language, it faces several complications. For example, in many languages it is observed that central vowels are targets for assimilation, and vowels often neutralise to central vowels. These facts point towards the unmarked status of central vowels. Still, it cannot be stated that central vowels are always the unmarked segments. Numerous inventories lack central vowels and have only a two-way distinction of frontness-backness. Such a system is observed in Modern Greek, where front vowels pattern as unmarked (Rice 1999). This problem may be explained away by positing some kind of markedness hierarchy, and if the least marked sound is unavailable in a language, the next sound is considered. Nevertheless, the lack of central vowels in a number of languages still remains a problem.

The problem of markedness reversal surfaces in connection with the universal versus language-specific debate. In general, markedness reversal happens when

elements that pattern as unmarked in one language appear as marked in other languages. The notion is relevant in theories which assume universal markedness values, such as Chomsky and Halle (1968). Although nothing explicitly bans markedness reversals in their theory, the framework does not provide a way to formally incorporate markedness reversals. In the case of such theories, instances of markedness reversal may decrease the power of the universal view.

Markedness reversals may be the consequence of not taking every relevant detail into consideration. As Hayes and Steriade point it out, "at least some of the markedness scales relevant to phonology must be built on representations that contain language specific phonetic detail" (2004: 21), as phonologically irrelevant elements may be the cause of different behaviour.

Opponents of the universal view emphasise the importance of certain language-specific aspects of markedness. The fact that different languages may choose different unmarked elements might be taken as evidence that markedness does not hold universally. According to Hume, markedness, at best, can only help determine "the degree of probability that a particular element will surface as unmarked cross-linguistically" (2003: 16). If one takes the language-specific stance, then even though markedness can be useful in describing phenomena in individual languages, it cannot express typologically relevant generalisations, and cannot be used as a general explanatory principle.

2.1.4 Theory-dependency

A thorough inspection of markedness reveals many problems; one is the question of whether markedness should be included in phonological theories or a separate theory of markedness is needed. Some frameworks actually incorporate markedness so that it gains a specific theory-dependent meaning. Such an approach is Chomsky and Halle's generative model (1968), which works with universal binary markedness values assigned to features of segments. Prince and Smolensky's Optimality Theory (1993) also utilises the concept of markedness in the form of violable markedness constraints. Violability suggests that Optimality Theory allows more variability of markedness phenomena. While it is true that candidates do not have to satisfy markedness constraints, the evaluation (on the basis of the raking of constraints) gives only one optimal candidate.

Theories may consider markedness a binary choice or a scalar concept. As the descriptions in (1) exemplify, markedness is often seen as a gradient phenomenon; nevertheless, it is possible to create binary pairs, such as normal/abnormal. Whether strict categorisation into such pairs reveals the true nature of markedness phenomena or overlooks a significant characteristic of markedness is a further question. Generative accounts usually use the marked-unmarked distinction (e.g., Chomsky & Halle 1968, Kean 1975), whereas phonetically-grounded accounts tend to use markedness scales (e.g., Steriade 1997).

Another related issue is the necessity of markedness phenomena to surface in each and every case. Generativist accounts are inclined to consider markedness as something categorical, which manifests in exceptionless laws. By contrast, phonetically-based approaches are more flexible: though the markedness of a sound may prevent it from surfacing in one language, the same sound may be permitted in

other languages. Even in one language the same sound may appear in particular environments and not in others. Nevertheless, there are certain restrictions that hold in phonetically-based accounts: if an element is allowed in a more marked position, it should also be allowed in a less marked position.

2.1.5 Implications

The notion of implication is tied to the universal view of markedness and can only be determined by examining a number of languages. The presence of a marked sound, feature, etc. implies the presence of its unmarked counterpart in a given system. For example, if a vowel inventory has mid vowels, it should also contain high vowels. Similar implicational relations hold between nasals and liquids. It seems that if a language has liquids, it should also have nasal sounds (apparent exceptions exist though) (Rice 2007). Jakobson suggests several typological laws that give predictions on the possible sounds of languages; for example, a language can only have nasal vowels if it also has oral ones (Battistella 1996).[2]

The markedness of syllables has been extensively studied from an implicational perspective. For instance, Kaye and Lowenstamm (1981) consider the markedness values of different onsets and rhymes, and based on the observed sets of syllables in different languages, they arrange syllable types into an implicational relationship. Implicational relations have predictive power; for example, they predict that if a language has CVC syllables then it will also have CV syllables.

The significance of implications may be criticised on several grounds. First, not all of the world's languages have been examined so there is always the possibility that there are exceptional cases that have not been found. A different argument is that

[2] The question arises whether these typological laws hold on the phonetic or the phonological level.

language users cannot be aware of cross-linguistic markedness phenomena, hence, it cannot influence their competence. Nevertheless, implications may play an indirect role: implicational relations help determine what is a possible human language. A further criticism of the importance of implications is given by Hayes and Steriade, who state that "in most cases the laws originate as generalisations over known languages, not as principles explaining why these laws should be expected to hold" (2004: 6).

2.1.6 Criteria

An obvious problem is what to base markedness on. Different criteria have been proposed such as frequency, language acquisition, phonological processes, phonetic factors, etc. One way to approach this question is to treat markedness diagnostics as *correlational*. Battistella notes that markedness may be seen as "an open-ended set of properties [that] results in a kind of 'and/or' view in which markedness has no central definition" (1996: 14). However, markedness may also be considered *criterial*, in which case certain characteristics need to be inspected. If one restricts the number of criteria, it may narrow the scope of phenomena that may be examined and exclude patterns that are traditionally thought to be related to markedness.

There is a widely held view that statistical patterns, principally frequency, correlate with markedness. One of the most important advocates of frequency as the relevant criterion is Joseph Greenberg, who claims that markedness phenomena are directly linked to frequency, in a way that unmarked members are the most frequent (Bybee 2001). Besides the obvious problem of what data to use when measuring frequency, it is also a question whether type or token frequency is the relevant factor. Furthermore, there is no agreement on the question of whether cross-linguistic or language-internal frequency measures provide better insight. One might accept

frequency as the only relevant criterion of markedness, as advocated by Greenberg, but then the question arises whether one needs the concept of markedness at all, since it could be explained by referring to frequency alone (Battistella 1996).[3]

Some suggest that frequency reflects markedness only indirectly. Frequency may be indicative of the markedness of an element because it is in a sense the consequence of it: elements that are easier in some way tend to appear more frequently. What one means by easier is not evident; Frisch (2004), for example, attributes some statistical patterns to processing difficulty. Berent, Lennertz and Smolensky (2011) discuss syllable markedness and mention that the typological frequency of syllables like *blif* is the result of ease of perception and production.

Language acquisition is another commonly cited phenomenon that may constitute a criterion of markedness. The claim is that language development takes place in predictable stages, the acquisition of certain sounds always precedes that of other sounds regardless of the acquired language. One might assume that the acquisition of a marked element presupposes that the corresponding unmarked element has already been acquired (Jakobson 1941, Hume 2011). On the basis of this, it is possible to claim that some sounds are easier or more basic than others. However, language development seems to be dependent on language-specific factors as well.

Phonological patterns (such as alternations, assimilation, distribution) may also be considered criterial for markedness. For the Prague school, the most important phonological pattern that was used to test markedness was neutralisation, which is claimed to yield the unmarked member of an opposition in a particular context. Generally, an asymmetrical patterning is observable in phonological processes, and the

[3] It is another conceptual issue that in standard generative theory frequency is considered an irrelevant factor in the study of competence; it belongs to performance.

cross-linguistic tendency is that unmarked elements are the targets of phonological processes, while marked elements tend to resist change (Hume 2011).

Phonetic factors are often cited to be the relevant criteria to markedness. While articulatory factors are usually included in accounts of markedness phenomena (e.g., Chomsky & Halle 1968), so that there is a preference for articulatorily less complex sounds whose production involves fewer gestures, the effect of perception has often been neglected. Recent phonetically-based accounts (e.g., Kaun 2004, Flemming 2004) emphasise the importance of perception, and base their explanations partly on the perceptual salience of sounds (detailed in section 3.2.1).

Hume (2011) emphasises the importance of cognitive factors in determining markedness. In her opinion, factors such as information content and entropy may be utilised to gain a clear picture of markedness patterns (discussed in more detail in section 3.2.2).

3 Approaches to markedness

As we have seen, markedness is a multifaceted issue that may be approached from various perspectives. Different approaches involve formalised accounts, which incorporate markedness in a more coherent way, and non-formalised accounts, which often include markedness only covertly.

3.1 Formalised accounts

In this section, accounts that incorporate markedness into a theory in a formalised way are presented. Our examination includes structuralist, generativist, underspecification as well as optimality-theoretic accounts.

3.1.1 Structuralist accounts

Two prominent figures of the Prague school of structural linguistics have to be mentioned: Nikolai Trubetzkoy and Roman Jakobson. The term *markedness* is attributed to Trubetzkoy, according to whom markedness has a phonological as well as a phonetic nature. Though markedness is admitted to be related to articulatory factors, statistics, and functional load, the most relevant criterion for him is neutralisation.

Trubetzkoy's extensive study of neutralisation resulted in the distinction of two main types: contextually conditioned neutralisation, in which case neutralisation is dependent on the presence of an adjacent segment, and structurally conditioned neutralisation, which happens in different prosodic positions. In his theory, neutralisation develops to be the defining criterion of markedness, in a way that the outcome of neutralisation will have the unmarked value. This calculation may result in languages having more than one unmarked value, as different neutralisation contexts may yield different unmarked elements. Since neutralisation patterns may be different cross-linguistically, his approach to markedness is a language-specific one (Gurevich 2001).

Roman Jakobson also worked on the concept of markedness. In his view, markedness affects the inventories of languages, phonological rules, sound change and also language development (Rice 1999). In the description of features, he generally uses plus marks for the marked value and minuses for the unmarked, while blanks mean predictable features (Battistella 1996). As Bybee remarks "Jakobson's theory of markedness is a good example of a structuralist theory since it focuses on the structure of categories and proposes that all categories have the same structure" (2001: 115), in that marked members of a category will always be marked with plus values. He proposes a universal feature hierarchy, which manifests itself in language acquisition.

Jakobson also took the phonetic content of sounds into consideration. As Battistella notes, "Jakobson's work on phonological relations emphasizes the existence in language of a series of universally ranked oppositions based in part on the intrinsic acoustic and perceptual content of sounds" (1996: 31).

3.1.2 Generative phonology

The traditional generative approach aims to provide an explicit, truly scientific and formal framework for the description of languages. Generative models characterise segments with binary feature values and make use of feature matrices to derive the existing surface forms from abstract underlying representations with the help of ordered rules.

Chomsky and Halle (1968) develop their theory of markedness in *The Sound Pattern of English* in response to concerns that their model is too formal and does not reflect the content of features. They themselves admit that their approach raises some important issues; for example, that the number of features needed to define a class does not reliably reflect the 'naturalness' of classes. A further related concern is that phonetic principles are not reflected in the theory. Hence, there was a need to modify the model so that it can accommodate the content of features.

For this reason, markedness conventions were introduced, which assign features marked or unmarked values. Feature specifications contain *u* for unmarked and *m* for marked, which are in turn translated into the usual ± binary features (examples are shown in (4) (1968: 405)).

(4)

[*u* nasal] → [- nasal]

[*u* high] → [+ high]

Markedness conventions are designed in a way that unmarked values do not contribute to complexity; thus, it becomes possible to count only the marked features, which give the markedness value of the whole segment. These markedness conventions are universal, which entails that they do not constitute part of a specific grammar.

