ON HUNGARIAN MORPHOLOGY

András Kornai

Abstract

The aim of this study is to provide an autosegmental description of Hungarian morphology. Chapter 1 sketches the (meta)theoretical background and summarizes the main argument. In Chapter 2 phonological prerequisites to morphological analysis are discussed. Special attention is paid to Hungarian vowel harmony. In Chapter 3 a universal theory of lexical categories is proposed, and the category system of Hungarian is described within it. The final chapter presents a detailed description of nominal and verbal inflection in Hungarian, and describes the main features of a computer implementation based on the analyses provided here.

0 Preface

This thesis was written in 1984-1986 – the first publicly circulated version (Version 1.4) was defended at the Hungarian Academy of Sciences (HAS) Institute of Linguistics in September 1986. An extended Version 2 was submitted to the HAS Scientific Qualifications Committee in August 1988, and was formally defended in September 1989. Version 2.1, the 1994 paperback edition in *Linguistica: Studia et Dissertationes* (ISBN 963 8461 73 X) was different from Version 2 in two main respects: an English version of the official Summary Kandidátusi Értekezés Tézisei replaced the old Introduction, and the section on implementation reflected the system that was running at Xerox PARC at the time of the formal defense. The current Version 3.0 (pdf only) is based on the source of the paper version with minimal reformatting (the addition of hyperrefs, and a slight revision of this preface – salient changes *italicized*).

To leave the material in the 1986 thesis in manuscript format until 1994 would have required more patience than the author can lay claim to. The abstract analysis of feature systems in chapter 2.1 forms the basis of the more detailed analysis in Kornai 1993. The analyses of Hungarian vowel harmony in 2.1 and 2.6 have been published in Kornai 1987 and Kornai 1991. Part of the material on syllable structure in 2.4 can now be found in Kornai 1990, and a more extended version of the material in 2.5 on postlexical phonology is in Kornai and Kálmán 1989. The analysis of lexical categories in 3.1 and 3.2 has been published as Kornai 1985. Except for chapter 4, which describes the morphological rules in detail and discusses a two-level implementation, almost all the material has been published elsewhere *by 1994*. What was the purpose of the paper edition in 1994 and what is the purpose of the present (2007) pdf version?

In 1994 it was hoped that by bringing the parts together in a convenient monographic format the reader could use the volume as a reference work detailing nearly all aspects of the inflectional and many aspects of the derivational morphology of Hungarian. *In this regard, the book has been clearly superseded by* Siptár and Törkenczy 2000. There are a number of places where references to this volume and to later developments could be incorporated, especially in the treatment of underspecification (see Steriade 1995), multiple specification (see Lieber 1987) and feature geometry (see Clements and Hume 1995). In general, I restricted myself to updating the references by citing the published version of works that I originally consulted in manuscript form.[1]

Since the rule system developed for the thesis has been extensively tested on over twenty thousand stems covering virtually the whole lexicon of contemporary Hungarian, it was also hoped that future models of Hungarian morphology can, if nothing else, learn from the mistakes of the 1989 model presented in the 1994 edition. *This hope was certainly fulfilled:* the underlying database has become part of the bloodstream of research in Hungarian phonology and morphology, providing one of the three main sources for the current morphdb.hu, the other two being the Dictionary of Hungarian Inflections by

[1] In this process I had to trade in philological accuracy for usability: the sections quoted from the mss may no longer be present verbatim in the published works, but the latter are generally very close to the former in spirit, and are, unlike the mss, easily accessible to the reader.

László Elekfi and László Németh's Hungarian ispell dictionary. Since a fair amount of the material here serves as motivation for later developments such as the annotation scheme used in the morphdb database, linguistically oriented users of this resource may still find Version 3.0 a relatively painless introduction to the subject.

It was also hoped that the overarching plan of this work, as presented in chapter 1, might still have some relevance for theoretical linguistics. In a fast moving field like generative grammar authors seldom get a chance to contemplate whether the topics they chose to investigate and the methods they applied a decade *(now two decades)* ago would still make sense. In the eighties the author was lucky enough to pick a problem domain, morphology, that was just about to come in from the cold, and a technical framework, autosegmental theory, that has since grown from an esoteric branch of Africanist phonology into a mainstay of generative grammar. In the nineties some of the fundamental judgement calls made in this work, in particular the unashamedly procedural mode of description and the unbridled minimalism/reductionism that are the common threads binding it all together, came under renewed attack. Whether the work stands up against these lines of criticism was supposed to be decided by the 1994 reader. *In 2007 the author feels free to draw his own conclusions:* proceduralism is dead but rule systems are alive and well. Eighties phonology was not particularly concerned with the psychological reality of intermediate stages of derivations, unless they were the output of a specific lexical level, and contemporary phonology shows the same indifference to Gen candidates thrown out by Eval, unless they matter for sympathy. (Actually, the two issues as strongly related, see McCarthy 2000, Kiparsky 2007.)

0.1 Textual conventions

Examples, rules, figures, and formulas are numbered consecutively in each chapter: for instance, (2:18) refers to item (18) in chapter 2. Given the availability of hyperref, the footnote/endnote distinction of Version 2.1 is no longer maintained. When there is no danger of misunderstanding, Hungarian forms are given orthographically. Glosses appear in single quotes, and phonemic transcriptions (in cases where the spelling is insufficient) in slashes. In a few cases, phonetic transcriptions are given in square brackets. Angled brackets are used with features exclusively: tree structures are often given as bracketings.

