Hungarian Vowel Harmony

András Kornai

Stanford University

0 Introduction

The basic characteristics of Hungarian vowel harmony are well known – the data introduced in the pioneering work of Vágó (1975,1976) has been subjected to a large number of analyses. In this paper I will assume the reader to be familiar with this literature and with the basic data concerning 'binary' suffixes such as the dative *nak/nek*. (For an overview, see v.d. Hulst 1985 or Siptár 1984).

In Section 1, I will exemplify a number of additional facts concerning the harmonic behavior of 'ternary' suffixes such as the allative *hoz/hez/höz* and 'quaternary' suffixes such as the accusative *at/et/ot/öt*. Quaternary suffixes, presented here for the first time, appear to be problematic both for the standard accounts and for the more recent theory of underspecification. Ternary suffixes, although discussed by Vágó (1975), have largely been ignored in the subsequent literature – here the familiar data concerning binary and ternary suffixes will be reorganized along the lines dictated by the new facts.

Section 2 gives an autosegmental account of the harmonic patterns in Hungarian. The use of a single diacritic feature, a '*floating* –U' will make it possible to treat the regular and the exceptional patterns together. Negative feature values and core specification are used only for the exceptional cases – the basic system uses only single-valued features that leave no room for underspecification. The analysis brings into sharp relief a hitherto unnoticed parallelism between high and mid vowels that extends even to the exceptional cases.

Section 3 discusses the behavior of neutral vowels. The status of neutral vowels in harmony systems is problematic for autosegmental theory in general, and the present treatment is no exception. As we shall see, each vowel in Hungarian can show 'regular' or 'exceptional' behavior, and the exceptional forms can be derived by the same rules as the regular forms by adding a single diacritic to the underlying representation of exceptional stem vowels. In neutral vowels we can observe a third kind of behavior, so we will have to adjust the underlying representations in an extraordinary fashion (namely, by core-specification) to account for the two kinds of 'abstract' *i* and *e*.

In the presentation of the data, the emphasis will be shifted from vacillating stems to exceptional, but non-vacillating stems. The reasons for this shift are methodological. The exceptional 'Class I' and 'Class IV' stems show the same unambiguous behavior irrespective of sentence stress, syntactic environment, register, etc. for every speaker of standard Hungarian. In the phonological study of vacillation it would be necessary to control for all of these factors, and perhaps for others as well. The existing literature, based largely on anecdotal evidence, is of little help here, as it makes no systematic distinction between dialect mixture, vacillation in production, and varying degrees of acceptance. In fact, empirical work on vacillating stems is still in its infancy[1], while the information on exceptionality is readily available from standard dictionaries.

*I would like to thank Don Churma, Sharon Inkelas, László Kálmán, Paul Kiparsky, Will Leben, Marcy Macken, Ádám Nádasdy, Livia Polányi, and Péter Siptár for their comments and criticisms on various versions of the manuscript. This work was in part supported by a grant from the Hungarian Academy of Sciences.

[1] For the first steps this direction, see Kontra and Ringen 1985,1986.

1 The data

The vowel system of Hungarian is given in (1) below – for a description in phonetic terms, see e.g. Nádasdy (1985). In this paper, forms will be given in orthographic, rather than phonetic, transcription.

(1)

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

1.1 The feature system

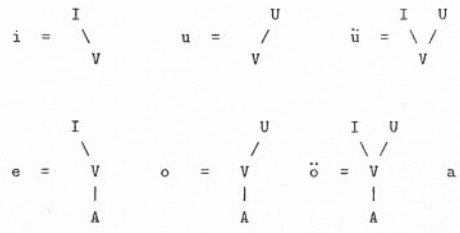
In general, discussions of Hungarian vowel harmony have revolved around what I will call the 'binary' suffixes, i.e. those that have two surface realizations. Ternary suffixes (i.e. those having three alternants) are at best mentioned in passing, and quaternary suffixes are usually completely ignored. For the binary suffixes it is of course sufficient to deal only with one autosegment (or one distinctive feature), and leave the larger issue of the featural composition of Hungarian vowels untouched. But ternary and quaternary alternations involve at least two autosegments, and the choice of feature analysis (which will determine the rest of the analysis to a surprising extent) is far from trivial. For the reader's convenience, let me tabulate here the two main proposals (Vágó 1980, Becker-Makkai 1970):

(2)

	a	á	e	é	i	í	o	ó	u	ú	ü	ű	ö	ő
Back	+	+	-	-	-	-	+	+	+	+	-	-	-	-
High	-	-	-	-	+	+	-	-	+	+	+	+	-	-
Low	+	+	+	-	-	-	-	-	-	-	-	-	-	-
Round	+[2]	-	-	-	-	-	+	+	+	+	+	+	+	+
Long	-	+	-	+	-	+	-	+	-	+	-	+	-	+
Diffuse	-	-	-	-	+	+	-	-	+	+	+	+	-	-
Flat	−[3]	-	-	-	-	-	+	+	+	+	+	+	+	+
Tense	-	+	-	+	-	+	-	+	-	+	-	+	-	+
Grave	+	+[3]	-	-	−[4]	−[4]	+	+	+	+	-	-	-	-

In this paper I will use a tridirectional feature system as proposed by Rennison (1984), Kaye & al (1985), and others. In essence, tridirectional features amount to an autosegmental version of the Jakobsonian feature system given above.

(3)



[2] Derived from underlying -Round. Cf Vágó (1980) 1.1.1. Essentially the same feature matrix is given in Szépe (1969: 399).

[3] "Could be disputed on articulatory grounds" (Becker-Makkai 1970:639)

[4] "The two blank spaces represent redundancies which do not need to be marked" (loc cit). Redundancy rules would presumably add the value –.