The markedness of segments or classes of segments can be defined in this way, but a further mechanism is needed to express the 'naturalness' of rules. The notion of linking is introduced for this reason, which ensures that markedness conventions, which change the *u/m* values of features into ± values, are linked to the rules that contain these features. If the environmental conditions are met, feature specifications are assigned. Chomsky and Halle claim that linking conventions help predict rule plausibility.

Although marking conventions can help predict 'natural' patterns, there still remain some problematic points. It seems that marking conventions cannot account for all the relevant phenomena and there exist "overriding considerations of symmetry and feature hierarchy that must be brought to bear in establishing what constitutes an optimal phonological system" (Chomsky & Halle 1968: 409). As Chomsky and Halle note, the optimal five-vowel inventory should consist of /a i u e o/, but the theory suggests that any two of /e o æ ü ɔ ɨ/ could be chosen in addition to /a i u/, since they all have the same markedness value. While many such inadequacies may be resolved by minor adjustments, they seem ad hoc and do not follow from any general principle.

Though the theory of markedness introduced in Chomsky and Halle's *Sound Pattern of English* is developed to compensate for formality and abstractness, markedness conventions have been still criticised for not reflecting phonetic content, and being circular (see section 2.1.2 above).

3.1.3 Underspecification theory

Underspecification accounts build on the generative framework and share most of its assumptions. The general idea of underspecification is that certain feature specifications do not need to be present underlyingly[4] since they are predictable. The unspecified features of segments are later filled in by redundancy rules. The main benefit of underspecification resides in the enhancement of lexical minimality.

Steriade (1995) distinguishes three cases in which one may appeal to underspecification. First, a feature may be predictable because it always appears together with another feature; these cases are referred to as feature co-occurrence. For instance, it is possible not to give any specification to the [voice] feature of sonorants, since they are always voiced. It is often assumed that features frequently co-occurring in the majority of languages are included in Universal Grammar and do not constitute a part of individual grammars.

The second potential case is context-free underspecification, which relates to the markedness of elements. Features may be considered in themselves, regardless of how they combine with other features. For example, in the case of nasality, the unmarked value is assumed to be [-nasal] as the feature [+nasal] has a more limited distribution. Steriade suggests that these features should rather be treated as privative, in which case underspecification is genuine.

The third case to consider is positional neutralisation. Certain feature values may be predicted in positions where neutralisation happens; this may be considered a case of temporary underspecification. Positional neutralisation is also linked to markedness, as the predictable value is usually the same as the unmarked value.

[4] There exist accounts which consider it possible to leave features unspecified even in surface representations; in such cases, the unspecified value is given an interpretation during the pronunciation process.

While predictability is a strict requirement if one wants to leave features unspecified, Steriade argues that not all feature values that are predictable should be left unspecified. In her approach, only feature co-occurrence phenomena are subject to underspecification.

3.1.4 Optimality Theory

Prince and Smolensky (1993) provide a constraint-based theory that also incorporates the notion of markedness. Standard Optimality Theory assumes that constraints are included in Universal Grammar, and the ranking of them defines possible grammars. Originally, constraints are very general formulations, and are of two types: faithfulness constraints, which ensure the preservation of the input (in some respect), and markedness constraints, which enforce the well-formedness of the output. As constraints are violable and may be conflicting, ranking determines which will be the winning candidate.

Optimality Theory provides a formal way of expressing that markedness is not exceptionless and categorical. Since constraints are violable, and it is possible to modify the ranking of constraints, languages with higher ranked markedness constraints will show more unmarked phenomena, while in other languages that rank the same constraints lower we may encounter more marked phenomena.

Markedness constraints may be violated for different reasons. It may be the case that a given markedness constraint is violated (and the relevant candidate can still win) because the candidate has to satisfy a higher ranked (faithfulness or markedness) constraint. Nevertheless, a given markedness constraint can still be violated (and the relevant candidate can still win) as long as competing candidates violate the constraint in the same way.

Optimality Theory is sometimes criticised on the grounds that one can invent almost any kind of constraint to get the right results, therefore, the theory is not restrictive enough. Moreover, constraints that can be applied only to a limited number of grammars and seem 'unnatural' may not be included in Universal Grammar.

As a consequence of the appearance of phonetic principles in Optimality Theory, certain modifications have been proposed. Hale and Reiss (2008) discuss the appropriate level of representation that should be used in Optimality-Theoretic accounts. The increasing influence of phonetics manifests itself in many phonetically motivated constraints. Such constraints (e.g., fortition or tapping) are formulated in purely articulatory terms, which may be difficult to map onto a surface representation. In order to better incorporate the phonetically-based approach, it may be necessary to revise and modify Optimality Theory constraints and the representations they evaluate.

3.2 Non-formalised accounts

Some of the markedness phenomena are explained without using formal apparatus. In this section, different phonetically-based analyses, usage-based accounts, the theory of Evolutionary Phonology and exemplar theory are considered.

3.2.1 Phonetic accounts

Phonological patterns often show cross-linguistic variability; however, variation is not unconstrained, and may be derived from general principles that govern speech production and perception. Phonetically grounded accounts provide explanations on the basis of articulatory and perceptual factors. As the speaker and the hearer are also taken into account, conflicting principles arise (minimal articulatory effort/ ease of articulation and maximal distictiveness/ ease of perception), which languages resolve

differently. Several accounts of markedness phenomena have been proposed, which are grounded in phonetic principles. A wide range of phenomena have been addressed, but typically processes such as assimilation, neutralisation and reduction are analysed (Jun 2004, Steriade 1997, 2001, etc.).

While there is no unified phonetic theory, some general characteristics may be mentioned: advocates of the phonetic approach suggest phonetic motivations for phenomena, they do not assume Universal Grammar, they often propose scales on the basis of perceptual cues, etc. These accounts differ as to whether they adopt some kind of formalised framework, but mostly they are compatible with Optimality Theory.

Such a phonetic account is provided by Flemming (2004), who emphasises two basic principles: minimisation of articulatory effort, and minimisation of possibility of confusion of the speaker. His explanation of phenomena is based on the proposal that markedness is not a property of individual sounds, but of contrasts. He provides an example for the distribution of vowels in inventories based on dispersion theory. In the case of front, back and central vowels, the front-back contrast is more distinctive (therefore more preferred) from a phonetic point of view than front-central or back-central contrasts; hence, central vowels are only possible if there are front and back vowels. However, if frontness-backness is not contrastive in a language, central vowels will be preferred.

Wright (2004) proposes the reformulation of the Sonority Sequencing Constraint as a scalar, perceptually motivated constraint. He emphasises the importance of maximising the robustness of perceptual cues. In his discussion, he reviews auditory cues to place of articulation, manner of articulation, voicing, and vowel quality. He identifies cue robustness as an organising principle. As speech usually occurs in some

noise, robust cues as well as redundancy of information are needed for successful communication.

Wright provides a phonetically-based analysis of the CV syllable. He identifies the CV syllable as the optimal (unmarked), since gestural overlap and optimal signal modulation provide salient cues to the quality of the vowel as well as the place, manner and voicing of the consonant. He also gives an explanation for the widely attested but problematic sCV sequences, which cannot be explained with the traditional view of sonority hierarchy (since clusters such as *sp, st* etc. show sonority falls instead of rises). Though in general CC clusters have worse cues than CV sequences, sibilants have strong internal cues, which helps the correct identification of the sound. In this account, then, the intrinsic properties of sounds help explain the behaviour of an otherwise unexpected phenomenon.

3.2.2 Usage-based accounts

Different factors attributed to usage (such as frequency or information content) are often cited to be criterial for markedness. Some accounts incorporate such factors to provide descriptions of markedness phenomena.

Hume (2004) presents a predictability-based approach, in which the influence of language experience is emphasised. She states that predictability positively correlates with unmarkedness, and that predictable elements are less vital to successful communication. Different factors play a role in determining predictability: perceptual and articulatory factors, functional load, social factors, and language experience, among which the last factor seems to be the most prominent. Her account appears to be of the language-specific type, as she bases predictability on two factors: one is the inventory

of a given language, and the other is the extent to which the elements contained in the inventory are used.

In her account, predictability has two types of effects: the first is termed the *instability of the predictable*, which refers to the observation that predictable elements are prone to change (e.g., reduction, deletion, assimilation processes), and the second is the so-called *bias towards the predictable*, which means that more familiar or frequent patterns have a greater chance of being produced or perceived (e.g., epenthesis, dissimilation).

In another account, Hume (2011) emphasises the significance of information content and entropy, and utilises the concept of probability. Entropy is defined as the amount of uncertainty arising in connection with selecting the right outcome. The higher the entropy of an element, the greater its uncertainty.

The most likely target of a phonological process is the segment that adds little to the entropy of the system; this is achieved if the element has either high probability (low information content) or low probability (high information content and low rate of occurrence). The information content comprises of the frequency of the element, its phonetic salience, and the attention given to it by the language user in context.

Hume presents the example of vowel epenthesis in French. The epenthetic vowel in French is a mid front rounded vowel ([ø] or [œ]), which seems difficult to explain given the assumption that front rounded vowels are marked. It is possible to explain this phenomenon by claiming that front rounded vowels are the most frequent in French, and they are also the most confusable with other vowels. Hence, they have very low information content, so they become the preferred epenthetic segments. The reason why the same segment may be likely to be deleted and epenthesised is that such segments

have low entropy, and those elements are preferred which contribute little to the entropy of the system.

This approach also provides a solution to the problem that perceptually more salient and less salient elements are also regarded as unmarked. According to Hume, the salience of a segment is calculated on the basis of the probability that the segment is correctly identified. Less salient sounds will have a low information content value, and add little to the entropy of the system.

3.2.3 Evolutionary Phonology

Blevins's (2006) theory of Evolutionary Phonology aims to provide explanations for the cross-linguistic distribution of sound patterns. In her examination of typological generalisations, she lays emphasis on the fact that, on the one hand, there are a number of almost universal patterns which have few exceptions and, on the other hand, there are rare, almost impossible patterns with few attested languages. Typological examinations often reveal that certain patterns are non-existent; nevertheless, one cannot draw the conclusion that they are impossible.

In explaining phonological patterns, Blevins incorporates diachronic aspects of languages, phonetic factors as well as extralinguistic factors. She defines five sources of similarity between two languages. Similar patterns may arise due to development from the same language, phonetically motivated development, innate phonological knowledge (largely dependent on speech perception and production), extralinguistic factors (prescriptivism, literacy, etc.), and coincidence. In her model, non-phonological explanations have priority over phonological ones. Blevins also inspects the sources of sound changes; they may be the result of misperception (termed *change*), resolution of

ambiguous signals (termed *chance*), and phonetic variation (termed *choice*). She emphasises similarities between sound patterns and types of sound change.

3.2.4 Exemplar theory

Exemplar theory originates from psychology, but it has been adopted to help explain linguistic phenomena as well. It is mainly used for explaining the organisation and storage of linguistic elements. Categories are represented in the minds of speakers as clouds of memorised tokens of a sound. The categories constitute an organised network-like map, in which similar members are closer to each other. The most typical tokens of a category occupy central positions, while less typical tokens are more peripheral. Exemplars are stored with detailed phonetic knowledge (Pierrehumbert 2000).