0.2 Acknowledgments

This work owes its existence to a grant of the Hungarian Academy of Sciences (HAS) that enabled me to study morphology at the HAS Institute of Linguistics (NYTI) while on leave from the HAS Computer Science and Automation Institute (SZTAKI). I would like to thank Előd Knuth†, then head of the Information Systems Department, and Tibor Vámos, then director of SZTAKI, for their generosity in providing access to the computer facilities of SZTAKI — this work would have been impossible without their support. I am deeply indebted to many linguists for their suggestions, comments, and criticisms over the years. My former colleagues at the NYTI, and in particular the erstwhile collective of Room 13, are to be thanked *in toto*.

Improvements from Version 1 to 1.4 were due largely to László Kálmán (NYTI), Ádám Nádasdy (Eötvös Loránd University), and Péter Siptár (NYTI). Version 2 benefited a great deal from the opinions of my opponents at the 1986 defense, Ferenc Papp†, and Péter Siptár (NYTI) as well as from the comments and criticisms submitted in writing or made in person at the defense by József Hermann†, István Kenesei (University of Szeged), Ádám Nádasdy (Eötvös Loránd University), and Péter Siptár (NYTI), who went over so many versions that mentioning his name a third time is still less than adequate.

Version 2.1 also benefited from the opinions of my opponents at the formal defense, Ferenc Papp, and György Szépe (NYTI), as well as the comments and criticisms of those on the committee: László Elekfi (NYTI), György Hell (Budapest Institute of Technology), László Hunyadi (University of Debrecen), István Kenesei (University of Szeged), Katalin É. Kiss (NYTI), Mihály Péter (Eötvös Loránd University), and László Pordány (University of Szeged). Kari Swingle (UCSC) is to blame for the errors that still remain.

Extending the system from 200 to 20,000 stems would have been impossible without the SZOTA1R database. I am indebted to many people who contributed to SZOTA1R, especially to the late Ferenc Papp, whose *Debrecen Thesaurus* provided the majority of stems in SZOTA1R, and Mihály Füredi, whose *Frequency Dictionary* was extensively used in testing and refining KIMMO1. I would like to thank Lauri Karttunen, whose TWOL system provided the basis of the 1989 implementation. Without his patient instruction and untiring support, I would still be working on Hungarian morphology with paper and pencil.

The 1985 lectures of Nick Clements on syllable structure at the Salzburg International Summer School have had a decisive influence on my treatment of the material in 2.4. It is a pleasure to acknowledge my indebtedness. The ideas contained in Appendix 2.6 and in particular the treatment of ternary harmony are the result of joint work with Donca Steriade (MIT) and Harry van der Hulst (University of Amsterdam). Errors and omissions are my own.

Finally, I would like to thank my advisor, Ferenc Kiefer, for his constant help and encouragement.

1 Introduction

Aside from a few isolated attempts, such as Kiefer 1970, morphology played a very limited role in the early development of generative grammar. The basic reason for this was that the standard generative model (Chomsky 1965) treated sentences as strings of morphemes: both base (rewriting) and transformational rules operated on morphemes. This one-step model (originating in the work of Harris 1946, 1951) was gradually replaced by a two-step model in which sentences are treated as strings of words, and words are treated as strings of morphemes, much as in traditional grammar.

In order to (re)introduce 'word' as an explanatory category, the class of possible models had to be delimited so that the division of labor between rules of syntax, on the one hand, and rules of phonology/morphology on the other, become clear. This is accomplished by the Lexicalist Hypothesis (Chomsky 1970), now usually called the Lexical Integrity Hypothesis (LIH). Although the LIH exists in many versions (see Scalise 1985), for our purposes it will be sufficient to state the following basic requirements, which are common to nearly all versions of the LIH:

(1) Rules of syntax (and semantics) cannot make reference to the phonological content of words.

(2) Rules of syntax (and semantics) cannot modify the phonological and morphosyntactic features of words.

Thus, in addition to forbidding rules like Affix Hopping (Chomsky 1957), the LIH also forbids the derivation of word-forms by syntactic rules. But if syntax can not derive word-forms, each and every word-form must be supplied by the lexicon. According to the traditional view of the lexicon as a list, this would mean that every (paradigmatic) form of a word must be listed. Given that such forms often number in the thousands, listing them all appears to be impractical, if not impossible.

But the traditional file-card based technology of lexicography has gradually been replaced by a computer-based technology that can handle several orders of magnitude more data, and the reason why generative morphology avoids listing all word-forms is not a practical but rather a theoretical one. It is the *Principle of Brevity*, stated by Chomsky and Halle (1968:12) as follows:

> "Regular variations (...) are not matters for the lexicon, which should contain only idiosyncratic properties of items, properties not predictable by general rule."

Listing all word-forms (and in particular, all paradigmatic forms) thus contradicts the Principle of Brevity, while the LIH, apparently, requires exactly this.

Generative morphology resolves this contradiction by treating the lexicon not as a static list but as a dynamic (generative) component of the grammar. The LIH requires only that syntax must get fully formed words from the lexicon, but does not require that such forms be stored in the lexicon. On the contrary, the Principle of Brevity demands that no form that can be produced by a regular operation should be stored. Thus we must distinguish the output and the content of the lexicon. The output of the lexicon is the set of well-formed words. If we treat compounding or recursive derivational processes as productive, this will be an infinite set, which makes it impossible to think of the lexicon as a list. The content of the lexicon, however, will be reduced to the list of irreducible elements (morphemes) and the rules operating on them (suppletive and other irregular forms are also included here).

Therefore, the fundamental goal of generative morphology is to characterize the basic elements and operations of the lexicon — the contributions of the present dissertation to this goal are discussed in 1.2 – 1.4.