Long vowels will be treated as geminates, because their harmonic behavior is identical to that of the corresponding short vowels.[5]

An important advantage of this system is that it predicts what kinds of binary alternations are possible in Hungarian. If an alternating pair is defined by the presence vs. absence of the feature I, (3) shows that *a* will be paired with *e*, *ó* will be paired with *ő* and *u* will be paired with *ü* – these are precisely the pairs attested in Hungarian, as can be seen from (4) below. Moreover, since *i* is paired with the empty vowel, which does not exist in Hungarian, the system predicts that it cannot take part in alternations. (For an exceptional case, see Section 3.)

1.2 The major alternations

Suffixes will be subcategorized according to their *arity*, i.e. the number of harmonic alternants they have. Unary suffixes, i.e. those that do not show harmonic alternation, will be ignored. Some suffixes with higher arity are listed below:

(4)

BINARY	
nak/nek	'DAT'
nál/nél	'ADE'[6]
tól/től	'ABL'
nok/nök	'professional characterized by' e.g. *szó* 'word', *szónok* 'orator'
ul/ül	'as' e.g. *ember* 'man', *emberül* 'as a man'
TERNARY	
hoz/hez/höz	'ALL'
on/en/ön	'SUE'
QUATERNARY	
at/et/ot/öt	'ACC'
ak/ek/ok/ök	'PL'
am/em/om/öm	'1SG POS'
as/es/os/ös	'having'

This list is representative: there are no alternations involving other (sets of) vowels. Thus, ternary suffixes will always show *o/e/ö*, and quaternary suffixes *a/e/o/ö* alternation.[7] Quaternary suffixes, though ignored in standard treatments, are in fact anything but marginal: the most frequently encountered suffixes, such as the accusative, the 1st and 2nd sg possessive, and the plural are all of this form. Nor are they restricted to inflection: the last example, *as/es/os/ös* 'having, having to do with' is a high-frequency derivational suffix that forms adjectives from nouns.

1.3 The harmonic taxonomy

Stems will be subcategorized according to the suffix-alternants they select. The taxonomy developed here will be applicable not only to stems, but also to fully formed words that can undergo further suffixation. The five classes presented below offer a theory-neutral descriptive framework encompassing all the non-vacillating Hungarian data.

In the binary case, all stems can be divided into two classes, 'Back' and 'Front', according to the quality of the alternant they select. This taxonomy is based on the fundamental observation that the binary alternants can be arranged in parallel series: if a stem takes

[5] With the possible exception of neutral vowels in vacillating forms.

[6] The abbreviations for the less familiar case endings are as follows: ADE(ssive), ABL(ative), ALL(ative), SU(per)E(ssive).

[7] In vowel-initial suffixes, the vowel is dropped regularly after stems ending in vowels, and occasionally after consonant-final stems – for the investigation of vowel harmony, this phenomenon can safely be disregarded.

-nak in the dative, it will take *-nál*, rather than *-nél* in the adessive, *-tól*, rather than *-től* in the ablative, etc.[8]

With the introduction of quaternary suffixes, a four-way partitioning results, according to the quality of the vowel in *ak/ek/ok/ök*. Such a partitioning is then justified by the fact that other quaternary suffixes will take the alternant with the same vowel. Indeed, no stem can subcategorize for *-ak* in the plural but *-ot* in the accusative, and in general the distributions of the quaternary alternants, as far as it can be established,[9] are completely parallel.

This classification predicts the harmonic behavior of binary suffixes: all stems that take *-ak* (Class I) or *-ok* (Class II) in the plural take 'Back' suffixes and the rest take 'Front' suffixes. These classes also select for the *o*-alternant in ternary suffixes. In general, stems that take *-ek* (Class III) take the *e*-alternant in ternary suffixes, and stems that take *-ök* take the *ö*-alternant in ternary suffixes (Class V). But, most strikingly, the four-way classification established on the basis of quaternary suffixes does not fully predict the distribution of ternary suffixes. There are a number of stems (Class IV), that take *-ek* in the plural but take the *ö*-alternant, rather than the *e*-alternant with ternary suffixes.

It should be emphasized that the behavior of Class IV stems is qualitatively different from that of vacillation: forms like **hölgyhez, *tőgyhez, *sülthez*, and **tüzhez* are unacceptable in every idiolect, and in ECH [10] the forms **hölgyöt, %tőgyöt*,[11] **sültöt*, and **tüzöt* are clearly unacceptable.

In order to characterize the range of possible stem vowels within each class, monosyllabic examples were chosen. It will be apparent from (5) that the quality of the stem vowel determines the harmonic behavior of the stem to a large extent. If the stem vowel is *ö* or *ü*, the stem must be in Class IV or Class V, and if it is *a, o* or *u*, the stem will belong in Class I or Class II. As long as the last vowel of a polysyllabic stem (or word) is not *i* or *e*, the quality of the last vowel will predict harmonic behavior the same way as with monosyllabic stems. If the last vowel is neutral, the situation is more complex – the details will be discussed in Section 3.

For +I +U vowels the choice between Class IV and Class V is lexically determined, as is the choice between Class I and Class II (for –I vowels). However, only the selection of Class I or Class IV has to be marked in the lexicon – the default case is Class II for stems in *a, u,* and *o*; and Class V for stems in *ü* and *ö*. This is particularly clear for Class IV, which contains roughly 20 monomorphemic stems, as opposed to the thousands of monomorphemic stems in Class V. That Class IV is the marked class can also be seen from the fact that all recent loans in *ü* or *ö* are in Class V. Although Class I is much larger (it contains more than a thousand monomorphemic members), it is still considerably smaller than Class II. It is also closed: nonce-words and recent loans in *a, u,* and *o* always belong in Class II. For the same reasons, the default is Class III for monosyllabic stems in neutral vowels.