Exemplar theory can explain perception phenomena and has a way of accounting for frequency effects and variation; storage characteristics may partly account for the different behaviour of marked and unmarked elements within a language if their behaviour is reflected in their frequency. However, van de Weijer (2009) also highlights some of the theory's shortcomings: exemplar theory makes no predictions on cross-linguistic patterns; therefore, it cannot account for typological observations or cross-linguistic markedness patterns, and it has no production mechanism either.

3.3 Summary

What we have seen so far is that the concept of markedness is used to express some fundamental characteristic of language. It describes an asymmetrical pattern, which involves more basic, natural or expected elements and peripheral, less natural or less expected elements. Markedness manifests itself in all kinds of elements and structure: it may be applied to features, segments, classes, rules, inventories and syllables. Whereas

definitions are often vague, and criteria and source are difficult to pinpoint, in phonological theory there is an intuitive notion of (and broad agreement about) the markedness of phonological entities.

Besides its broad interpretation, markedness has also gained theory-dependent connotations. Several formalised as well as non-formalised accounts make use of the concept of markedness in some way. Analyses, especially formalised ones, often run into problems with markedness phenomena, since it has an unexpected and non-optimal nature. As it seems to depend at least partially on non-phonological factors, treating markedness as purely phonological in nature may be a misconception. These issues will be expanded on in the next section.

4 The source of phonological markedness

Markedness observed in phonology may depend on the internal organisation of phonological elements as well as on some non-phonological factors (e.g., phonetics[5] or cognitive factors). Different criteria, such as frequency, probability, or implicational relations, have been proposed to help identify and describe markedness phenomena. While they may show close connection with markedness, it is to be pointed out that the *source* of markedness is not necessarily the same as the *criteria* (or diagnostics) with which we can categorise or identify markedness. These criteria are based on language-internal or cross-linguistic observations only, and do not follow from independent principles or some identifiable source. They may help make useful generalisations, but have little explanatory power in themselves. Some criteria mentioned above are reviewed here again, with the focus on the nature of their relationship with markedness phenomena.

[5] Some approaches consider phonetics to be an integral part of phonology, in which case articulatory and perceptual influences also count as phonological factors.

Though phonological processes (assimilation, neutralisation, epenthesis, deletion, etc.) have been proposed to help establish the markedness of elements, certain problems arise with them as markedness criteria. First, the accuracy of phonological processes in determining the markedness values of segments is doubtful, since one and the same process may not yield outputs of the same markedness value all the time.

Hungarian voicing assimilation is an example for having marked and unmarked outputs of a process at the same time. Hungarian has a regressive assimilation rule and neutralises the voicing contrast of obstruents in pre-obstruent position. Assimilation may yield voiceless targets if the trigger is also voiceless, or may result in voiced ones if the trigger is also voiced.[6] Some examples are shown in (5).

(5)

tűz [tyːz] 'fire'	*tűztől* [tyːstøːl] 'id.' ablative
vas [vɔʃ] 'iron'	*vasból* [vɔʒboːl] 'id.' elative

Assimilation may be considered an unmarked process from a functional perspective, but its outputs do not necessarily reflect the unmarked value of a feature. While such processes may reveal some markedness asymmetries as they can show the different behaviour of elements, they cannot reliably identify marked and unmarked elements. Another problem is that one cannot explain why only certain processes are relevant in determining markedness and what the connection is between the selected processes. Claiming that markedness originates in phonological processes would not help explain this issue unless some independent motivation is found for their use.

Another criterion, frequency, may influence linguistic patterns in complex ways. On the one hand, the frequency of a pattern or element may be considered the

[6] The minor deviations of /v/ and /h/ are disregarded here.

consequence of how difficult or complex it is. In their examination of the cross-linguistic distribution of segments, Coupé, Marsico and Pellegrino note that frequencies "are the emergent properties of an underlyingly organized structure" (2009: 143). However, type and token frequency are connected to linguistic elements in different ways. While type frequency may show the effects of underlying structure, token frequency is mostly associated with usage: one and the same process may apply differently to different tokens depending on their frequency (consider vowel reduction in English *proton* /'prəʊtɒn/ vs. *atom* /'ætəm/).

On the other hand, the frequency of an element may play a role in how it is stored in the brain and how it behaves from a linguistic viewpoint (compare Exemplar theory 3.2.4). Several experiments have shown that frequent and infrequent elements behave in different ways: frequent patterns may be stored in the memory as chunks and are resistant to change, whereas infrequent elements are created online and tend to regularise. Furthermore, Maddieson claims that "as the frequency of any segment or pattern in a given language increases, the more familiar speakers become with it. This familiarity reduces the complexity of the item" (2009: 100). Consequently, frequency induced storage characteristics may in turn contribute to the different linguistic behaviour of elements.

Implications are sometimes also included in markedness criteria. They display an important characteristic of markedness, namely, that the presence of a marked element usually implies the presence of the unmarked one. What they seem to show is that if more complex elements are present in a given language, less complex elements are also expected. The problem with this criterion is that implications are based on cross-linguistic observations, and as data are incomplete, they do not reliably reflect

exceptionless laws. They may help identify basic and complex elements, but they do not provide an independent source for markedness.[7]

Language acquisition again shows that elements that are more basic are learnt easier and earlier, but the acquisition process may be influenced by other factors as well (e.g., frequency). Language acquisition can provide data on what causes difficulties for children, but the acquisition process does not give an explanation for why certain elements are easier or more difficult to learn. It has been suggested that markedness patterns are observable in second language acquisition as well: the unmarked characteristics of the first language are more easily transferred to the L2 system and are harder to unlearn when second language patterns are different from those of the first language (Carey 2005).

4.1 Phonetics as a source

Phonetic factors are also cited among the criteria of markedness. Articulatory difficulty and perceptual distinctiveness have been suggested to influence sound patterns. Ease of articulation is often claimed to underlie processes such as assimilation, while the lack of perceptual distinctiveness may ban contrasts in certain positions, as in vowel reduction processes (more details in Phonetic accounts 3.2.1).

One may accept that phonological markedness is at least partially determined by phonetics. In view of this, the fact that it is not only phonology that displays markedness patterns becomes interesting. Different levels of language may be (and are) studied separately, still one cannot avoid the question why we see similar asymmetric patterning in morphology, syntax, semantics, etc. One obviously would not like to claim that

[7] Phonetically-based approaches may also make use of implicational relations (see e.g., Steriade 2001). In such a framework, implications may provide a more satisfactory explanation as they are connected to phonetic difficulty.

phonetics has an influence on these matters as well. What seems to be a more acceptable explanation is that the source of markedness is some kind of complexity or difficulty that manifests in phonological behaviour, statistical patterns and, consequently, in a different representation in the brain. Such an approach is considered by Maddieson, who claims that "a given linguistic element or pattern is more complex than another if it is more difficult to execute, more difficult to process, more difficult to learn, or more difficult to retain in memory" (2009: 102).

4.1.1 Calculating phonetic complexity

Calculating phonetic difficulty or complexity is by no means a straightforward task. As phonetics has a gradient nature, integrating phonetic aspects into phonology – a categorical system – may be difficult. Different suggestions have been made as to how to determine the phonetic complexity of an element: it may be possible to identify the complexity of segments on the basis of articulatory difficulty; hence, consonants may be classified as basic, elaborated or complex. Another way of classification may be to distinguish only basic and complex segments. An element may be considered complex if you can remove some feature of the segment and it still defines an existing segment; for example, in the case of nasals: if nasality is removed, an existent non-nasal consonant remains (Maddieson 2009).

4.2 Phonology as a source

As features are based on articulatory/acoustic properties, the phonetic makeup of sounds is an obvious factor in determining markedness. However, the content of features may not be the only factor influencing the markedness of segments. The organisation of segments within an inventory may also determine their phonological behaviour. Certain

phenomena, such as vowel harmony, may shed some light on the possible phonological sources of markedness. All contrastive vowels are assumed to potentially participate in vowel harmony; nevertheless, some vowels may show different, unexpected behaviour (transparency, opacity). These vowels are usually in an asymmetrical position within the vowel inventory. The issue is more complex, as Kiparsky and Pajusalu mention that "a vowel may be neutral even if it has a harmonic partner in the language's inventory: the relevant notion of contrast is a contextual one" (2003: 219). What this seems to show is that not only vowel inventory specifications but also other phonologically relevant issues, such as contextual contrast, may be at play.

4.3 Other possible sources of markedness

Claiming that phonetic factors are influential in asymmetrical patterns does not exclude other potential factors. Identifying phonetics as the only functional factor determining markedness may be an oversimplification, since phonetics may not be in a one-to-one relationship with the phenomena perceived as markedness. The diverse and conflicting nature of markedness phenomena advance the assumption that other factors may also be at play. For instance, sequences of identical or similar elements may be difficult to process. Frisch (2004) suggests that certain similarity avoidance phenomena are connected to processing difficulties, which may arise from the cognitive functioning of the brain. Such constraints may also influence markedness patterns.

5 Variation and limits

Typological studies reveal that there is too much and too little variation of linguistic patterns at the same time. On the one hand, not all logically possible phonological patterns exist and there are remarkable similarities between languages (see e.g.,

Chomsky 1966). On the other hand, particular languages can display great differences of structure and patterns (e.g., Joos 1957). Coupé, Marsico and Pellegrino note that the "coexistence of numerous viable types of linguistic elements and structures [...] reveals that language is a system poorly constrained, or at least presenting numerous degrees of freedom" (2009: 143). Variation is interesting from the viewpoint of markedness examinations because if marked elements present more difficulty, they are expected to be rare. This is exactly what cross-linguistic examinations seem to show; however, the pressure of markedness on linguistic structures cannot be very high, as marked elements may still appear in a number of languages.

5.1 Explaining variation

Different assumptions are made as to what causes variation and limitations on possible patterns. It has been suggested that some innate, universal component is responsible for the observed similarities and differences (Chomsky 1965). A different assumption is that the causes lie in articulatory and perceptual principles and possibly some cognitive factors (see e.g., Newmayer 2005). In either case, it is surprising to see so great differences between languages, since all humans should possess the same innate component or have basically the same articulatory, perceptual and cognitive functions.

One of the reasons why certain languages can allow more complex, marked elements, while others cannot may be in connection with the fact that perception seems to be language specific to some extent. Several studies have shown that language users' perception is biased and has some language specific characteristics. As the child acquires his mother tongue, he takes notice of the existing contrasts of his language and disregards those that play no role in the phonology of his mother tongue. A well-known case is Japanese, in which no phonological contrast exists between /l/ and /r/.

Accordingly, Japanese native speakers have difficulties differentiating the two sounds (e.g., Goto 1971). Ultimately, then, it is the contrasts and distribution of sounds in a language that influence perception and contribute to the variation observed cross-linguistically.

Language-specific phonological restrictions may also cause variation in markedness patterns. Specific constraints (e.g., phonotactic restrictions) may ban patterns that would otherwise be considered unmarked. For example, open syllables are considered less marked than closed syllables; nevertheless, English does not allow word-final open syllables if the nucleus is a plain lax vowel.

5.2 Explaining limits

Cross-linguistic investigations also reveal that certain patterns are not (or rarely) found in languages. There are two kinds of explanations for missing marked patterns. One explanation considers missing patterns to be the result of some general restriction in the phonological component (assuming some innate, universal aspects of language). A possible critique of this view is that universal restrictions are often deduced from data, which is incomplete. Moreover, there may be patterns ruled out by the phonological component that still exist.