1.1 The methods of the investigation

We have seen above that the LIH and the Principle of Brevity, taken together, will naturally lead to a generative view of the lexicon. Given that traditional grammar makes as sharp a division between phonology and morphology as is made by the LIH between morphology and syntax, it seems possible to gain a better understanding of the structure of the lexical component by enforcing more principles of separation than just the LIH. This is the method of 'natural' generative phonology, where the most important principle of separation is the Morphophonemic-Allophonic Principle introduced by Koutsoudas et al. 1973. More recently, research in 'natural' morphology (Dressler 1985) attempts to isolate a third, morphophonological component between phonology and morphology.

The basic method of the present investigation is exactly the opposite of the strategy of natural phonology/morphology. My fundamental assumption is that phonology and morphology form an indivisible unit which I will call 'the lexicon' or just 'morphology'. This assumption seems to be contradicted not only by the obvious difference in the size of the basic units (phonemes vs. morphemes) but also by the different nature of the typical phonological operations (e.g. assimilation or deletion) and the typical morphological operations (e.g. affixation or compounding). Therefore it is necessary to discuss briefly why 'mainstream' generative grammar treats phonology and morphology homogeneously.

Jakobson pointed out that we find a number of rules (such as the rule of word-final devoicing in Russian, see Halle 1959) that are morphophonemic and allophonic at the same time. With the introduction of context-sensitive rules (Chomsky 1956), the homogeneous formal treatment of phonological and morphological rules became possible and necessary, since there was no separate morphological component at the time. The principle of cyclic rule application made clear that each morphological rule (affixation) triggers phonological rules such as stress shift. Furthermore, Lexical Phonology (Kiparsky 1982) provided a unified treatment of cyclic rules and the phonological processes taking place at morpheme boundaries and showed that the rules of phonology and morphology are arranged in the same stratal structure.

The most important argument in favor of a homogeneous treatment of phonology and morphology was provided by the fundamental transformation of phonology in the last decade. This transformation started with the introduction of a separate tier for tone (Leben 1973, Goldsmith 1976). This was followed by tiers for harmonizing vowel features (Clements 1976), for aspiration (Thráinsson 1978), for nasalization (Hyman 1982), for syllabicity (Clements and Keyser 1983) and so on. The multi-tiered representations thus formed made it possible to treat the infixing morphology of Semitic languages by purely phonological means (McCarthy 1979) and to reduce reduplication to concatenative affixation

(Marantz 1982). With the aid of multi-tiered representations, other processes, hitherto assumed to be purely morphological, also became amenable to a treatment in terms of phonologically motivated operations.

The methodological basis of the present dissertation is the principle of parsimony (Occam's razor). Thus, when we seek a characterization of the basic elements and operations of the lexicon, we seek an answer to the following questions. What are the fundamental units that *must* be stored in the lexicon? What are the operations which are *indispensable* for the task of generating every word-form? Starting from the phonological form of words we can argue that the words can be decomposed into syllables, the syllables can be decomposed into phonemes, and the phonemes can be decomposed into distinctive features. The distinctive features are atomic, and thus will necessarily be part of the lexicon. Starting from the meaning of words, the minimal units having both phonological and semantic content are (by definition) the morphemes, so the lexicon will have to contain these too.

However, the question whether morphemes are built from phonemes or directly from features is arguably open: for instance, in Hungarian, backness is a property not of the individual vowels, but of the whole morpheme (see Hetzron 1972). Autosegmental phonology expresses this fact by locating backness on a separate tier. For instance, the segmental content of *apa* 'father' and *epe* 'bile' will be *ApA* in both cases, where *A* is the same *a/e* archiphoneme that we find in the dative suffix *nak/nek*. In the full representation of these morphemes, both the segmental and the backness tiers are present, together with the association lines between them:

```
 +B          -B
 / \         / \
 ApA         ApA        nAk

(apa)       (epe)      (DAT)
```

As can be seen, the dative morpheme is not associated to the +B or –B feature — the basic rule of vowel harmony says that the *A* of *nAk* has to be associated to the backness feature of the stem. Therefore association (and also delinking) will have to be listed among the fundamental operations, the more so because they appear in the description of other phonological processes (such as assimilation or compensatory lengthening) as well. Similarly, a theory of the lexicon can not do without the operations of feature insertion and deletion. The use of these operations and of concatenation has been extensively justified in autosegmental phonology already in the description of tone languages, so the present dissertation could take these to be given.

For the reasons sketched above, I could not take it for granted that phonemes are primitives. With the introduction of a root tier, autosegmental phonology made it possible to identify a phoneme with the set of features associated to a root node, and to treat the (traditionally problematic) affricates and diphthongs as well as the long (geminate) phonemes in a homogeneous manner together with short phonemes (Clements and Keyser 1983). Accordingly, I did not permit rules that operate on phonemes rather than directly on features. With the introduction of a morphemic tier (McCarthy 1979), parentheses, and in general segmentoid boundary markers also become eliminable (for the syllable boundary see Kahn 1976, for boundary markers of various strength see Mohanan 1984).

The central aim of the dissertation is to show that Hungarian morphology can be described with the extremely limited inventory of representations and operations outlined so far. Since the most powerful tools of the standard theory (such as transformations, curly brackets, and Greek letter variables) were already eliminated from autosegmental theory, the dissertation concentrates on the remaining two strongest tools, namely diacritic features and feature changing rules. Since my aim was to limit the number of tools available, the choice between alternative analyses was always dictated by the criterion of using the least number of *ad hoc* tools.