In (5) below the examples are all monomorphemic noun stems, but every non-vacillating Hungarian word falls into one of these classes, irrespective of morphemic composition or lexical category. The plural suffix is representative of the quaternary type: whenever the initial vowel is present in the other quaternary suffixes, it is the same as in the plural. The allative is representative of the ternary type, and the dative, adessive and formal suffixes have been chosen to represent the three major alternating pairs *a/e, o/ö,* and *u/ü*. For the treatment of non-standard (morphophonemic) alternation, see Section 3.

[8] This observation has been contested on several occasions, but I was unable to find non-vacillating conterexamples that would take, say, *-nak* in the dative but *-től* in the ablative. At any rate, it is an extremely robust generalization with no systematic counterexamples.

[9] The alternating vowel is lost after certain stems for some of the quaternary suffixes, e.g. we have *borok* 'wine-PL' but *bort* 'wine-ACC'.

[10] Educated Colloquial Hungarian, the standard (Budapest) dialect. See Nádasdy (1985)

[11] This form has a decidedly *vidéki* 'rural' flavor for Budapest speakers, but is, perhaps, acceptable in some dialects. Note that different dialects can have different vowel inventories.

(5)

	stem	4-ary	3-ary	2-ary *a/e*	2-ary *o/ö*	2-ary *u/ü*	gloss
I							
	had	hadak	hadhoz	hadnak	hadtól	hadul	'army'
	ház	házak	házhoz	háznak	háztól	házul	'house'
	lyuk	lyukak	lyukhoz	lyuknak	lyuktól	lyukul	'hole'
	kút	kutak	kúthoz	kútnak	kúttól	kútul	'well'
	hold	holdak	holdhoz	holdnak	holdtól	holdul	'moon'
	ló	lovak	lóhoz	lónak	lótól	lóul	'horse'
	in	inak	inhoz	innak	intól	inul	'tendon'
	híd	hidak	hídhoz	hídnak	hídtól	hídul	'bridge'
	héj	héjak	héjhoz	héjnak	héjtól	héjul	'crust'
II							
	bab	babok	babhoz	babnak	babtól	babul	'bean'
	báb	bábok	bábhoz	bábnak	bábtól	bábul	'puppet'
	rum	rumok	rumhoz	rumnak	rumtól	rumul	'id.'
	húr	húrok	húrhoz	húrnak	húrtól	húrul	'chord'
	bot	botok	bothoz	botnak	bottól	botul	'stick'
	drót	drótok	dróthoz	drótnak	dróttól	drótul	'wire'
	zsír	zsírok	zsírhoz	zsírnak	zsírtól	zsírul	'fat'
	cél	célok	célhoz	célnak	céltól	célul	'goal'
III							
	hit	hitek	hithez	hitnek	hittől	hitül	'belief'
	víz	vizek	vízhez	víznek	víztől	vízül	'water'
	fej	fejek	fejhez	fejnek	fejtől	fejül	'head'
	érv	érvek	érvhez	érvnek	érvtől	érvül	'argument'
IV							
	hölgy	hölgyek	hölgyhöz	hölgynek	hölgytől	hölgyül	'lady'
	tőgy	tőgyek	tőgyhöz	tőgynek	tőgytől	tőgyül	'udder'
	sült	sültek	sülthöz	sültnek	sülttől	sültül	'roast'
	tűz	tüzek	tűzhöz	tűznek	tűztől	tűzül	'fire'
V							
	tök	tökök	tökhöz	töknek	töktől	tökül	'pumpkin'
	bőr	bőrök	bőrhöz	bőrnek	bőrtől	bőrül	'skin'
	füst	füstök	füsthöz	füstnek	füsttől	füstül	'smoke'
	bűn	bűnök	bűnhöz	bűnnek	bűntől	bűnül	'sin'

1.4 The possessive paradigm

In the morphology of Hungarian, ternary and quaternary suffixes are fully integrated with the rest of the system: not only do the various kinds of suffixes appear in the same position, e.g. as case markers, regardless of arity, but they can also change arity as a result of morphological processes. These facts are exemplified by the possessive paradigm given in (6) below. Only one stem is given for each harmonic class: *úr* 'master' (Class I); *sógor* 'brother-in-law' (Class II); *ember* 'man' (Class III); *hölgy* 'lady' (Class IV); and *őr* 'guard' (Class V).

The 1st, 2nd, and 3rd person forms are given in the 1st, 2nd, and 3rd column respectively. With each stem, the first and third rows contain singular, and the second and fourth rows contain plural possessor forms: the difference between the two is in the number of the possessed (the stem), which is singular in the first two rows and plural in the last two. The reader is encouraged to analyze the forms into component suffixes before looking at the solution offered in 2.3.

(6)

	1st	*2nd*	*3rd*	possessor	possessed
I					
	uram	urad	ura	*sg*	*sg*
	urunk	uratok	uruk	*pl*	*sg*
	uraim	uraid	urai	*sg*	*pl*
	uraink	uraitok	uraik	*pl*	*pl*
II					
	sógorom	sógorod	sógora	*sg*	*sg*
	sógorunk	sógorotok	sógoruk	*pl*	*sg*
	sógoraim	sógoraid	sógorai	*sg*	*pl*
	sógoraink	sógoraitok	sógoraik	*pl*	*pl*
III					
	emberem	embered	embere	*sg*	*sg*
	emberünk	emberetek	emberük	*pl*	*sg*
	embereim	embereid	emberei	*sg*	*pl*
	embereink	embereitek	embereik	*pl*	*pl*
IV					
	hölgyem	hölgyed	hölgye	*sg*	*sg*
	hölgyünk	hölgyetek	hölgyük	*pl*	*sg*
	hölgyeim	hölgyeid	hölgyei	*sg*	*pl*
	hölgyeink	hölgyeitek	hölgyeik	*pl*	*pl*
V					
	őröm	őröd	őre	*sg*	*sg*
	őrünk	őrötök	őrük	*pl*	*sg*
	őreim	őreid	őrei	*sg*	*pl*
	őreink	őreitek	őreik	*pl*	*pl*