The other explanation claims that there are no restrictions on patterns imposed by phonology and marked patterns are lost simply because they are more difficult than others (de Lacy 2006). Transmissibility problems then cause a marked pattern not to be continued (compare Evolutionary Phonology 3.2.3). Problems with this approach involve cases in which functional restrictions do not hold but patterns still do not exist. For example, the English diphthong [aʊ] may not be followed by a non-coronal segment, though there is no such restriction on other combinations of diphthongs and

non-coronal segments, and it does not seem to stem from any phonetically motivated constraint. The need to impose arbitrary restrictions on such patterns may decrease the power of purely functional explanations.

6 The role of phonetic principles in phonological theories

Identifying phonetics as a source of markedness does not automatically mean that it has to play a significant role in phonological theories. There remain many open questions about the necessity of markedness to be integrated into phonological theories. There are different paths to follow: functional motivation may provide an explanation for markedness phenomena; however, some authors have raised doubts about the importance of utilising the source of markedness for different reasons.

After the relevance of phonetics is settled, one reasonable way to proceed may be to make the phonetic bases of markedness more transparent. Markedness may not have a purely phonological nature; hence, instead of artificially incorporating the concept of markedness into phonological theory, its source (i.e., phonetic principles) should be incorporated. This view then involves bringing phonetics and phonology closer together, and integrating some aspects of phonetics into phonology. This idea is adopted by many phonetically-based accounts. These approaches have to face some challenges, such as the need to reconsider levels of representation, since phonetic details have been demonstrated to matter in asymmetrical patterns. Another area of future work may be some formal way of expressing phonetic motivation.

A different line of argumentation is that even though markedness patterns observed at the phonological level may be the result of phonetics, they do not necessarily need to be included in phonological theories as they may not be influential or significant in phonology after all. From the language learner's point of view, a given

language contains an arbitrary subset of marked phenomena which do not differ in behaviour from regular, unmarked patterns. Speakers learn marked and unmarked patterns in just the same way (Buckley 2000). Nonetheless, the order and speed of acquisition of linguistic elements may be taken as evidence that marked/unmarked elements do not present the same level of difficulty for the learners of a language. It has been noted, however, that "the phonological system of the target language as reflected in the input largely determines the way in which acquisition proceeds" (Fikkert 2000: 234).

A further argument against including markedness in phonological theories draws attention to the diachronic evidence that *marked* phenomena may be quite resistant to change in languages, which also suggests that these patterns are not treated differently from unmarked ones (Buckley 2000). One may then conclude that whatever the source of these patterns is, it is immaterial to phonology. Before such a conclusion is reached, however, it may be worth examining what the exact nature of the relationship of markedness to diachrony is.

A further argument against the incorporation of phonetics is that though phonetic principles may surface in languages, they are always subject to the phonology of a given language, which may alter (generalise, impose further constraints, etc.) phonetically motivated patterns, so that synchronically they will not reflect phonetics on the whole (Chitoran & Cohn 2009). Nevertheless, if the behaviour of an element may be derived from its markedness value (originating from phonetic and/or other properties), it may contribute to the better understanding of how language works.

A version of the *duplication problem* arises in connection with the nature and place of phonetic motivation. Some authors have raised doubts about the necessity of treating markedness patterns both as phonetically motivated (arising from articulatory

and perceptual restrictions) and as phonologically conditioned (meaning that patterns are subject to certain general principles or restrictions). As phonological restrictions may be rooted in phonetics, the same matter should not be included in the model twice (Chitoran & Cohn 2009).

6.1 Examining the role of phonetic principles

In what follows, we will consider two phenomena that involve the behaviour of vowels: Hungarian vowel harmony and English vowel reduction. These phenomena may be worth studying from a markedness perspective as they involve processes that are often regarded as determiners of the markedness values of elements. Vowel harmony may be seen as long distance assimilation or neutralisation, and vowel reduction is also a neutralisation process. We will examine the phonetic grounding of these two phenomena and compare the predictions made by phonetics with the phonological patterning of elements.

7 The phonetic bases of vowel harmony

Linebaugh (2007) reviews different accounts of the phonetic grounding of vowel harmonies. One such suggestion is that vowel harmony enhances the perception of word boundaries. Experiments on Finnish vowel harmony pointed to the fact that speakers use vowel harmony to determine word boundaries. Though vowel harmony may help identify boundaries in certain cases, it fails to do so in many others (e.g., in sequences of front harmony words).

A different phonetic explanation is provided by Kaun (cited in Linebaugh 2007), who claims that the extension of a feature over a domain helps identify the correct feature value of segments. As the feature has a longer duration, listeners have a greater

chance of identifying it correctly. Once identified, the quality of the following vowels is predictable within the domain. Kaun's proposal is partly based on Steriade's (cited in Linebaugh 2007) findings that contrasts that are perceptually difficult are only allowed in certain positions where identification is easier (i.e., word-initial, word-final, and metrically strong positions). These positions are supposed to have greater duration, which helps identification.

Another phonetically-based account claims that vowel harmony is the result of listeners' misapprehension of vowel-to-vowel coarticulation. Research has shown that speakers are sensitive to the acoustic effects of vowel-to-vowel coarticulation and compensate for it in the determination of the makeup of vowels. The claim is that when language users cannot attribute the coarticulation effects to a neighbouring vowel, they will perceive it as the intrinsic property of the vowel. Vowel harmony systems then develop as the phonologisation of coarticulation effects (Linebaugh 2007).

In Lloret's (2007) view, the driving force of vowel harmony is either ease of articulation, which results in gestural uniformity and, consequently, in perceptual advantages or, alternatively, vowel harmony may originate from the drive to increase the perceptual salience of some feature and gestural uniformity follows only as a consequence. In the first case, the trigger is in strong position, and the targets are in weak positions (strong/weak positions may refer to stressed/unstressed syllables as well as morphologically strong/weak elements). In the second case, the harmonic feature appears in a weak position and spreads to strong positions. Kaun's (2004) account of rounding harmony (detailed in section 7.2.4) appears to take only the second case into consideration, since she assumes that good triggers show weak effects of rounding, while good targets show strong effects.

7.1 Vowel harmony typology

Vowel harmony is considered to be a cross-linguistically common phenomenon (Finley & Badecker 2008). The types of vowel harmony may be differentiated on the basis of what feature displays harmony: rounding, backness, height, tongue root, and complete harmonies may all be observed in the world's languages, though with different cross-linguistic frequencies (Rose & Walker 2011). Backness harmony involves the agreement of the feature [back] and is found in many languages, such as Turkish, Hungarian, Finnish. Rounding harmony involves the feature [round], and is often subject to specific restrictions. Languages exhibiting rounding harmony patterns are likely to show harmony for some other feature as well (Rose & Walker 2011). Rounding harmony languages include Khalkha Mongolian, Turkish, Tuvan, Tunica, etc.

Height harmony depends on vowel height and is often of the dominant type (detailed below). It usually applies only to a subset of the vowels, and contextual limitations may restrict its scope of application. A further characteristic of height harmony is that harmony for the feature [low] is cross-linguistically rare. Languages with height harmony include Lena, Kisa, etc. Tongue root harmony is based on the feature [ATR], and is found, for instance, in the Pulaar dialect of Fula, or in Maasai. Tongue root harmony languages usually apply to mid vowels, [+ATR] low vowels and [-ATR] high vowels are generally avoided. Languages with complete harmony (also called *vowel copy harmony*) show the agreement of all the features of vowels, and usually involve some restriction on the intervening consonants (Rose & Walker 2011).

Vowel harmony patterns may also be classified according to the way vowel harmony works. The terms *directionality* and *dominance* are employed to describe the way the trigger is chosen and how it spreads its feature(s). Languages with directional vowel harmony exhibit patterns with the leftmost vowel spreading some feature

rightwards or the rightmost vowel spreading its feature leftwards. Dominant systems work in a different way: a dominant feature value is selected, and if it appears somewhere within the domain, it spreads in both directions (Finley & Badecker 2008). Stress-triggered harmony may be considered a subtype of dominance harmony (Finley & Badecker 2008). Backness harmonies may show directional or dominant patterns, but height harmonies are usually of the dominant type (Linebaugh 2007).

7.2 Hungarian vowel harmony

Through the example of Hungarian vowel harmony we will explore the extent to which phonetics can predict the characteristics of a phonological process and how it influences the markedness of elements. Direct functional motivations may be hard to find for all the characteristics of such a phonological process, as phonology categorises and generalises variable and gradient phonetic patterns. We will try to demonstrate that phonetics, phonology and possibly morphology are at complex interplay in the determination of the Hungarian vowel harmony pattern.

7.2.1 The Hungarian vowel inventory

Hungarian contains 14 vowel phonemes which constitute 7 short-long pairs. On the basis of the horizontal position of the tongue, one can distinguish front /i, iː, y, yː, ɛ, eː, ø, øː/ and back /u, uː, o, oː, ɔ, aː/ vowels. Vertically, three distinctive heights must be differentiated: high vowels /i, iː, u, uː, y, yː/, mid vowels /o, oː, ø, øː, eː/, and low vowels /ɔ, aː, ɛ/.

Phonetically, [aː] is produced with a lower tongue position[8], and the mid height can be divided into upper-mid [eː, øː, oː] and lower-mid categories [ø, o], as long

[8] Mády (2008) argues that the articulation of [ɔ] is also more open.

vowels are less open. Phonetically, short and long vowels are not of the same quality, which is more noticeable in the case of low vowels.

Rounding is a distinctive feature in Hungarian; rounded vowels include /o, oː, ø, øː, y, yː, u, uː, ɔ/ while unrounded ones comprise of /i, iː, ɛ, eː, aː/. Vowel length is also a distinctive feature in Hungarian (Siptár & Törkenczy 2000).

(6)

	front		back	
	unrounded	rounded	unrounded	rounded
high	i, iː	y, yː		u, uː
mid	eː	ø, øː		o, oː
low	ɛ		aː	ɔ

7.2.2 Backness and rounding harmony in Hungarian

Hungarian displays both backness and rounding harmony. Backness harmony applies to all the vowels, whereas rounding harmony applies only to a restricted set of vowels. Hungarian vowel harmony is of the directional type, and it proceeds from left to right.

An essential distinction has to be made between vowel harmony in roots and vowel harmony in suffixes, since they show slightly different patterning. Many roots contain a combination of back vowels plus front unrounded vowels ([i], [iː], [ɛ], or [eː]), in which case harmony does not seem to be operative. However, on the basis of their phonological behaviour, front unrounded vowels may be considered neutral since they behave transparently with respect to backness harmony and take back suffixes. Some examples are provided in (7).

(7)

víz+nAk[9]	'water' dative	*víznek*
bál+@k	'ball' plural	*bálok*
ipar+hOz	'industry' allative	*iparhoz*
papír+rA	'paper' sublative	*papírra*

Nevertheless, front unrounded vowels are not transparent to the same extent, and may be arranged into a transparency hierarchy: [i(ː)] > [eː] > [ɛ] (termed the *height effect* by Hayes and Londe (2006)). The longer the sequence of neutral vowels in a back + neutral stem, the more likely it is that the suffix will have a front vowel (the *count effect* in Hayes and Londe's (2006) terminology).

While most of the non-harmonising roots can be disregarded as they contain back plus neutral vowels (so-called *mixed roots*), one can still find some words that do not contain neutral vowels and still do not show the effects of vowel harmony. These are disharmonic stems that contain back and non-neutral front vowels as well (e.g., *sofőr* 'chauffeur', *nüansz* 'nuance').