The method employed in the investigation of the meanings of words and morphemes was also dictated by considerations of parsimony. Although we do not know precisely what features distinguish the meaning of one morpheme (or word) from the meaning of another one (or, to put it differently, our decisions in these matters will greatly depend on the theory of lexical semantics we adopt), we can assume without further argumentation that the number of the ultimate features is finite (less than the number of elements that have to be listed in the lexicon). Since some operations that will insert features in larger structures (and further transform these structures) will be necessary for phonological purposes anyway, the dissertation employs the same operations in the investigation of meaning as well. The morphosyntactic features that transmit the information between syntax and morphology are also handled by these operations, thereby steering clear of the problem whether morphosyntactic features are to be treated as morphological (Kiparsky 1986), syntactic (Gazdar *et al* 1985), or semantic (Lapointe 1980).

1.2 Summary of new results

2.1 develops an algebraic treatment of phonological features which, through the investigation of the relation between phonological features and natural classes, provides a unified treatment of the original (Pāṇinian), the standard (SPE), and the modern (autosegmental) approaches. Here I will omit the algebraic details and provide a somewhat simplified model that employs only concepts from elementary set theory.

Let $P = \{p_1, p_2, ..., p_k\}$ be a set of phonemes and $F = \{f_1, f_2, ..., f_n\}$ be a set of binary features. The mapping $C: P \to 2^F$ will be a *feature analysis* if it satisfies the following criteria:

(3) Emicity. If $i \neq j$, then $C(p_i) \neq C(p_j)$.

(4) Compactness. If $N \subset P$ is a natural class, we can find a set of forbidden features T' and a set of required features B such that $p \in N$ iff $B \subset C(p) \subset F - T'$

In the dissertation I argue that features and phonemes can be treated uniformly in the algebraic model, and show that the standard theory of features makes two predictions concerning the set of natural classes:

(5) The number of natural classes is a (small) polynomial function of the number of phonemes ($k^{1.585}$)

(6) The set of natural classes is closed under intersection

I show that the Pāṇinian theory makes essentially the same predictions (with the function k^2 instead of $k^{1.585}$), and I develop a concept of 'natural class' for the autosegmental case that leaves (5) and (6) in force. The theory is illustrated on a feature analysis of Hungarian vowels based on the tridirectional features $\langle I, U, A\rangle$ and the model is interpreted with the aid of barycentric coordinates so that it relates phonological features to phonetic facts.

2.2 deals with the feature analysis of consonants by reformulating the traditional analysis in terms of the feature geometry suggested by Clements (1985). The investigation of consonant epenthesis undertaken here provides a new argument in favor of a proposal made by É. Kiss and Papp (1984) that in Hungarian *dz* should be treated as a sequence of two phonemes.

2.4 deals with syllable structure. Earlier results (chiefly Siptár 1979,1980) are discussed from the point of sonority, and the ID/LP model of constituency (Gazdar and Pullum 1982) is used to exclude final clusters such as *pj, kj* that appear only in inflected forms.

2.5 deals with postlexical rules. In addition to an outline of Hungarian consonant sandhi, the rules of Hungarian sentence intonation, jointly developed with László Kálmán, are also sketched.

1.3 Vowel harmony

The investigation of Hungarian vowel harmony traditionally concentrates on the binary alternations *a/e, á/é, o/ö, ó/ő, u/ü, ú/ű*, and mentions the ternary alternation *o/e/ö* only in passing. An important empirical result of the present dissertation is that it includes the quaternary alternation *a/e/o/ö* which, following Vágó 1975, is treated in the literature as if it were independent of the problem of vowel harmony. As we shall see, a unified treatment of binary, ternary, and quaternary alternation is justified not only because in the surface representation the alternant *a* (as in *házat, házak, házam*) plays the same role as the alternants *o,e,ö*, but also because this way suffix-combinations can be described by the same machinery used for single suffixes.

A fundamental difference between the earlier analyses and the one proposed here is that here *a* and *á*, as well as *e* and *é*, are treated as differing only in length. In order to show that there are no differences between *a* and *á* or *e* and *é* in (underlying) height or roundedness, I analyze two quantity-changing processes, Low Vowel Lengthening (*apa/apát, epe/epét*) and Stem Shortening (*tél/telet, nyár/nyarat, tűz/tüzet, víz/vizet, nyúl/nyulat*).

For the description of a seven vowel system at least three features must be used. A feature analysis using only three features can be reconstructed from the phonological pattern of Hungarian: the feature A separates the vowels of the quaternary archiphoneme from the rest, the feature I separates the two alternants of binary alternations, and the feature U will group the vowels according to ternary alternation.

If we also take the proximal/distal alternations *itt/ott, ez/az, így/úgy* into account, we arrive at a feature chart that is equivalent to a Jakobsonian analysis: I = ⟨–grave⟩, A = ⟨–diffuse⟩, U = ⟨+flat⟩.

The next step in the analysis is to show that the traditional palatal/velar as well as the more modern front/neutral/back stem classifications are not sufficiently detailed and that in fact we need five stem classes. The stems *had, bab, hit, hölgy,* and *tök* all have to be in different classes, since no two of these will get the same alternants for every (binary, ternary, and quaternary) suffix (cf *hadat, babot, hitet, tököt*, and *hithez, hölgyhöz*). Stems of the *had* and *hölgy* type (which are treated in Vágó (1980) with the aid of a 'Minor Lowering' rule triggered by a diacritic ML) require separate classes not only because they are numerous, but also because this treatment permits all stems (including derived ones) to be unambiguously classified. Although from a diachronic point of view, these classes are closed, in a synchronic description they must be treated as productive, since stems bearing plural (or possessive) suffixes all belong here (cf **babokot, *babomot, *tökököt* etc.).