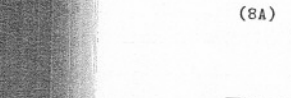
2 The analysis

In this section I will develop an analysis of the above data in a step-by-step fashion. The starting point will be the observation that the difference between the ternary and quaternary suffixes can always be expressed by a single feature. This leads to two maximally simple spreading rules that interact with a single exception feature. In order to capture the full pattern reflected in the taxonomy developed above, further rules and representations will be introduced along the way. The resulting system is then applied to the description of the possessive paradigm.[12]

2.1 The basic system

Closer inspection of the data in (5) reveals that whenever there is a difference between the quality of the ternary and quaternary suffix vowels (*-at* vs *-hoz* in Class I, and *-et* vs *-höz* in Class IV), the quaternary suffix will have the –U. The most straightforward analysis (which will have to be supplemented by other rules later), is to spread the feature values for I and U onto the suffixes, and adjust the underlying representation of exceptional stems so that the spreading of U is blocked for them. This can be achieved by marking the stems in Class I and Class IV by a floating –U. Thus, the cornerstone of the analysis is the following pair of spreading rules:

(8A)



(8B)

[12] For an exhaustive analysis of the nominal and verbal paradigms in Hungarian using the same rules, see Kornai (1986 Ch 4).

Supposing that the vowel of quaternary suffixes is specified only for A, the rules in (8) will derive e.g. *tökök* from *tök* simply by spreading the I and U features of the stem vowel. This solution can readily be extended to ternary suffixes by supposing that these are represented underlyingly with a floating U which can dock only if the I feature of the stem did not spread.

The most important problem with this simple solution is that it does not capture the pattern of exceptions. Without additional 'cleanup' rules we would have to mark every element of Class II in the lexicon, in spite of the fact that Class II is productive, and Class I is closed. In order to deal with this problem, we must posit an independent source for the feature U in recent borrowings. The field of computer science provides many examples such as *fájl/fájlok* 'file/PL', *ram/ramot* 'random acces memory/ACC' – these forms make it clear that there has to be an U in the representation of the quaternary suffixes themselves.

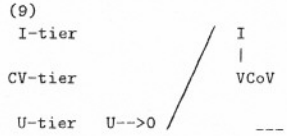
Adopting a proposal of Halle-Vergnaud (1982), I will take this U to be specified in the phonemic core.[13] For those features, like A in Hungarian, that do not harmonize there is no good reason to establish a separate tier. I suppose that the unmarked place for a segmental feature is on the segmental tier, and we autosegmentalize a feature only if there is some evidence for this. However, the segmental core remains a possible location even if the feature has been lifted to a separate tier – the present analysis will make use of this option only in the case of exceptional elements.

In sum, the underlying representation of quaternary suffixes contains the same features, namely A and U, as that of the ternary suffixes: the difference being that in the quaternary case the U is in the core and in the ternary case it is floating. This 'geometrical' difference will surface only after exceptional stems and, as we shall see later, after certain suffixes. Before turning to these, let me show first how the non-exceptional forms are derived.

In Class II, *zsír* and *cél* are exceptional (there are less then 10 monomorphemic *i* and *e* stems there) – the Class II pattern is regular only for stems in *a*, *o* and *u*. Since these do not contain the feature I, (8A) is inoperative. Whether (8B) actually spreads the feature U in the case of stems in *u* and *o* can not be decided on the basis of these forms, since both ternary and quaternary suffixes have an underlying U (albeit on different tiers).

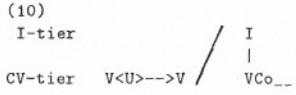
I-spread gives the right result in Class III, provided that the U floating over *hoz/hez/höz* is stopped from linking up. This is achieved by a rule of floating U deletion:

(9)



In the same environment, the core U of *ak/ek/ok/ök* must also be deleted:

(10)

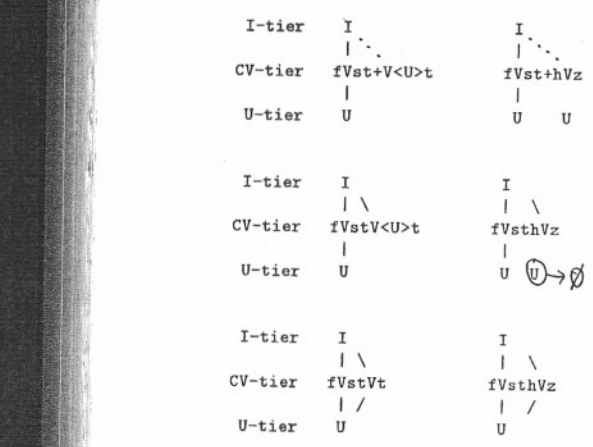


Finally, in Class V, I-spread (8A), U-deletion (9-10), and U-spread (in this order) give the right result.[14] This is illustrated in (11) by the derivation of the forms *fűstöt, fűsthöz*. Notice, that the derivation is essentially the same for forms like *tököt, tökhöz*, as the presence or absence of A-specification plays no role in any of the rules. Although the suffixes in question are specified for A, the high-mid parallelism makes it possible to omit the As from the display altogether.

[13] As geminates will not be discussed, the root tier is omitted from the display, and the segmental core is identified with the CV tier. Nothing hinges on these simplifying assumptions.

[14] Roughly speaking, (9) and (10) have the same effect as the 'Rounding Harmony' rule of Vágó (1980: 1.7).