In root plus suffix combinations, suffix vowels are expected to show harmony with the rightmost root vowel. If this vowel is neutral, the preceding vowel determines the frontness/backness of the suffix. Antiharmonic stems demonstrate a special case, as they contain only front unrounded vowels but take back suffixes, as in the case of *híd* 'bridge' – *hídak* 'id.' pl. This is unusual since all neutral roots are expected to take front suffixes.

Different views have been expressed on the domain of application, since root-internally vowel harmony seems to be less restricted and it does not always apply.[10]

[9] We use the traditional notation (capital letters) to refer to front as well as back alternants of suffixes (as in Törkenczy 2011).

Some authors suggest that the domain of application is the phonological word (e.g., Siptár & Törkenczy 2000); thus vowel harmony applies within morphemes, across synthetic and analytic suffix boundaries, but does not apply across word boundaries.

Rounding harmony is claimed to be a more restricted phenomenon than backness harmony. Rounded-unrounded pairs can only be found among front vowels; hence, rounding harmony does not involve back vowels. Front vowels of any height may act as a trigger, but ternary and quaternary alternating suffixes only contain mid and low vowels (o/ø/ɛ/ɔ), among which only mid vowels may be front rounded, which may be targeted by rounding harmony.

Rounding harmony applies only to suffixes and does not cause morpheme-internal harmony. Front unrounded vowels do not behave as neutral with respect to rounding harmony, they do not show transparency effects, therefore, it is the roundedness of the final front vowel of the root that determines the roundedness of the suffix. A further characteristic of rounding harmony is that it does not show vacillation (Törkenczy 2011).

(8)

kürt+@k	'horn' plural	*kürtök*
körte+hOz	'pear' allative	*körtéhez*
likőr+@k	'liqueur' plural	*likőrök*

Suffix vowels may be alternating or non-alternating, and may further be distinguished on the basis of how many vowels participate in the alternation. Alternating suffixes may then be classified into binary suffixes (such as y/u, ø/o, ɛ/ɔ), ternary suffixes (e.g., ø/o/ɛ), and quaternary suffixes (ø/o/ɛ/ɔ). Non-alternating or

[10] Active application of vowel harmony is observed in epenthetic stems, such as *bokor* 'bush' – *bokrok* 'id.' pl., *ökör* 'ox' – *ökrök* 'id.' pl. (Törkenczy 2011).

invariant suffixes include productive suffixes (containing the neutral vowels i, iː, eː) and non-productive ones (containing, oː, ɔ). There is one productive invariant suffix that has a non-neutral vowel (*-kor*) (Törkenczy 2011).

7.2.3 The stress pattern of Hungarian

Stress in Hungarian is usually signalled by amplitude and pitch. The first syllable of open-class words is potentially stressed, but the word's position in the sentence determines whether stress manifests itself or not. Most authors, such as Kálmán & Nádasdy (1994), recognise only stressed and unstressed syllables, and they argue against the phonological relevance of secondary stresses in Hungarian.[11] They note that though it is possible to stress non-initial syllables (e.g., in compound words), these stresses are always optional and play no role in phonological patterns.

Stressed vowels tend to have a longer duration and greater intensity; therefore, they are perceptually good positions (i.e., they facilitate recognition). Unstressed vowels, on the other hand, are less prominent and are considered perceptually weak as they are shorter and less loud. Word-initial syllables are prominent not only because they are the first syllable of a word, but also because they are stressed.

What we observe in Hungarian is that in such positions no neutralisation of vowel contrasts occurs, and any vowel may appear. Though non-initial unstressed positions are perceptually bad positions, Hungarian vowels do not show reduction phenomena. However, these are the positions that show certain restrictions on the appearance of specific vowel contrasts. By suspending the contrast, features become predictable.

[11] Some authors suggest that there is secondary stress in Hungarian, which applies to every second syllable after the primary stress, while others suggest that Hungarian secondary stress is weight-sensitive (Grimes 2007).

7.2.4 Contrasts and inherent properties

Besides the inherent cues of segments, the existing contrasts in an inventory may also influence the perception of sounds. For instance, back vowels are usually rounded, which may imply that back unrounded vowels are dispreferred (marked) sounds. As it has been noted, perception is a language-specific process to some extent; as it is largely dependent on the existing contrasts of a given language. Therefore, back unrounded vowels may not be perceptually difficult if rounding is non-contrastive in back vowels. If rounding is contrastive, however, perceptual problems may arise (Flemming 2004). Kaun (cited in Linebaugh 2007) also remarks that the backness and rounding contrast of vowels is less perceptible if rounding is contrastive in a language. In systems that contain only rounded back vowels and unrounded front ones, the presence of lip rounding causes lower F2 values on back vowels and the absence of lip rounding preserves higher F2 values in the case of front vowels, which contributes to the better perception of backness and rounding. If rounding is contrastive, however, back vowels with higher F2 values and front vowels with lower F2 values are also possible, which endangers the perception of F2 contrasts.

This may be demonstrated in Hungarian as well: in view of the fact that vowel harmony is a process that helps the correct identification of vowels in perceptually weak positions, it is not surprising that only front vowels participate in rounding harmony and back vowels do not, as they do not have unrounded counterparts, which would make perception difficult. As front rounded vowels are perceptually closer to back rounded vowels, the insufficient F2 contrast may cause perceptual problems. Backness and rounding harmony then help the unambiguous identification of the features [round] and [back] by transferring the contrast to an earlier, more prominent position.

F2 perception problems may also shed light on the behaviour of neutral vowels with respect to their occurrence in roots. Front unrounded vowels may freely occur in roots that contain back vowels, such as in *papír* 'paper,' *iroda* 'office.' The front-back vowel co-occurrence restrictions in roots hold only for back and rounded front vowels,[12] whose F2 values are more easily misperceived. As front unrounded and back rounded vowels have the most distinct F2 values, the accurate identification of backness value may not be difficult in these cases.

Kaun (2004) presents a typology of rounding harmony, which Hungarian does not seem to fit. She suggests that a typical rounding harmony setting involves non-high triggers and high targets (a good target shows dramatic effects of lip rounding, while a good trigger displays little effect of harmony), and the trigger and the target usually agree in height. The phonetic grounding for this phenomenon is that high and back vowels are perceived as more rounded. Rounding harmony increases the likelihood that the rounded quality of the trigger will be accurately identified, as non-high triggers have a weak acoustic effect of lip rounding, while the roundedness of high targets is easier to perceive. The tendency to avoid cross-height harmony is explained by gestural uniformity over the scope of harmony.

Nevertheless, this is not what we see in Hungarian: triggers may be mid or high, while targets can only be mid. It may then be possible that Hungarian rounding harmony works in a different way and is not motivated by the weak cues of the trigger. Moreover, high vowels may not be good targets from every perspective. Walker claims that high vowels may present more perceptual difficulty: "[a]ll else being equal, high vowels are expected to be lower in amplitude and shorter in duration than non-high vowels, because of their narrower aperture and lesser jaw lowering" (2011: 28). As the

[12] Rare exceptions include *sofőr* 'chauffeur,' *importőr* 'importer,' etc.

intensity and duration of high vowels is less than that of non-high vowels, they may be considered perceptually weaker sounds even if they show lip rounding more.

7.2.5 Duration of vowels

Duration is an important factor in the perception of vowels, as longer duration can help the accurate identification of vowel quality. Length is considered a distinctive feature of Hungarian vowels, since short-long vowels distinguish meanings, as in *agy* 'brain'–*ágy* 'bed,' *kor* 'age'–*kór* 'disease.' However, it has been shown that the phonetic length of Hungarian vowels is variable and may be dependent on the length of words, more precisely, on the presence and number of following vowels. For example, some experiments show that the duration of a monosyllabic root vowel decreases if suffixes are added to it (show in (9))(Gósy 2004: 109).

(9)

tát	tátogatóknak
[aː]	[aː] [o] [ɔ] [oː] [ɔ]
210	110 80 90 110 140 (ms)

Other experiments, however, do not confirm this finding. White and Mády (2008) examined words of different lengths in continuous speech and did not find a correlation between vowel duration and the length of words. Phrase-final lengthening effects are observable in Hungarian though. Length considerations then do not definitely point to the strong status of the initial syllable or the final one.

Short-long vowel pairs differ in quality as well as quantity, nevertheless, several experiments have shown that the identification of short-long pairs is very uncertain. Short and long low vowels differ from each other in quality more clearly, consequently,

they are usually identified correctly. In the case of mid and especially high vowels, the quality difference of short-long pairs is very small.

Formant structure examinations reveal that there is not only an overlap of /u/-/uː/, /i/-/iː/, and /y/-/yː/, but long mid vowels also partly overlap with high vowels (Mády & Reichel 2007). Moreover, it has been suggested that the units of perception may be larger than those of production, which potentially leads to more confusion in perception (Meunier & Frenck-Mestre & Lelekov-Boissard & Le Besnerais 2003).

Length and height show a notable connection: Kassai (1994) mentions that there seems to be a tendency to identify vowels of shorter duration as higher. Since phonetic length of vowels is variable, duration may not constitute a reliable cue not only to the identification of short-long pairs, but also to the height of vowels. The height restrictions imposed on suffix vowels may be attributed to the potential confusability of non-low vowels (caused by their denseness in vowel space and the perception of variable length).

7.2.6 Motivation of Hungarian vowel harmony

Authors providing analyses of vowel harmony patterns often base their accounts on a distinction between strong and weak positions. Steriade (cited in Linebaugh 2007) identifies word-initial, word-final and metrically strong positions as prominent, whereas Lloret (2007) differentiates strong and weak positions on the basis of stress and morphology. Walker (2011) mentions four positions which may be considered strong or prominent: stressed syllables, initial syllables, final syllables, and morphological roots/stems. She remarks that final syllables seem to display mixed effects: on the one hand, the potential of final phonetic lengthening supports their identification as a strong position, on the other hand, this position often shows drops in pitch and intensity, loss

of voicing, etc., which imply its weak status. Walker lists several different phenomena for the justification of roots and stems as strong positions (phonological contrasts preserved in roots, absence of vowel deletion in hiatus position in roots, etc.), which show that morphological characteristics also influence phonological phenomena.

In Hungarian, the first syllable of a word is indeed a strong position (which may be in connection with lexical processing) and it is also stressed (it is not necessarily reflected in the length of the stressed vowel though). As lexical stress falls on the first syllable of words, it cannot be tested whether it is the initial position or stress that exerts its influence; nevertheless, any vowel may appear in this position.

While certain restrictions hold on non-initial root syllables, they still show more combinatorial possibilities of vowels than suffixes. This may be a morphological influence: roots are treated differently from suffixes, even though non-initial root syllables are just as unstressed as suffixes. The strength of roots may be motivated by some cognitive factors, such as processing characteristics or lexical access. In morphologically weak positions Hungarian vowels can only contrast for height, as the backness (and roundedness) of the vowels is dependent on the vowel of the root syllable. Final position is expected to show mixed effects: Hungarian vowels may lengthen in word-final position (according to some sources), but the same restrictions apply to them as to other non-final suffixes. Depending on the actual word forms, harmonising suffixes may be in final or non-final position; their uniform behaviour may be attributed to the generalising and categorising nature of phonology.

7.2.7 Summary

What phonetics can predict is that more contrasts are allowed in stressed/initial positions. Considering contrasts and properties of sounds, one suggestion is that

distinctive rounding causes less distinct F2 values, which may contribute to the motivation of backness and rounding harmony. Moreover, the potential confusability of non-low vowels may provide some motivation for the height restrictions imposed on suffix vowels.