The simplest possible harmony rules are:

I			U		
|	⋱		|	⋱	
V	C_0	V	V	C_0	V

How far can an analysis employing these rules be pushed? It would go against the Principle of Brevity to leave the stems selecting the *a*-alternant unmarked and mark the stems selecting the *o* alternant with a diacritic — we must suppose that for back stems the unmarked member of the opposition *a/e/o/ö* is *o*. Therefore the diacritic ML turns [+A,+U] into [+A] (in the *had* class) and turns [+A,+U,+I] into [+A,+I] (in the *hölgy* class). It follows from our basic aim of eliminating diacritics that the *ad hoc* ML has to be replaced by the phonologically motivated ⟨–U⟩.

Given the existence of such ML stems as *lyuk*, the only way to reconcile the +U feature of the stem vowel and the –U required by the exceptional harmonic behavior of the stem is to specify the same feature on different tiers. In other words, the parsimonious treatment of features results in the introduction of an extra tool such as the *core specification* (in the sense of Halle and Vergnaud 1983) that is used in the analysis given in 2.3. Thus, it seemed advisable to provide an alternative analysis based on four vowel quality features. This analysis, given in 2.7, is based on the standard features:

	a	e	i	o	u	ü	ö	a/e	o/ö	u/ü	o/e/ö	a/e/o/ö
back	+	-	-	+	+	-	-	0	0	0	0	0
high	-	-	+	-	+	+	-	-	-	+	-	-
low	+	+	-	-	-	-	-	+	-	-	-	0
round	0	0	-	+	+	+	+	0	+	+	0	0

The role of the diacritic ML is played by a (floating) ⟨+low⟩. It can be seen that the spreading of ⟨+low⟩ will narrow down the quaternary archiphoneme *a/e/o/ö* to the binary archiphoneme *a/e*. The rules of (privative) I-spread are replaced by a pair of (equipollent) spreading rules:

$$\begin{array}{ccc} \mathrm{B} & & \\ | & \ddots & \\ \mathrm{V} & \mathrm{C}_0 & \mathrm{V} \end{array} \qquad\qquad\qquad \begin{array}{ccc} \mathrm{F} & & \\ | & \ddots & \\ \mathrm{V} & \mathrm{C}_0 & \mathrm{V} \end{array}$$

and there is a separate rule for ternary harmony:

$$\begin{array}{lcll} \text{backness tier} & \mathrm{F} & & \\ & | & \ddots & \\ \text{CV tier} & \mathrm{V} & \mathrm{C}_0 & \mathrm{V}\langle -H \rangle \\ & | & \text{⋰} & \\ \text{rounding tier} & \mathrm{R} & & \end{array}$$

While the analysis based on three features had to use feature changing rules, this analysis employs only monotonic (feature adding) rules. As a result, the spreading of ⟨+low⟩ onto the ternary archiphoneme is blocked because *o/e/ö* is underlyingly specified as ⟨–low⟩. Since the standard analysis does not distinguish between the ternary and the quaternary archiphoneme, the Minor Lowering rule of Vágó 1980 would generate the incorrect form **hölgyhez* in such cases.

1.4 Summary of conclusions

Chapter 3 discusses the notions *word* and *lexical category*. The explanatory value of the notion *word* is argued to stem from the fact that several logically independent methods of segmentation yield essentially the same word-sized units. Lexical categories are defined morphologically which has the advantage of yielding a feature analysis (X-bar theory) of lexical categories as a byproduct of morphological analysis. The (morphosyntactic) features defining lexical categories are argued to have their own ‘geometry’. The basic restriction on the tree structures thus formed is that only “+” (marked) nodes can have daughters. The theory is illustrated on the category system of Hungarian, with special emphasis on the problem of defective paradigms.

4.1 and 4.2 provides a detailed description of the Hungarian verbal and nominal paradigms. The diacritic ML was eliminated in Chapter 2 — here two other diacritics, governing Stem Shortening (as in *nyár/nyarat*) and Vowel Drop (as in *cukor/cukrot*) are eliminated in favor of more motivated features. In order to describe the 52 verbal forms discussed, we need 26 morphemes (ordered by the Elsewhere Principle). This is more than twice as many as the 12 morphemes that would be necessary in a fully agglutinative system, but only half of the 52 that would be needed in a purely inflecting system. Roughly the same degree of agglutination is found in the nominal (possessive) paradigm.

In sum, the investigation of diacritics in Hungarian arguably supports the thesis that it is often the arrangement, and not the substance of the atomic units that gives rise to different behaviors, much like in the case of graphite and diamond. The investigation of feature-changing rules, however, does not seem to yield the same kind of ‘conservation laws’ that are common in the physical sciences. Although Hungarian vowel harmony is amenable to a feature-adding treatment, there are places both in the verbal paradigm (such as the 3rd singular present definite *ja/i*) and in the nominal paradigm (such as the plural possessive *i*) that require the use of feature-changing rules under any analysis of vowel harmony.

2 Phonology

The aim of this chapter is twofold: on the one hand, I will discuss some phonological preliminaries to the morphological analysis developed in subsequent chapters, such as the feature analysis of vowels (2.1), consonants (2.2), and a treatment of vowel harmony (2.3). On the other hand, I will show that the proposed framework gives us insights that go well beyond what is strictly necessary for morphological analysis. To that end, I will outline a treatment of some phonotactic constraints (2.4), and a brief discussion of postlexical phonology (2.5).

Given the preparatory nature of this chapter, it did not seem necessary to pursue all these developments to the same level of descriptive detail. Vowel harmony, being probably the most pervasive, and certainly the most researched, feature of Hungarian phonology, receives a detailed description and not one but two analyses are provided for it.[2] Consonantal rules are discussed only in order to exemplify the basic autosegmental mechanisms one would want to employ in a truly detailed description: on a number of occasions the details are left as an exercise to the reader in the hope that he will be interested in 'executing' the 'leading ideas' of autosegmental phonology on the wealth of material provided by Hungarian.