(11)

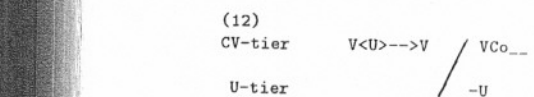


The lack of negative specification in the rules means that we can interpret U, I, and A as *simplex* features representing privative oppositions: I will return to this question in Section 4.

2.2 The exceptional forms

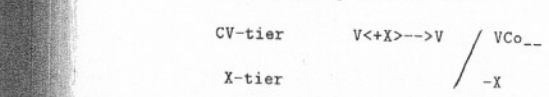
With the rules (9-10), the representation of non-exceptional items (*a á o ó u ú* in Class II, *i e é* in Class III, and *ö ő ü ű* in class V) was kept simple. The exceptional items are treated as follows. In Class IV, we have to stop the core U of quaternary suffixes from taking effect, and we do *not* want to derive the *e*-alternant for ternary suffixes. To do this, let us will suppose that the diacritic –U triggers the deletion of core Us:

(12)



In other words, for a negative feature 'spreading' to a positively specified core amounts to deleting the offending feature from the core, and I will assume that the negative feature also disappears. This can be formulated as a general rule:

(13) Core/Neg Annihilation



With this rule, ordered after the others, Class IV is taken care of. As for the remaining exceptional elements, the *i* of the *hid*-type words is specified for I in the core (therefore it will not spread); in addition, it will have a floating -U, which will derive the correct *hidat (*hidot)*. The *i* of the *zsír*-type words is also specified in the core *(*zsírnek)*, but will have no other exceptional property: therefore, we derive the correct *zsírok (*zsírak)*. The exceptional *é* of *héj* has I in the core, and floating –U: this will give us the correct *héjak, héjhoz*. The only stem in *e* or *é* paralleling the behavior of *zsir* is *cél* 'goal': this has I in the core but no other exceptional marking.

Every other exceptional element will be marked by a floating –U : the U specification (where present, e.g in *lyuk, hold, hölgy*) must be relegated to the core (CV tier). This simple and unified treatment of back stems in Class I is a highly desirable result, given the fact that Class I contains more than a thousand monomorphemic noun stems (and several thousand compounds) of this kind, while the remaining exceptional types are only sporadically represented.

Moreover, the use of the feature –U unifies the treatment of the exceptional classes: the only vowels we do not find in Class I are *ü* and *ö*: when these are marked by –U, they belong in Class IV. In addition, the tridirectional feature system captures the hitherto unnoticed parallelism between the set *u ü i* of high vowels and the set *o ö e* of mid vowels. Since no rule makes reference to the feature A, we expect to find paired elements always in the same class. As can be seen from the data presented in (5), this expectation is fulfilled not only by the regular vowels, but by the irregular ones as well.

In the case of the feature A core-specification was chosen because A never spreads – other features have appeared in the core only exceptionally, just in case they do not undergo the spreading rules in (8). But for I, something like core-specification is necessary in non-exceptional cases as well, since the *feature* I spreads, but the *vowel i* is generally permeable ('transparent'). This latter fact is best exemplified by the denominal suffix *-i* 'characterized by the location' that forms adjectives:

(14)

ház	házat	házi	háziak	'house'
kert	kertet	kerti	kertiek	'garden'
nyár	nyarat	nyári	nyáriak	'summer'
tél	telet	téli	téliek	'winter'
Fót	Fótot	Fóti	Fótiak	'the village Fót'
Tök	Tököt	Töki	Tökiek	'the village Tök'

In fact, the suffixes *am/em/om/öm* and *ak/ek/ok/ök* both turn Class II and Class V forms into Class I and Class IV forms, respectively,[15] so these latter classes, although exceptional for underived nouns, contain at least two inflected forms for every noun.

(15)

bot	botom	botot	botomat	*botomot
füst	füstöm	füstöt	füstömet	*füstömöt

Adding the above mentioned suffixes to Class I or Class III stems results again in Class I and Class III forms. In the case of Class IV, there is some vacillation: forms like *%könyvekhöz* 'book-PL-ALL', *%hölgyekhöz* 'lady-PL-ALL' are acceptable for most native speakers, but the ECH forms appear to be *könyvekhez, hölgyekhez*. Therefore, suffixation of Class IV stems leads to class III forms in ECH and perhaps to Class IV forms in certain dialects.

The analysis can be extended to capture the behaviour of suffix-combinations simply by marking quaternary suffixes (and the *-i* discussed above) with a floating –U. In this way suffix-combinations such as those in (15) can be derived without further complications. The adjective-forming denominal suffix *as/es/os/ös* 'having (to do with)', which must also be marked this way, offers a particularly good way of testing the proposed mechanism.

In isolation, a form containing this suffix can exhibit behavior characteristic of Class II forms:

[15] The accusative suffix, being absolute word-final, cannot be tested.

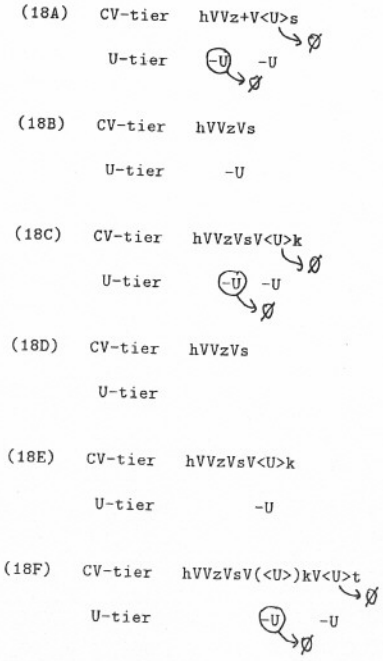
(16)
ház 'house' házas/házasok 'married/PL' *házasak

However, if the form appears in a non-lexicalized meaning, as in

(17) A kertes könyveket jobbra tedd, a házasakat pedig balra!
'Put those books that have gardens (in them) to the right side,
and those that have houses to the left'

the form surfaces as *házasak*.[16] This behavior can be explained only if we suppose that the exceptional marking introduced by *-as* is lost in the lexicalized form *házas*, but can be present if the form is derived anew, as required in (17). This derivation is given in (18).