Vowel harmony shows prosodic and morphological influence as well. Considering positions, the first syllable of a word may then be considered a strong position as it is initial and always stressed in Hungarian, and as we have seen, all vowel contrasts are licensed in word-initial syllables. The strong status of non-initial root syllables (signalled by more combinatorial freedom) may be attributed to lexical access or some processing factors.

Vowel harmony exemplifies the neutralisation of certain contrasts (backness and rounding) in specific dependent positions. On the basis of the harmony process, however, it does not seem possible to unequivocally state which of the feature values (or segments, or classes of segments) are the marked and unmarked ones, as neutralisation does not always yield the same result.

Neutral vowels show more combinatorial possibilities as they may appear in all positions within a word and may combine with back as well as front vowels in roots, which implies that they are the unmarked elements in the vowel inventory. Their distribution does not give an explanation as to why these specific vowels behave in the way they do. In their examination of vowel harmony systems, Kiparsky and Pajusalu (2003) also propose that neutral vowels have the unmarked feature specification of the harmonising feature. Their account is rooted in optimality-theoretic grounds and so they base their opinion on the general idea that markedness constraints (which enforce neutrality in their account) produce unmarked outcomes. While purely phonological approaches may not be able to account for the behaviour of neutral vowels, their

unmarked quality may be shown to have a phonetic motivation (as described in section 7.2.3). The unrounded quality of neutral vowels makes them more distinctive on the F2 dimension which may be a cause of their greater combinatorial freedom.

7.2.8 Suggestions for further research

We have focussed our attention on the perception side of vowel harmony, but articulatory characteristics (namely, gestural uniformity) are also mentioned among the possible motivations of such processes. Therefore, articulatory gains should also be examined in connection with Hungarian vowel harmony. As it is suggested in Kaun's (2004) study, in sequences of identical vowels, languages with vowel harmony exhibit gestural uniformity as opposed to languages that do not show harmony. In such a case, gestural uniformity is expected to show its effects on the intervening consonants as well, which may be examined in Hungarian, too.

Another articulation-related question concerns the variability of production. In connection with vowel reduction, Flemming (2009) observes that schwas in non-contrastive positions show more variability than those that contrast with high vowels. The question arises whether the production of Hungarian vowels in positions where all vowel contrasts are possible (in the first syllable of a root) is more constrained than in predictable non-contrastive positions (i.e., in suffixes).

A further area of research may be the phonetic length and its connection with stress and position in Hungarian. As conflicting evidence has been found with respect to the duration of vowels and word length, further experiments may shed light on the length of vowels in different positions and how it influences the vowel harmony pattern.

8 The phonetic bases of vowel reduction

A cross-linguistic examination of vowel reduction reveals that reduction mostly affects the height dimension of segments, and does not neutralise backness and rounding contrast (Flemming 2005). Unless all contrasts are neutralised, backness and rounding maintain their distinctive quality. If all contrasts are neutralised, vowel reduction usually results in a central vowel. Reduction phenomena may be interpreted as the effect of the ease of articulation principle (e.g., van Bergem 1994), or the effect of perceptual principles (e.g., Burzio 2007). Attributing vowel reduction to only one of these principles may be an oversimplification, as several other factors may also influence vowel reduction, such as stress, frequency, speaking style, syllable type, position of the vowel within the word, phonemic context, etc. (van Bergem 1994).

Flemming (2005) bases his account of phonological vowel reduction primarily on three constraints: maximisation of the distinctiveness of contrasts, minimisation of articulatory effort, and maximisation of the number of contrasting vowels. He provides an explanation of vowel reduction on the basis of the shortness of unstressed vowels: it is more difficult to produce distinct vowel qualities for shorter vowels, as one needs faster articulator movements, which is more effortful.

Stress seems to show a strong connection with reduction phenomena, but Flemming claims that a direct relation may not be posited, since it is a *correlate* of stress (namely, vowel duration) which is the primary cause of reduction. The distinction may be significant as Flemming mentions languages in which the shortness of vowels induces reduction independently of stress.

Another cross-linguistic investigation of reduction patterns suggests that different functional motivations underlie different types of reduction. Crosswhite (2004) distinguishes two different types of vowel reduction: contrast enhancing reduction and

prominence reduction. Contrast enhancing reduction involves cases in which perceptually undesirable vowels reduce in unstressed positions. Perceptually difficult contrasts mostly involve non-corner (i.e., mid) vowels. As a result of the reduction process, mid vowels are eliminated and usually a set of corner vowels ([i, u, a]) remains. On the one hand, these vowels are advantageous from a dispersion point of view; vowel reduction to corner vowels enhances distinctiveness and reduces the possibility of misperception. On the other hand, the acoustic and articulatory characteristics of these vowels also indicate their special status.

Other vowel reduction patterns disprefer certain vowels (especially [a]) and reduce them to schwa in unstressed positions. As Crosswhite notes, "prominence reducing vowel reduction is based on the desire to avoid particularly long or otherwise salient vowel qualities in unstressed positions" (2004: 208). She posits a scale of vocalic prominence (a>ɛ,ɔ>e,o>i,u>ə), and argues that sounds higher on this scale can appear in prominent (stressed) positions, while sounds occupying lower positions can only appear in non-prominent (unstressed) positions. However, the exact phonetic bases of the scale are still debated. Moreover, it has to be noted that such prototypical reduction patterns are hard to find and the two types may not be easily separated.

8.1 Phonological vs. phonetic reduction

Two kinds of reduction may be distinguished: phonological (or lexical) reduction and phonetic (or acoustic) reduction. In the case of phonological reduction, the speaker intends to produce a reduced vowel, while in the case of phonetic reduction, the target is a full vowel (and reduction is the result of undershoot). Phonological vowel reduction raises the question as to whether the vowel is stored in its reduced or full form. In addition to these two possibilities, van Bergem (1994) introduces a combined approach

as well, in which variants of one and the same word are stored with full as well as reduced vowels. Van Bergem also notes that abstract analyses may not give a full explanation to phenomena as performance considerations, such as frequency and speaking style, are also influential in reduction phenomena.

8.2 Vowel reduction in English

In the next section, we will review the vowel inventory and stress patterns of English, then we will explore the process of vowel reduction, focussing our attention on the possible phonetic motivations of this process. We will consider the contribution of phonetics to the determination of phonological markedness.

8.2.1 The English vowel inventory

Standard British English vowels may be classified into monophthongs /ɪ, e, æ, ʌ, ʊ, ɒ, iː, uː, ɜː, ɔː, ɑː/ and diphthongs /aɪ, ɔɪ, eɪ, aʊ, oʊ, eə, ɪə, ʊə/. They may all appear in stressed syllables, therefore they can be described as full vowels. Unstressed syllables can only contain reduced vowels /ɪ, i, ʊ, u, ə/. A further classification of full vowels involves tense /iː, uː, ɔː, aɪ, ɔɪ, eɪ, aʊ, oʊ, eə, ɪə, ʊə/ and lax vowels /ɪ, e, æ, ʌ, ʊ, ɒ, ɜː, ɔː, ɑː/. While the traditional classification is based on the tenseness of vocal folds, it has been suggested that the difference lies in tongue root position. English vowels may also be classified according to the horizontal and vertical position of the tongue (shown in (10)). Rounding is not a distinctive feature in English; front and central vowel are unrounded, while back vowels are rounded, except for /ɑː/, which is unrounded.

(10)

	front	central	back
high	iː		uː
upper mid	ɪ, ɪə	ə	ʊ, ʊə, oʊ
lower mid	e, eɪ, eə	ɜː	ɔː, ɔɪ
low	æ	ʌ, aɪ, aʊ	ɒ, ɑː

English vowels may also be grouped into short and long vowels, but they do not constitute short-long pairs in the way Hungarian vowels do. The phonetic length of vowels is largely dependent on context: pre-fortis vowels have a shorter duration. In pairs such as *seat* [sit] – *sit* [sɪt] the difference is in quality rather than in quantity (Cruttenden 2001).

8.2.2 Stress patterns in English

Stress and reduction may not show a direct connection, nevertheless, from a phonological point of view, the appearance of reduced vowels is dependent on stress, as only unstressed syllables may contain reduced vowels. There is variation as to how many degrees of stress are distinguished in English. Some authors (e.g., Pater 1995) only distinguish three degrees: primary and secondary stressed syllables, and syllables with reduced vowels, while others differentiate primary stress, secondary stress, tertiary stress and zero stress (e.g., Nádasdy 2006). Primary, secondary, and tertiary stressed syllables contain full vowels, syllables with zero stress contain reduced vowels. In the latter approach, tertiary and zero stressed syllables are differentiated exclusively on the basis of the vowels they contain. If tertiary stress is not considered an independent

degree of stress, the differentiation between non-major stressed syllables that contain full vowels and reduced vowels becomes lexical.

As to the phonetics of stress, the most important articulatory correlate of stress is muscular effort. From the perception side, the perceived prominence of syllables is influenced by their loudness, duration, and pitch. In addition to these, Roach (1991) also mentions that in a sequence of identical vowels different vowel quality also contributes to a syllable's prominence. He highlights the fact that reduced vowels are statistically very frequent in English, which may provide a background against which full vowels seem prominent. The above mentioned factors are not equally significant in marking stressed syllables: pitch is the most essential, length appears to be the second most important, while loudness and quality play a less significant role.

There are several different approaches to stress assignment as well. Stress is weight sensitive in English and stress assignment follows certain rules; nonetheless, some authors consider the stress pattern of English words an arbitrary and lexical characteristic because of the great number of exceptions.

8.2.3 Vowel reduction patterns in English

The class of English reduced vowels consists of [ə], [ɪ], and [ʊ] (the tense pairs of [ɪ] and [ʊ] may also appear); reduced vowels can only appear in unstressed syllables by definition (e.g., *comma* [ˈkɒmə], *effect* [ɪˈfekt], *Portugal* [ˈpɔːʧʊgl]). Nevertheless, [ɪ] and [ʊ] also appear as full vowels in stressed syllables. The distribution of reduced vowels seems to be lexically determined. While several morphemes may show variation (e.g., *-ity* [ɪtɪ] or [ətɪ]), reduced vowels are not freely interchangeable.

Whether one considers reduced vowels to be the surface forms of underlying full vowels is a question of analysis. In a number of cases there is alternation evidence as to

which vowel can be considered the underlying full vowel (e.g., *academy* [əˈkædəmɪ] - *academic* [ˌækəˈdemɪk]). There are several words, however, which have only one form (such as *ago* [əˈgəʊ], *alone* [əˈləʊn]), or do not show a variable stress pattern, therefore, it is problematic to posit an underlying full vowel for them.

8.2.4 Contrasts

As we have seen in connection with vowel harmony, vowel space considerations and the existing contrasts in a language may also influence phonological patterns. The examination of the vowel inventory of English reveals that the articulatory characteristics of schwa are very similar or even identical to that of [ɜː], the main difference being the duration of the two sounds. As [ɜː] may appear in non-major stressed syllables, such as in *extrovert* [ˈekstrəvɜːt], in such positions [ɜː] and [ə] contrast only in length.

The question of whether schwa can contrast with full vowels depends on how many degrees of stress one presupposes. If we recognize primary, secondary, tertiary and zero stresses, it is possible to say that schwa does not contrast with full vowels, as they appear in different environments: full vowels appear in stressed syllables and reduced vowels in unstressed syllables. On the other hand, if one denies the existence of tertiary stress, then full as well as reduced vowels appear in non-major stressed syllables.