2.1 The feature system: vowels

The feature analysis of Hungarian is by no means an uncontroversial question. The solution offered here differs from more or less traditional proposals not only in the choice of underlying feature set, but also in the manner in which the feature composition of a given phoneme is interpreted. Since the proposal to be made here departs from the conventional *linear* mode of representation, as exemplified by SPE (Chomsky and Halle 1968:8.2) and exploits the possibilities of the *autosegmental* approach to a considerable degree, it is extremely hard to 'linearize' it in order to facilitate comparison with more standard solutions.

Therefore, it will be necessary to give a formal rendering of the 'feature analysis' problem in well-defined mathematical terms: this way it will be possible to expose the metatheory of features. The less mathematically minded reader should not worry, for the formalization will be a light one: the algebra and the geometry employed here is conveniently summarized in chapters 9 and 12 of Birkhoff and Bartee 1970, where the definitions of vectors, fields, linear spaces, and lattices are given in detail; and chapter 13 of Coxeter 1961, where barycentric coordinates are defined rigorously, or Steinhaus 1960:119, where they are treated more intuitively. (In fact, the mathematical rigor of these concepts is not essential in what comes later.) Moreover, each step will be illustrated on a concrete example, that of the Hungarian vowel system. The consonant system is discussed in 2.2.

In general, there is very little disagreement concerning the phoneme inventory — for Hungarian, it is almost universally accepted that the (surface) phonemes are the following:

[2] Since the second of these is not wholly the author's intellectual property, it is relegated to an Appendix to chapter 2.

(1A)

p	b			t	d	ty [c]	gy [ɟ]			k	g	
		f	v	sz[s]	z			s [š]	zs [ž]			
				c[t^s]	dz [d^s]			cs[č]	dzs [ǰ]			
	m				n		ny[ɲ]					
					r l							
						j [y]						h

(1B)

a [ɔ]	e [ε]	i	o	u	ü	ö
á [a:]	é [e:]	í[i:]	ó [o:]	ú [u:]	ű [ű:]	ő [ö:]

This gives us a 14-member vowel system. Phonetically, there is little discernible difference between the quality of *i* and *í*; *o* and *ó*; *u* and *ú*; *ü* and *ű*; or *ö* and *ő*. However, the long vowel *á* is low, central, and unrounded, while the short vowel *a* is mid-low, back, and rounded. The long vowel *é* and the short vowel *e* do not differ in backness or rounding (they are both front unrounded), but they differ in height: *é* is mid-high and *e* is mid-low. On the basis of these differences it has been suggested (e.g. in Abondolo 1988:1.1.9.2) that 'marginal' phonemes [a] (and [e]) also belong to the vowel inventory. Since the minimal pairs

vért/vért	[vert/ve:rt]	'blood-ACC/armor'
párt/párt	[part/pa:rt]	'pair-ACC/party'
Móczárt/Mozart	[mot^{s}a:rt/mot^{s}art]	'Móczár-ACC/Mozart'

suggested by Endre Tálos (pc) can*not* be told apart by native speakers, I will use the traditional 'systematic' inventory, in spite of the fact that certain other minimal pairs (based on the short variants) e.g

arra/ara	[ɔ:rɔ/ɔrɔ]	'that way/bride'
erre/ere	[ε:rε/εrε]	'this way/his vein'

as suggested by Ádám Nádasdy (pc), apparently can be. The problem of marginal vowels is clearly linked to the existence of phonemic *ë* [e] in a number of dialects.[3]

The nearly unequivocal position of grammarians with respect to the *cardinality* of the set of vowels and their phonetic properties is in sharp contrast to the diversity of the feature analyses they propose. There is no agreement concerning the *names* of the individual features (e.g. ⟨grave⟩ vs. ⟨back⟩), the *arity* of certain features (e.g. the ones for vowel height), the *values* a feature has for a given vowel, the *number* of features to be used, etc. For the reader's convenience, let me tabulate here the two main proposals (Vágó 1980, Becker-Makkai 1970):

[3] The 14-member vowel system is taken for granted in nearly every grammar dealing with the standard (Budapest) dialect of Hungarian: the list includes Lotz 1939, Hall 1944, Majtinskaja 1955, Tompa (ed) 1961, Abondolo 1988. The most important exception is Szépe 1969, which argues for the recognition of underlying *ë* in the standard dialect as well.

(2)	a	á	e	é	i	í	o	ó	u	ú	ü	ű	ö	ő
back	+	+	-	-	-	-	+	+	+	+	-	-	-	-
high	-	-	-	-	+	+	-	-	+	+	+	+	-	-
low	+	+	+	-	-	-	-	-	-	-	-	-	-	-
round	+	-	-	-	-	-	+	+	+	+	+	+	+	+
long	-	+	-	+	-	+	-	+	-	+	-	+	-	+

(3)	a	á	e	é	i	i	o	ó	u	ú	ü	ű	ö	ő
diffuse	-	-	-	-	+	+	-	-	+	+	+	+	-	-
flat	-	-	-	-	-	-	+	+	+	+	+	+	+	+
tense	-	+	-	+	-	+	-	+	-	+	-	+	-	+
grave	+	+	-	-	-	-	+	+	+	+	-	-	-	-

Originally, (2) was suggested in Szépe 1969:399 as part of the first attempt at developing a generative phonology of several Hungarian dialects. Vágó 1975 introduced the idea of taking *a* as underlyingly ⟨–round⟩, and started to use SPE feature names. In the subsequent literature, (2) has been adopted almost universally without any attempt to justify it. The emergence of autosegmental phonology did not change this situation: the 'standard' system was retained (with a separate tier for front/back) essentially in the same form as in (2) until quite recently, when the interest in feature geometry and underspecification called for a re-evaluation of feature systems.