The only rule that will play a role in the derivation is Core/Neg Annihilation, which derives (18B) as the representation of the form *házas* from the underlying (18A). If the exceptional marking contributed by the suffix *as/es/os/ös* is retained, we derive (18C) on the next cycle. If the exceptional marking is lost, so that we start with (18D), we derive (18E) as the result of plural suffixation. Finally, since the plural form is not in the lexicon, the exceptional marking contributed by the plural suffix can not be lost, so in the last cycle we derive (18F) in both cases.



2.3 The possessive paradigm

The analysis presented above enables us to derive the possessive paradigm in a surprisingly simple fashion. The 1st and 2nd sg markers differ from the accusative and plural markers only in the segmental contents of their C-slots:

[16] In ECH, %házasokat is tolerated, but hardly ever produced in (17) and similar contexts.

(18)

```
V<U>C    V<U>C
|  | ;  |  |
A  m    A  d
```

The 3rd sg, 1st pl, and 3rd pl suffixes show the ordinary *a/e* and *u/ü* binary alternations, and the 2nd pl is given by

(19)

```
      U

V<U>CVC
|  | |
A  t k
```

One might derive this form by adding the regular plural to a devoiced 2nd sg – this would show that quaternary vowels do not necessarily have a floating –U. Since the rule can not be generalized to other persons, I will not pursue this matter here.

Now, the forms in rows 3 and 4 can be derived from those in rows 1 and 2 with a single rule which deletes the U (if any) of the first V and adds an A instead, and infixes an *i* after it:

(20)

```
V -->  VV<I>
       |          /  __ +POS PL
       A
```

What happens here is that both the quaternary *a/e/o/ö* and the binary *u/ü* are replaced by the binary *a/e* at the beginning of the suffix: these replacements can be treated as one and the same process if we suppose that what happens is that the feature U is replaced by A. This process can be decomposed into three steps: first the feature U (if present) is deleted, no matter whether it was in the core or on its own tier, and second the feature A is added. (As a third step, if the resulting V is doubly specified for A, the OCP collapses the two As.)

The idea of treating the plural marker of the possessed as an infixed element goes back to Antal (1959,1963). Rácz (1974) argues that Antal has to posit e.g. *ank* and *enk* alternants in addition to the regular *unk* and *ünk* forms in 1st pl – the present rule obviates his criticism. Yet another counterargument to infixation is based on the 3rd sg pos pl forms of V-final stems. In general, we have *hajói* 'his ships' instead of **hajójai*. There is a great deal of vacillation *(ajtói %ajtai %ajtajai* 'its doors', *ekéi %ekéjei* 'his ploughs', *lépcsői %lépcsőjei* 'its stairs', *hajdúi %hajdújai* 'his attendants'), and in 'Radio' Hungarian the *-jai* forms are acceptable only after *i*-final stems in general. But these forms appear to be problematic regardless of infixation, and the present model can handle them with a lexically governed rule that deletes *ja*.

3 Neutral vowels

The treatment of neutral vowels in polysyllabic stems remains a problem area. It appears to be a very real problem for speakers of Hungarian as well: most of the 'vacillating' cases involve *i* and *e* in non-initial syllables. The solution presented in Section 2. was based on the assumption that *i* and *e* have the feature I in the core unless they are stressed (= appear in the first syllable). However, wo do not have independent evidence that stress and harmony interact in Hungarian, and this makes unclear whether the present analysis contributes to our understanding of the problems posed by neutral vowels.

In this Section I will give a somewhat speculative account of neutrality based on the concept that archiphonemes are not only the sites, but also the triggers of the harmonic process. This shift in focus is perhaps best motivated by the existence of 'morphophonemic' harmonic alternation where it is clear that the proper surface forms can not be derived by simple spreading but require some additional rule(s) which must be tied to the alternating morpheme.

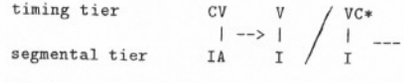
The 3rd sg, 2nd pl and 3rd pl present definite suffixes in Hungarian provide such an example. The following table gives the present tense paradigm of the stems *vár* 'wait', *kér* 'ask', and *tűr* 'suffer'. The items in the first, second, and third columns are first, second, and third person forms, respectively. With each stem, the first two rows give the singular, and the last two rows give the plural forms: the indefinite conjugation is in the odd rows, and the definite conjugation is in the even rows.

	1st	*2nd*	*3rd*	number	conjugation
(21A)	várok	vársz	vár	*sg*	*indef*
	várom	várod	várja	*sg*	*def*
	várunk	vártok	várnak	*pl*	*indef*
	várjuk	várjátok	várják	*pl*	*def*
(21B)					
	kérek	kérsz	kér	*sg*	*indef*
	kérem	kéred	kéri	*sg*	*def*
	kérünk	kértek	kérnek	*pl*	*indef*
	kérjük	kéritek	kérik	*pl*	*def*
(21C)					
	tűrök	tűrsz	tűr	*sg*	*indef*
	tűröm	tűröd	tűri	*sg*	*def*
	tűrünk	tűrtök	tűrnek	*pl*	*indef*
	tűrjük	tűritek	tűrik	*pl*	*def*

The first person suffixes show standard harmonic alternation, and the same holds for the first three second person suffixes and the indefinite suffixes in third person. The third person definite forms (both for singular and plural) and the 2nd pl definite ending behave differently however.