The phonetic basis of the distinction between tertiary and zero stressed syllables may only be that non-major stressed syllables with full vowels are phonetically more prominent than zero stressed syllables because of their longer duration. Nevertheless, duration plays an important role in determining stress and is apparently a strong cue in English, as it is enough to maintain the contrast of, for example, [ɜː] and [ə].

8.2.5 Motivation of vowel reduction

Analyses of English vowel reduction focus almost exclusively on the behaviour of schwa. Schwa is described as a mid central vowel with neutral lip position, but its quality is quite variable and depends on the context in which it appears. As the pronunciation of schwa is dependent on neighbouring sounds, the nature of weakening may not simply mean getting close to the centre of vowel space. Several experiments have provided evidence that vowel reduction may be seen as an assimilation to the segmental context (Flemming 2009).

Flemming argues that two kinds of schwa may be found in English: those that show centring (*mid central schwa*) and those that show assimilation to the context (*variable schwa*). Flemming notes that the two types "differ in their phonological patterning: mid central schwa usually minimally contrasts with higher vowel qualities (e.g. [i, u]), whereas variable schwa occurs primarily in contexts where all vowel quality contrasts can be neutralized" (Flemming 2009: 79). Flemming claims that it is the word-final position in which schwas can have a consistent vowel quality, whereas schwas in other positions show much more variability, especially in their F2 values (i.e., in their backness, depending on the neighbouring segments). Vowel reduction to schwa may be analysed as the neutralisation of F2 contrast: neutralisation of the second formant happens in contexts where it is too difficult to maintain distinct realisations of vowels. The context may be determined in terms of duration: when a vowel is very short (Flemming 2004).

Mid central schwa and variable schwa are claimed to arise from the same process, they only differ in the *degree* of assimilation to the segmental context. Reduction is induced by short duration: as the articulators have to move fast, there is no time for distinct realisation of vowel qualities. The two types of schwa may show

different realisations because they differ in duration, and shorter vowels induce more assimilation to the context. In their examination of reduced vowels in American English, Flemming and Johnson (2007) measured the length of schwas in different positions, and they found that the average duration of non-final schwas is much shorter (64 ms) than those in word-final position (153 ms). This may be partially attributed to phrase-final lengthening effects, nevertheless, the duration difference may still be significant.

Flemming (2009) observes that schwas show more variation in contexts where they do not contrast with other vowels. He explains this by suggesting that "there is no motivation to resist the pressure to assimilate to context if there is no need to realize vowel quality contrasts" (Flemming 2009: 91). Word-final schwas, on the other hand, contrast with other vowels and are less variable.

While Flemming (2009) calls attention to the articulatory aspects of vowel reduction, Burzio (2007) takes a perceptual approach. Burzio emphasises that besides the short duration of English unstressed vowels, they also have low energy levels. Low energy level endangers the perceptibility of unstressed vowels, and the motivation behind vowel reduction is that "perceptually ineffective articulatory effort is suppressed" (Burzio 2007: 156). Burzio claims that neutralisation happens in favour of the unmarked segment (i.e., schwa). Vowel reduction also influences the perceptual cues of the following consonant; Burzio notes that vowel reduction is possible only if the following consonant is coronal, which is considered the unmarked place specification.

8.2.6 Summary

Vowel reduction is restricted to unstressed positions in English. Stress may not be the direct motivating factor though, it has been suggested that it is shorter duration that induces vowel reduction. Among reduced vowels in English, schwa has been given the most attention. It has been described as a central vowel with a transitional nature, and its examination reveals that its articulation is different in positions where it contrasts with high vowels and where it is non-contrastive. Reduction involves the neutralisation of F2 contrast, and happens in positions where it is difficult to maintain distinct realisations of vowel qualities.

In English it is stress that seems to mark strong and weak positions. Stressed syllables may be considered strong or prominent as they provide better cues and allow more vowel quality contrasts, whereas unstressed syllables may be considered weak and may only show limited vowel contrasts.

Determining the unmarked element on the basis of vowel reduction seems to be a more straightforward case (than in the case of Hungarian vowel harmony) as the outcome of neutralisation is always the same. While there is certainly some lexical variation, the choice of the reduced vowel is not dependent on any other segment. In this case, neutralisation yields a mid central vowel, which may be seen as a targetless vowel whose articulation depends on the environment. Its production involves low articulatory effort as it simply provides a transition between the preceding and the following sound. Its configuration and transitional nature may contribute to its unmarked value. Its phonological behaviour as well as its phonetic characteristics are nevertheless dependent on stress and/or duration.

8.2.7 Suggestions for further research

Further research is needed on the correlates of stress as well as on how exactly stress, vowel length and vowel reduction are related. It is suggested that shorter duration is the result of the absence of stress and it entails reduction. The fact that low vowels have a longer duration (because more time is needed for the articulator movements), while the production of high vowels involves less time may contribute to their phonological patterning: high vowels may appear in unstressed positions but low vowels reduce to schwa in such positions.

Another area of research may involve the inspection of the phonetic characteristics of syllables with reduced vowels and syllables with tertiary stressed vowels. An examination as to whether the phonetic correlates of stress show significant differences, besides the apparent difference in vowel duration, may shed more light on the vowel reduction phenomenon in English.

As the examination of reduced vowels is usually restricted to schwa in English, the articulatory/acoustic characteristics of [ɪ] and [ʊ] could also be examined. These sounds have the widest distribution as they may appear in stressed as well as unstressed syllables. Their phonetic features should be compared in stressed and unstressed positions to see whether they show systematic differences.

9 Relating vowel harmony to vowel reduction

At first sight, vowel harmony and vowel reduction do not seem to be related. Vowel harmony is a word-domain process that involves the spreading of some feature of a vowel in a prominent syllable (in this case the first syllable of the root[13]), while vowel

[13] The trigger of Hungarian vowel harmony is usually the first syllable of the root, but in the case of neutral + back roots and disharmonic words, the trigger is the last (non-neutral) vowel. Anti-harmonic

reduction is a kind of weakening that is not dependent on other segments and happens automatically in unstressed syllables. Nonetheless, the two phenomena may be connected to the same underlying motivation. What seems to connect the two processes is that the targets of vowel harmony and the targets of vowel reduction may both be considered weak or dependent positions. These are disadvantageous positions in which the quality of the vowel is (at least partially) predictable.

In the case of vowel harmony, the first syllable of a root may be considered a strong or prominent position for two reasons. On the one hand, any vowel may appear in this position, and on the other hand, it restricts the appearance of certain features in the following syllables: dependent vowels get their values from the vowel in strong position. From a functional perspective, it is in word-initial position and is stressed, which results in a louder and possibly a longer vowel. A vowel in a perceptually good position then extends its features to the following vowels which are less salient.

In English, syllables that contain a full vowel may be considered strong positions. These positions allow for a greater range of vowels (though contextual restrictions may apply). Unstressed positions may be considered weak as only a limited set of reduced vowels may appear there. The phonetic grounding of weak positions is the absence of stress, and consequently, shorter duration.

There are different ways to help the identification of vowels in weak positions; one may aid these positions by restricting the set of possibly occurring vowels from which one has to choose (as is done in the case of vowel reduction). In this case, segments in weak positions are given default values that do not depend on the context. Another option, adopted by vowel harmony systems, is to transfer the contrast to a more perceptible vowel. While the set of possible vowels in a weak position is greater, the

---

words also present a difficulty from this respect, but the back quality of the suffix cannot be explained on the basis of the root vowels.

dependent segments' values are derivable from the vowel in strong position. The two phenomena do not provide clear cases as Hungarian vowel harmony is not of the vowel copy harmony type: certain contrasts are suspended in weak positions (backness, rounding) but others are still maintained (height). Apparently, vowel reduction in English does not provide a prototypical case either, as schwa is not the only reduced vowel and there is some lexical restriction on which of the three reduced vowels may appear in weak positions.

It is important to note that the configuration of vowels appearing in weak positions also conforms to the ease of articulation principle: unstressed vowels have a shorter duration, in which case more articulatory effort is required to complete the same movement, therefore, configurations which require little movement are preferred. In the case of harmonic vowels there is no need to make great adjustments of articulators, and in the case of vowel reduction, reduced vowels have a neutral articulation and a transitional nature that reduces articulatory effort.

10 Conclusion

As we hope to have shown, markedness is a multifaceted issue that may be linked to several elements and manifests itself in widely different ways. The concept of markedness is controversial, as it has gained different theory-dependent meanings and may not provide a definite and satisfactory explanation to phenomena. We have reviewed the relevant criteria that may possibly determine markedness and discussed a related problem that the source of markedness is often neglected or equated with the criteria that is used to determine it. In this way, markedness may not provide deeper insight into the phonological behaviour of elements.

We have shown that phonetics is a significant factor in determining phonological patterns. With linking markedness patterns to phonetics, one can explain cross-linguistically similar patterns with the help of general principles, the explanatory power of which rests on independent functional grounds. In this sense, then, phonetically based markedness may be considered universal, as the same general phonetic principles apply in all languages (language-specific aspects of articulation and perception are attested though).

We have exemplified the influence of phonetics with two markedness-related phenomena: Hungarian vowel harmony and English vowel reduction. These two phenomena show different kinds of contrast neutralisation, which happens in weak positions. The identification of these positions is possible (partly) on the basis of phonetics as these processes show the influence of articulatory and perceptual principles. Nevertheless, phonetic principles may not be the exclusive determiners of markedness observed at the phonological level, since phonological patterns are not fully predictable from phonetic principles.

The question of including phonetic principles in phonological theory has been raised in section 6. Phonetics as a source of markedness poses certain difficulties: it is not directly reflected in phonology: the effect of phonetic principles may emerge in certain cases, while it may fail to do so in others where it would be expected. As phonology generalises phonetically variable patterns, tracing back everything to phonetics seems to be impossible. However, an at least partially phonetically motivated approach may provide satisfactory explanation for markedness phenomena. The incorporation and systematic use of phonetic principles may provide an independent and empirically testable source for patterns, and may compensate for the "make-it-up-as-you-go-along" nature of phonology (Ohala 1990: 167).

References

Battistella, Edwin L. (1996) *The Logic of Markedness*. Oxford: Oxford University Press.

Berent, Iris & Tracy Lennertz & Paul Smolensky (2011) Syllable markedness and misperception: It's a two-way street. In C. E. Cairns & E. Raimy (eds.) *Handbook of the Syllable*. Leiden: Brill. 373-394.

Blevins, Juliette (2006) A theoretical synopsis of Evolutionary Phonology. *Theoretical Linguistics* 32. 117-166.

Brasington, R. W. P. (1982) Markedness, strength and position. In David Crystal (ed.) *Linguistic controversies*. London: Arnold. 81-94.

Buckley, Eugene (2000) What should phonology explain? Handout from SUNY Buffalo Linguistics Colloquium. Available from http://www.ling.upenn.edu/~gene/papers/buffalo.pdf.

Burzio, Luigi (2007) Phonology and phonetics of English stress and vowel reduction. *Language Sciences* 29. 154-176.

Bybee, Joan (2001) *Phonology and Language Use*. Cambridge: Cambridge University Press.

Carey, M. (2005) Interlanguage Phonology: Sources of L2 Pronunciation "Errors". Available from http://clas.mq.edu.au/phonetics/phonology/interlanguage/pronerrors.html.