The most important attempt at developing a different feature system is that of Becker-Makkai 1970:639, who presents her system, given in (3), in Jakobsonian terms. Unfortunately, this paper went largely unnoticed, in spite of the fact that it contains some of the quaternary data ignored in mainstream treatments. In the original system, the ⟨grave⟩ values for *i* and *í* are underspecified[4], and I have added the value '–' in (3) in order to make it comparable to the fully specified SPE system in (2).

In spite of the different feature names, the two systems are not dissimilar: ⟨tense⟩ is the same as ⟨long⟩, ⟨diffuse⟩ is the same as ⟨high⟩, and ⟨grave⟩ is the same as ⟨back⟩ (cf SPE 4.2.1). Moreover, although Becker-Makkai notes that the ⟨–flat⟩ specification for *a* "could be disputed on articulatory grounds", she opts for an abstract solution which makes her ⟨flat⟩ equivalent to Vágó's underlying ⟨round⟩. The main difference between the two systems is in the treatment of short and long *a* and *e*. In (3), the only difference is in tenseness (=length), while in (2) there are other differences as well.

Working in the framework of SPE, which permits an underlying as well as a surface inventory, Vágó did not have to make a decision that flies in the face of phonetic facts in a case such as the roundness of *a*. Nevertheless, there are distinctions in Vágó's analysis which are completely arbitrary from a phonetic point of view. Nádasdy (1985) describes the phonetic height difference between *a* and *á* as mid-low vs. low: Vágó treats both as ⟨+low⟩. Nádasdy describes the phonetic height difference between *e* and *é* is mid-low vs. mid-high: Vágó treats *e* as ⟨+low⟩ but *é* as ⟨–low⟩.

[4] "The two blank spaces represent redundancies which do not need to be marked"

But if phonetic facts cannot serve as an infallible guide in establishing a feature analysis, what is the motivation for the systems given in (2) and (3) above? What does it mean to assign a '+' or a '–' as the value of some feature for a phoneme? In the remainder of this section I will discuss this question from a general perspective — the phonological argumentation in favor of (3) will be deferred until 2.3.

A feature analysis has essentially two aims: one is to encode the phonemes with the aid of the features, and the other is to embed this analysis in the universal theory of features. In the case at hand, the aim is to choose certain features from a finite and fixed set, and describe the vowels of Hungarian in terms of these in such a manner that the following conditions are met:

(4A) Different phonemes have different feature matrices (Emicity),

(4B) Natural classes of phonemes are expressed by the combinations of these features (Compactness).

In addition to the above criteria, one would prefer a feature analysis in which the assignment of feature values is somehow "natural": for example, the contrast *o/ó* should involve only one feature, and the marked value of this feature should be assigned to *ó*. To see what is involved here, it will be necessary to give a formal rendering of feature assignment. For the sake of simplicity, I will suppose that the features involved are all binary. It will be convenient to use the values "0" and "1" instead of "–" and "+". These numbers, together with the ordinary rules of modulo 2 arithmetics (in which 1+1 is 0, but everything else happens in the usual manner), will give us the finite field GF(2), and questions of feature assignment can be translated into problems involving finite dimensional linear spaces over this simplest of finite fields.

Given a finite set F of purportedly universal features $f_1, f_2, ..., f_n$, these features form a basis in the n-dimensional linear space GF(2,n) over GF(2). In other words, it is possible to treat the (fully specified) feature matrix of a phoneme as the vector given by its coordinates. Formally, *feature assignment* can be defined as a mapping C from a given set P of phonemes $p_1, p_2, ..., p_k$ *into* the linear space GF(2,n). The points of this space are in one-to-one correspondence to the vertices of a (hyper-) cube in ordinary (Euclidean) n-dimensional space.[5] Naturally, not every vector corresponds to the feature matrix of some phoneme (in some language): the image of P under C will be denoted by Q. Condition (4A) above guarantees only that C must map different phonemes on different vectors, i.e. that C(p) and C(q) must be different if the phonemes p and q are different.

More precisely, feature assignments C_i for the (possibly overlapping) phonemic inventories of various languages and dialects should be treated as a family (set) of mappings of the above sort. The strongest possible condition on a universal feature system is that phonemes having similar (acoustic) properties should have the same feature matrix regardless of the language (dialect) to which they belong: in our formalism, this would mean that whenever phoneme p of language i is (acoustically) similar to phoneme q of language j, we must have $C_i(p) = C_j(q)$.

A weaker, but more realistic, condition is that the feature assignment must be *monotonic:* whenever we have two languages i and j with phonemes p_i, q_i, p_j, and q_j differing only in a single feature F, if

[5] For the earliest analysis of features along these lines, see Cherry 1956, 1957:3.4.

p_i is to the right of q_i on the phonetic scale corresponding to F, and p_i and p_j are both (un)marked with respect to F, then p_j must also be to the right of q_j. One effect of this condition would be to reduce the number of those vectors that appear as the image of some phoneme under some C_i. But even without this condition it is quite clear that certain feature combinations do not correspond to anything attested (or even possible) in natural languages.[6]

It should be emphasized here that the parameters that have to be fixed in this formal model reflect the choices that must be made by the linguist, rather than the choices confronting the language learner. To quote Wheeler (1972:88): "...it may be an important task of a descriptive linguist to decide which features are appropriate in the case of the language he is studying". For instance, the linguist describing the vowel system of Hungarian has to choose a particular set of features: he can adopt the system proposed by Jakobson, Fant, and Halle 1952, Halle and Clements 1983, or any of the innumerable proposals in between. Becker-Makkai 1970 opts for the former, and Vágó 1980 employs the SPE system. The language learner, on the other hand, is supposed to know the 'true' feature system, which is only approximated by the proposals in the literature.