The 3rd sg definite suffix appars as *i* with every stem that takes front suffixes, and as *ja* with all other stems (i.e the ones that take back suffixes). As we have seen in (2), the regular alternant of *a* is *e*. Thus, we would expect *ja/je* alternation, and the *i* form will have to be derived from *je* by a special rule.[17]

(22) 'i/ja harmony'



The first element (which drops out, together with the A in front contexts) is truly morphophonemic: in addition to taking part in nonstandard harmonic alternation, it also assimilates to a preceding sibilant. 'Ordinary' *j* stays unchanged: compare *rázza* 'he shakes it' to *kézjel (*kézzel)* 'handsign'. This is captured by leaving the I unassociated. Since in Hungarian the default consonant for empty Cs is *v*[18], we can not say that the unassociated I of the initial C is not present at all, for this would give us **várva*. Moreover, the floating I provides the obvious source for the front alternant. In the rule of i/ja harmony it is sufficient to say that

[17] Another solution would be to take the *i* alternant as basic, but in Hungarian suffixes with *i* generally do not show harmonic alternation, so we would still need a special rule that turns *i* into *ja* in back vowel contexts. For the sake of concreteness, I'll take the *ja* form as basic here, but nothing hinges on this assumption.

[18] Cf the 'abstract w' of Vágó (1980).

the initial C and the A are deleted in front harmonic contexts: the I will link up with the remaining V and give us the right result without further stipulations.

Irrespective of the details of the analysis, it is quite clear that the change from one alternant to the other can not be accomplished just by spreading the feature I. Nevertheless, the notion 'front harmonic context', only sketchily formulated in (22), is the same for nonstandard harmony rules as for the standard cases. The preceding treatment of vowel harmony was based on the assumption that neutral vowels are specified for the feature I in the core, except when they appear in initial (stressed) syllables where the I of neutral vowels has to appear on the I tier. Given this assumption, unbounded spread of the feature I will give the desired results (if we suppose that the core I-s that follow do not block spreading). A 'front harmonic context' means that there is an I on the I-tier which is not followed by vowels that do not bear I in the core or on the I-tier.

These assumptions, although not unreasonable, are hard to reconcile with a tentative generalization[19] concerning stems that end in two or more neutral syllables. In order to keep the following discussion separate from the analysis developed in the previous section, I will use finite automata, rather than standard autosegmental notation in the description of vacillation. I will restrict myself to the binary case: for the sake of simplicity, *a o u* will be called *back*; *i e* will be called *neutral*; and *ü ö* will be called *front*[20] vowels.

The first generalization that we have to take into account is that neutral vowels are transparent. More precisely, stems in which the penult is back (front) take back (front) suffixes if the final vowel is neutral. There are a number of vacillating stems that do not fit into the general pattern, and it is unclear when vacillation means idiolectal variation (such that each individual speaker uses either the front or the back alternants quite consistently), and when it means that one and the same speaker uses both back and front alternants.

The second generalization is that a stem containing only neutral vowels takes the front alternants of every suffix – there seems to be no vacillation here. Therefore, in disyllabic stems with a final neutral vowel, if the penult is neutral, we expect (and get) front suffixes. But the status of trisyllabic stems in which the first vowel is back and the other two are neutral is left open by both generalizations, and the same holds for stems in which a back vowel is followed by a longer sequence of neutral vowels (examples are hard to find).

The third generalization, already alluded to, concerns these stems: the overall impression is that they show front harmonic behavior. This observation is easily accounted for by an ad hoc rule that assigns a non-core I to final neutral vowels if preceded by a neutral vowel.

This solution, however, is unsatisfactory: a rule that assigns a non-core I to penult neutrals if followed by another neutral would do just as well. The observed tendency makes more sense if we suppose that the direction of the harmonic process is right to left, i.e. if archiphonemes are 'looking for' the features that complete them. This is especially intuitive in the case of i/ja: here the 3rd sg def suffix must check the stem to see if (22) is applicable, and this 'checking' should proceed backwards from the stem-final vowel.

The checking process can be described by a three-state finite automaton that scans the stem right to left. In the initial state, called B, the back alternant is selected. As the automaton moves backwards, it can encounter front, neutral, or back vowels. If it encounters a back vowel first, the process has ended: the automaton stays in the initial state and we get the back alternant. Likewise, if it encounters a front vowel, it moves to state F where the front alternant is selected. The third, or N state comes into play when a neutral vowel is encountered first: in this case the automaton stays in state N and investigates the preceding vowel.

The results of this investigation are evaluated in the same manner: if the vowel is front, the automaton moves to F, if it's back it moves to B and if if it's neutral, it stays in the N state *or goes to F* – it is this choice that gives the vacillating behavior. The more neutral vowels it encounters, the more likely the automaton is to fall into the F state (which seems to

[19] The reader should keep it in mind that the 'generalizations' presented below are all based on impressionistic judgements, rather than exhaustive testing.

[20] Neutral vowels will *not* be called 'front', their phonetic properties notwithstanding.

be the desired generalization), and if there are no more vowels to scan, it moves to F. (This is equivalent to saying say that in state N the front alternant is selected.) This description of the automaton will account for the first two generalizations without further stipulations, and the third one follows if we suppose that from N a neutral vowel will *necessarily* take the automaton to F.

Thus, vacillation corresponds to a probabilistic choice: for those speakers who maintain the third generalization, the choice is deteministic. But the chief advantage of scanning right-to-left is not that we can describe different speakers by changing a simple parameter (although this is certainly desirable) – the crucial argument for this direction comes from non-vacillating stems. Disharmonic roots of the *föderativ* 'federal' type always take back, while those of the *zsonglőr* 'juggler' type always take front suffixes. Thus, the decisive factor is the last non-neutral vowel, although its effect might be obscured if two or more neutral vowels follow. The automaton given above will work with disharmonic stems without any modifications – the comparable automaton that scans left-to-right would have to be extremely complex to handle disharmonic stems.