Chitoran, Ioana & Abigail C. Cohn (2009) Complexity in phonetics and phonology: gradience, categoriality, and naturalness. In François Pellegrino, Egidio Marsico, Ioana Chitoran, and Christophe Coupé (eds.) *Approaches to phonological complexity*. Berlin and New York: Mouton de Gruyter. 21-46.

Chomsky, Noam (1965) *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam (1966) *Cartesian Linguistics: a chapter in the history of rationalist thought*. New York: Harper & Row.

Chomsky, Noam & Morris Halle (1968) *The Sound Pattern of English*. New York: Harper & Row.

Coupé, Christophe & Egidio Marsico & Francois Pellegrino (2009) Structural complexity of phonological systems. In François Pellegrino, Egidio Marsico, Ioana Chitoran, and Christophe Coupé (eds.) *Approaches to phonological complexity.* Berlin and New York: Mouton de Gruyter. 141-170.

Crosswhite, Katherine M. (2004) Vowel reduction. In Bruce Hayes, Robert Kirchner, and Donca Steriade (eds.) *Phonetically Based Phonology*. Cambridge: Cambridge University Press. 191-231.

Cruttenden, Alan. (2001) *Gimson's Pronunciation of English*. London: Arnold.

de Lacy, Paul (2006) Transmissibility and the role of the phonological component. *Theoretical Linguistics* 32 (2). 185-196.

Fikkert, Paula (2000) Acquisition of Phonology. In Lisa Cheng and Rint Sybesma (eds.) *The first GLOT International state-of-the-article book: the latest in linguistics.* Berlin/New York: Mouton de Gruyter. 221-250.

Finley, Sara & William Badecker (2008) Analytic biases for vowel harmony languages. In Natasha Abner and Jason Bishop (eds.) *Proceedings of the 27th West Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla Proceedings Project. 168-176.

Flemming, Edward (2004) Contrast and perceptual distinctiveness. In Bruce Hayes, Robert Kirchner, and Donca Steriade (eds.) *Phonetically Based Phonology*. Cambridge: Cambridge University Press. 232-276.

Flemming, Edward (2005) A phonetically-based model of phonological vowel reduction. Manuscript, Massachusetts Institute of Technology.

Flemming, Edward (2009) The phonetics of schwa vowels. In Donka Minkova (ed.) *Phonological Weakness in English: From Old to Present-day English.* Basingstoke: Palgrave Macmillan. 78-98.

Flemming, Edward & Stephanie Johnson (2007). Rosa's roses: reduced vowels in American English. *Journal of the International Phonetic Association* 37. 83-96.

Frisch, Stefan A. (2004) Language processing and segmental OCP effects. In Bruce Hayes, Robert Kirchner, and Donca Steriade (eds.) *Phonetically Based Phonology*. Cambridge: Cambridge University Press. 346-371.

Gósy, Mária (2004) *Fonetika, a beszéd tudománya*. Budapest: Osiris Kiadó.

Goto, Hiromu (1971) Auditory perception by normal Japanese adults of the sounds "l" and "r". *Neuropsychologia* 9 (3). 317–323.

Grimes, Stephen (2007) Word final consonant extrametricality in Hungarian. Manuscript, Indiana University.

Gurevich, Naomi (2001) A critique of markedness-based theories in phonology. *Studies in the Linguistic Sciences* 31 (2). 89-114.

Hale, Mark & Charles Reiss (2008) *The Phonological Enterprise*. Oxford: Oxford University Press.

Haspelmath, Martin (2006) Against markedness (and what to replace it with). *Journal of Linguistics* 42 (1). 25-70.

Hayes, Bruce & Donca Steriade (2004) Introduction: The Phonetic Bases of Phonological Markedness. In Bruce Hayes, Robert Kirchner, and Donca

Steriade (eds.) *Phonetically Based Phonology*. Cambridge: Cambridge University Press. 1-33.

Hayes, Bruce & Zsuzsa Cziráky Londe (2006) Stochastic phonological knowledge: the case of Hungarian vowel harmony. *Phonology* 23. 59-104.

Hume, Elizabeth (2003) Language Specific Markedness: The Case of Place of Articulation. *Studies in Phonetics, Phonology and Morphology* 9 (2). 295-310.

Hume, Elizabeth (2004) Deconstructing Markedness: A Predictability-based Approach. *Proceedings of the Berkeley Linguistics Society* 30. 182-198.

Hume, Elizabeth (2011) Markedness. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice (eds.) *The Blackwell Companion to Phonology*. Chichester: Wiley-Blackwell. 79-106.

Jakobson, Roman (1941) *Kindersprache, Aphasie und allgemeine Lautgesetze.* Språkvetenskapliga Sällskapets i Uppsala Förhandlingar, Uppsala: Uppsala Universitetet. [English translation: (1968) Child Language, Aphasia and Phonological Universals. In *Selected Writings I.* The Hague & Paris: Mouton. 328-401.]

Joos, Martin (ed.) (1957) *Readings in Linguistics I*. Chicago: University of Chicago Press.

Jun, Jongho (2004) Place assimilation. In Bruce Hayes, Robert Kirchner, and Donca Steriade (eds.) *Phonetically Based Phonology*. Cambridge: Cambridge University Press. 58-86.

Kálmán, László & Ádám Nádasdy (1994) A hangsúly. In Ferenc Kiefer (ed.) *Strukturális magyar nyelvtan 2: Fonológia*. Budapest: Akadémiai Kiadó. 393-549.

Kassai, Ilona (1994) A fonetikai háttér. In Ferenc Kiefer (ed.) *Strukturális magyar nyelvtan 2: Fonológia*. Budapest: Akadémiai Kiadó. 581-666.

Kaun, Abigail R. (2004) The typology of rounding harmony. In Bruce Hayes, Robert Kirchner, and Donca Steriade (eds.) *Phonetically Based Phonology*. Cambridge: Cambridge University Press. 87-116.

Kaye, Jonathan D. & Jean Lowenstamm (1981) Syllable Structure and Markedness Theory. In Adriana Belletti, Luciana Brandi, and Luigi Rizzi (eds.) *Theory of markedness in Generative Grammar*. Pisa: Scuola Normale Superiore. 287-315.

Kean, Mary-Louise (1975) *The Theory of Markedness in Generative Grammar*. Doctoral dissertation, Massachusetts Institute of Technology.

Kiparsky, Paul (2008) Universals constrain change; change results in typological generalizations. In Jeff Good (ed.) *Language universals and language change*. Oxford: Oxford University Press. 24-53.

Kiparsky, Paul & Karl Pajusalu (2003) Towards a Typology of Disharmony. *The Linguistic Review* 20. 217-241.

Liljencrants, Johan & Bjorn Lindblom (1972) Numerical Simulation of Vowel Quality Systems: The Role of Perceptual Contrast. *Language* 48 (4). 839-862.

Linebaugh, Gary Dean (2007) *Phonetic grounding and phonology: vowel backness harmony and vowel height harmony.* Doctoral dissertation, University of Illinois.

Lloret, Maria-Rosa (2007) On the nature of vowel harmony: spreading with a purpose. In Antonietta Bisetto and Francesco Barbieri (eds.) *Proceedings of the XXXIII Incontro di Grammatica Generativa*. Bologna: Universitá de Bologna. 15-35.

Maddieson, Ian (2009) Calculating phonological complexity. In François Pellegrino, Egidio Marsico, Ioana Chitoran, and Christophe Coupé (eds.) *Approaches to phonological complexity*. Berlin and New York: Mouton de Gruyter. 85-110.

Mády, Katalin (2008) Magyar magánhangzók vizsgálata elektromágneses artikulográffal normál és gyors beszédben. *Beszédkutatás* 2008. 52-66.

Mády, Katalin & Uwe D. Reichel (2007) Quantity distinction in the Hungarian vowel system – just theory or also reality? In J. Trouvain and W. J. Barry (eds.) *Proceedings of the 16th International Congress of Phonetic Sciences*. Dudweiler: Pirrot. 1053–1056.

Meunier, Christine & Cheryl Frenck-Mestre & Taissia Lelekov-Boissard & Martine Le Besnerais (2003) Production and perception of vowels: does the density of the system play a role? In M. Solé, D. Recasens and J. Romero (eds.) *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona: Casual Productions. 723–726.

Nádasdy, Ádám (2006) *Background to English Pronunciation*. Budapest: Nemzeti Tankönyvkiadó.

Newmayer, Frederick J. (2005) *Possible and Probable Languages: A Generative Perspective on Linguistic Typology*. Oxford: Oxford University Press.

Ohala, John J. (1990) There is no interface between phonology and phonetics: a personal view. *Journal of Phonetics* 18. 153-171.

Pater, Joe (1995) On the nonuniformity of weight-to-stress and stress preservation effects in English. Manuscript, McGill University.

Pierrehumbert, Janet B. (2001) Exemplar dynamics: Word frequency, lenition and contrast. In: J. Bybee and P. Hopper (eds.) *Frequency effects and the emergence of linguistic structure.* Amsterdam: John Benjamins. 137-157.

Prince, Allan & Paul Smolensky (1993) Optimality Theory: Constraint Interaction in Generative Grammar. Manuscript, Rutgers University and University of Colorado.

Rice, Keren (1999) Featural markedness in phonology: Variation. Part 1. *Glot International* 4 (7). 3–6. Part 2. *Glot International* 4 (8). 3–7.

Rice, Keren (2007) Markedness in phonology. In Paul de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press. 79-98.

Roach, Peter (1991) *English Phonetics and Phonology: A Practical Course*. 2nd edition. Cambridge: Cambridge University Press.

Rose, Sharon & Rachel Walker (2011) Harmony Systems. In John Goldsmith, Jason Riggle and Alan Yu (eds.) *Handbook of Phonological Theory*. Cambridge, MA: Blackwell. 240-290.

Siptár, Péter & Miklós Törkenczy (2000) *The Phonology of Hungarian*. Oxford: Oxford University Press.

Steriade, Donca (1995) Underspecification and Markedness. In John A. Goldsmith (ed.) *The Handbook of Phonological Theory*. Oxford: Blackwell. 114-174.

Steriade, Donca (1997) Phonetics in phonology: the case of laryngeal neutralisation. Manuscript, University of California.

Steriade, Donca (2001) Directional Asymmetries in Place Assimilation: A Perceptual Account. In Elizabeth Hume & Keith Johnson (eds.) *The role of speech perception in phonology*. New York: Academic Press. 219-249.

Törkenczy, Miklós (2011) Hungarian Vowel Harmony. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume and Keren Rice (eds.) *The Blackwell Companion to Phonology*. Malden, MA & Oxford: Wiley-Blackwell. 2963–2990.

van Bergem, Dick R. (1994) Reflections on aspects of vowel reduction. *Proceedings of the Institute of Phonetic Sciences* 18. 95-110.

van de Weijer, J. (2009) Optimality Theory and Exemplar Theory. In: The Phonological Society of Japan (ed.) *Phonological Studies 12*. Kaitakusha, Tokyo. 117-124.

Walker, Rachel (2011) *Vowel patterns in language*. Cambridge: Cambridge University Press.

White, Laurence & Katalin Mády (2008) The long and the short and the final: Phonological vowel length and prosodic timing in Hungarian. In Plinio A. Barbosa, Sandra Madureira and César Reis (eds.) *Proceedings of the Speech Prosody 2008 Conference*. 363-366.

Wright, Richard (2004) A review of perceptual cues ad cue robustness. In Bruce Hayes, Robert Kirchner, and Donca Steriade (eds.) *Phonetically Based Phonology*. Cambridge: Cambridge University Press. 34-57.