The algebraic structure of linear spaces does not correspond naturally to the linguistic structure of phoneme inventories. The (algebraic) sum of two or more phonemes will only accidentally belong to the inventory. In other words, inventories do not form subspaces. Nevertheless, the sum operation (which gives a vector containing 1 on those positions where the summands were different, and 0 on those where they were identical) can be useful. To see this, we will first identify the universal features with certain vectors. This identification is the obvious one: the feature f_k will correspond to the vector having coordinates 0 for l ($l \neq k$), and coordinate 1 on position k. This way, features can be treated on a par with phonemes. In certain cases (e.g. the feature ⟨vocalic⟩ and the English phoneme *e* in the SPE system), features and phonemes will be identical. In such cases I will say that the phoneme "stands for" the feature in question (and, symmetrically, the feature "stands for" the phoneme). Typically, however, features cannot be represented with the aid of the phoneme inventory in such a direct manner. The following definition is suggested by the minimal pair technique:

A feature f is *witnessed* by a pair of phonemes (p_i,p_j) iff $p_i + p_j$ = f.

For example, in the SPE analysis of English segments, the feature ⟨voice⟩ is witnessed by the pairs *(b,p), (f,v)*, etc.; the feature ⟨nasal⟩ by the pairs *(n,d), (b,m)*, etc.; and so on. Clearly, in the feature analysis of a given language only those features are relevant which are witnessed by at least one pair of phonemes in the language. In 2.2, as well as in most feature analyses, this principle is applied implicitly.

The linguistically relevant structure of phoneme inventories is provided by the natural classes appearing in them. Following an idea of Fudge (1967), I will define the *degree of naturalness* of a subset S

[6] Feature cooccurrence restrictions, or FCRs, as proposed by Stanley (1967) will have to be stated for every category (such as degree of voicing) for which the number of contrasts in a given language is not a power of 2. For instance the low/mid/high opposition in vowel height will require the exclusion of the combination *⟨+high, +low⟩. FCRs, however, are not necessarily the by-products of the consistent application of the binarity principle: it is simply an empirical fact that certain combinations of otherwise well-motivated features (corresponding to, say, uvularity and clicks, respectively) do not appear simultaneously in natural languages.

in a phoneme inventory P as the number of rules in which S occurs as context. This makes it necessary to enlarge the surface inventory so as to include underlying phonemes as well. But this can be done without increasing the complexity of the model — the set P will simply have to be replaced by some other set P'. Thus, the class {*b,u*} of Hungarian phonemes will have 0 degree of naturalness, while the class {*p,b,m*} has positive degree of naturalness because the phonemes in it trigger $n \rightarrow m$ assimilation (see ch. 2.5 of Vágó 1980). Needless to say, a subset S' of P' can be natural also by virtue of undergoing, rather than triggering, some process. From the formal point of view, the criteria we employ in distinguishing natural classes from 'unnatural' ones (cf. e.g. Rubach 1982 ch. 3.2) are unimportant: all that matters is that natural classes (and their degree of naturalness) are given externally. In other words, natural classes are present in the system (much like the phonemes) before we attempt to develop a feature analysis.

This is particularly clear for major classes, such as stops or vowels (listed in the first line of (1A) and in (1B) above): the system of major class features (to be discussed in 2.2) has to be constructed in accordance to these classes, rather than the other way round. Those classes that can be expressed by fewer features than their individual members will be called N-classes (cf. Halle 1964:328). Geometrically, N-classes are the intersection of Q with (hyper-) planes parallel to the coordinate axes, and, as such, will depend on the feature analysis (the mapping C) chosen. This definition makes it possible to reformulate the 'compactness' condition discussed above as

(4C) The N-classes given by C should coincide with the natural classes of P. (Compactness)

It is worth noting that the 'classical' feature analysis model, as outlined above, makes rather specific predictions concerning natural classes. First, there cannot be many. If there are p (surface) phonemes, we will need at least $f = \log_2 p$ features to satisfy condition (4A). Each N-class is given by assigning some specific value (+ or –) to some features, while leaving the rest unspecified. This gives us 3 options for each feature, so with the aid of f features, 3^f N-classes can be defined. Therefore, the number of natural classes is expected to be $3^{log_2 p} = p^{1.58}$. This is a very small fraction of the possible (natural or unnatural) subsets, of which there are 2^p. To get an idea of the order of magnitudes involved, suppose $p = 30$. The number of natural classes is $30^{1.58} = 220$ (or $3^5 = 243$ if we use 5 features, rather than the theoretical 4.91); while the number of classes is $2^{30} = 1,074,000,000$.

Second, the set of natural classes must be closed under intersection. More precisely, two natural classes, both containing some phoneme q will intersect either in q, or in a full natural class.[7] This is easy to prove for N-classes (no matter what function C we choose): therefore, the (meta)theory predicts that it must be true for natural classes as well. In practice, natural classes are defined by a few features, i.e. with the majority of the features left unspecified. Since the intersection of such classes will be specified for every feature for which at least one of the original classes were specified, we will eventually get classes in which only a few features are still unspecified — these are more often called *archiphonemes.*

For example, the set of high (diffuse) vowels contains *i,í,u,ú,ü,* and *ű* — this is clearly a natural

[7] It is a matter of definition whether we call one-member classes natural. If we take lexical items to be identity rules, as proposed by Kiparsky (1982), each phoneme will appear in a huge number of rules and will, therefore, have a very high degree of naturalness.