4 Conclusion

Stanley's (1967) objections against a 'ternary' use of binary features apply with even greater force in the framework of autosegmental phonology. Theoretically, in a context like

(23)

```
timing tier        C  V  C
                   |     |
F-tier            +F __ +F
```

the vowel in the middle can take place in a five-way opposition as given in (24):

(24)	(a)	(b)	(c)	(d)	(e)
$CV-tier$	V	V	V	V	V
	|	|			
$F-tier$	$+F$	$-F$		$+F$	$-F$

Here (a-c) are more or less standard, and lexical exceptions provide an important source of specifications like (d-e). Certain stems (like those in Class IV) must be marked in the lexicon, and it is an important task of autosegmental phonology to show that the lexical marking in such cases need not involve ad hoc rule exception features, but can be chosen from a restricted set of diacritics, namely that of floating features.

In this case, the theory forces the feature –U upon us for these stems, thereby accounting for the choice of *-et*, rather than **öt* in a straightforward manner. The *-höz* form could be derived only by positing some source other than the stem for the feature U in it: my proposal is that the U is present (as a floating feature) in the representation of *hoz/hez/höz*. This will also account for the lack of **haz*: if hVz receives no I or U specification from the preceding element, the 'default' U will automatically link up, and we get *-hoz*. (It was supposed that the A feature is present in the UR of every ternary and quaternary suffix.) But the use of negative feature values is strictly limited: in particular, we do *not* encounter minus-valued features in non-exceptional cases.

Unlike in Goldsmith's (1985) analysis, all three features are taken to represent privative oppositions: the absence of an autosegment is interpreted as negative specification for the given feature at every stage of the derivation. The tightness of the feature system leaves no room for underspecification – four features would give 16 possibilities, but with three features, 7 of the 8 possible combinations have to be taken 'at face value' in order to distinguish the vowels from each other.

The original aim of underspecification was to capture *archiphonemes* (i.e. underdifferentiated entities such as the harmonic pair *a/e*) that will be fully specified in the course of the derivation. This use of blanks is incompatible with the 'simplex' interpretation where the lack of specification is equated with negative specification. I suggest that archiphonemes are among the exceptional items inasmuch as they can also receive negative specification. For instance, if we take the dative suffix as underlying *-nak*, the alternant *-nek* can be derived by (8A): however, we have to block (8B) from applying. In other words the *a/e* archiphoneme has to be –U throughout the derivation – a convenient way of achieving this is by specifying it as –U in the core. This is compatible with the view that negatively specified autosegments do not spread but rather trigger the deletion of the following positively specified feature.

5 References

Antal L. 1959: Gondolatok a magyar főnév birtokos ragozásáról. *Magyar Nyelv* 55, 351-357

Antal L. 1963: The possessive form of the Hungarian noun. *Linguistics* 3, 50-61

Becker-Makkai, V. 1970: Vowel harmony in Hungarian reexamined in the light of recent developments in phonological theory. In: Becker-Makkai (ed): Phonological theory: Evolution and current practive. Holt, Rinehart, and Winston, New York, pp 634-648

Goldsmith, J. 1985: Vowel harmony in Khalkha Mongolian, Yaka, Finnish and Hungarian. *Phonology Yearbook* 2, 253-275

Halle, M.- J-R. Vergnaud 1982: On the framework of autosegmental phonology. In: v.d.Hulst-Smith (eds): The Structure of Phonological Representation. Foris, Dordrecht 65-82

v.d.Hulst, H. 1985: Vowel Harmony in Hungarian: a Comparison of Segmental and Autosegmental Analyses. In: v.d.Hulst – Smith (eds): Advances in Nonlinear Phonology. Foris, Dordrecht 267-303

Kaye. J. - Lowenstamm, J. - Vergnaud, J-R. 1985: Vowel systems. *GLOW Newsletter* No. 14.

Kontra, M. - Ringen, C. 1985: Stress and harmony in Hungarian loanwords. ms

Kontra, M. - Ringen, C. 1986: Hungarian vowel harmony: the evidence from loanwords. *Ural-Altajische Jahrbücher* 58, 1-14

Kornai, A. 1986: On Hungarian Morphology. Kandidátus Thesis, Instiute of Linguistics, Hungarian Academy of Sciences.

Nádasdy Á. 1985: Segmental phonology and morphophonology. In: Kenesei (ed): Approaches to Hungarian 225-246

Rácz E. 1974: A birtokos személyragozásnak a birtok többességét kifejező alakrendszere. In: Rácz E. – Szathmári I. (eds): Tanulmányok a mai magyar nyelv szófajtana és alakrendszere köréből. Tankönyvkiadó, Budapest. 135-151

Rennison, J. 1984: Tridirectional feature systems for vowels. *Wiener Linguistiche Gazette* 33-34

Siptár P. 1984: Vita a magyar magánhangzó- harmóniáról. *Magyar Nyelv* 80, 228-238

Stanley, R. 1967: Redundancy rules in phonology. *Language 41, 294-302*

Szépe Gy. 1969: Az alsóbb nyelvi szintek leírása. In: Károly S. – Telegdi Zs. (eds): Általános Nyelvészeti Tanulmányok VI, 359-466

Vágó, R.M. 1975: Hungarian generative phonology. IULC

Vágó, R.M. 1976: Theoretical implications of Hungarian vowel harmony. *Linguistic Inquiry* 7, 243-263

Vágó, R.M. 1980: The Sound Pattern of Hungarian. Georgetown University Press.