

# No lowering, only paradigms: a paradigm-based account of linking vowels in Hungarian<sup>1</sup>

*Bontási hulladékból ne építkezz!*

Szilágyi N. Sándor

There exist two basic approaches to grammatical, hence morphological phenomena (see Blevins 2016, Ackerman et al. 2009, 2016, Ackerman & Malouf 2016 on morphology; Kálmán 2007, 2014ab on grammar in general). One, typical of structuralist and especially generativist approaches, could be described as **atomistic**. This approach explains the patterns observed in languages by (de)composing them of elements that are themselves not further divisible. In morphology, this includes the Item and Arrangement or Item and Process models (cf. Hockett 1954) and other morpheme-based structuralist and generative approaches, where the primitives are the unambiguously demarcated morphs, morphemes, more recently morphomes (see Aronoff 1994). A theoretical consequence of this approach is a necessary distinction between the lexicon and the grammar. The lexicon contains all those properties of a linguistic element that influence its behaviour and that cannot be derived from other factors (cf. Bloomfield 1933). In these theoretical models grammar is essentially a set of rules or constraints applying to linguistic elements as input, determining their concatenation and realization, eventually producing the forms observable on the surface, i.e., these approaches hypothesize such input–output mappings and use manipulation techniques to derive surface patterns. As a result, the lexicon contains not only properties directly observable on the surface, but also abstract features or configurations in the underlying or intermediate representations that play a role in the constitution of surface forms.

---

<sup>1</sup> This paper makes an attempt at reconstructing and providing a context for László Kálmán's views on the interface of phonology and morphology. Like those of so many of his contemporary fellow linguists', Kálmán's first analyses in the 1980s were couched in the autosegmental framework. However, facing the limits of this and other similar approaches, he soon advocated the necessity of a grammar based on alternative grounds, not only in morphophonology, but in almost all other fields of the description of language (semantics and syntax) which he worked in. He was well acquainted with most of the models mentioned here (e.g., lexicalist and data-oriented/case-based approaches, analogy in grammar, construction grammar) and built them in his own theories. He published only a small part of his theoretical hypotheses in technical papers; the larger part he disseminated in personal communication and lectures, courses and summer schools of the Theoretical Linguistics Department jointly affiliated with Eötvös Loránd University and the Hungarian Academy of Sciences (MTA), as well as in popularizing papers aimed at a wider audience (e.g., Kálmán 2007, 2012, 2013, 2018cd, 2019a, 2020abcd, 2021ab, Kálmán & Forró 2014 for morphophonology, and Kálmán 2007, 2008, 2012, 2014ab, 2015, 2018ab, 2020e for grammar in general). In the introduction to this paper we give a summary of these alternative grammars echoing many of the views Kálmán expressed orally or in writing. The rest of the paper provides an analysis for a specific morphophonological phenomenon, which is strongly dependent on and indebted to László Kálmán's radical, uncompromising and novel approaches to Hungarian grammatical phenomena. Of course, we are taking the risk that he would not agree with all we have to say but believe that with the spirit he certainly would.

In the other approach, which is called **holistic** (Croft 2001, Crysmann & Bonami 2017, Kálmán 2007, Szilágyi N. 2004), the primitives are the (relatively) autonomous expressions (e.g., word, sentence) and their similarities are examined as manifested in directly observable linguistic data — as opposed to atomistic approaches, which focus on the properties of hypothesized underlying representations and/or formal objects hypothetically assigned to linguistic expressions. There are at least four essential features shared by the various models (lexicalist, data-oriented, case-based, exemplar-based, frequency-based, analogy-based or construction grammars, among others) that can fall under the holistic label: (i) no clear-cut separation of lexicon and grammar, (ii) surface-true generalizations, (iii) the *in-situ* character of analyses and (iv) the graduality of exceptionality and productivity.

(i) In a holistic model linguistic expressions are not necessarily configurations of partial expressions selected from the lexicon; instead (certain) expressions are the primitives of the theory in their entirety. In morphology, the word may be considered to be such a primitive concept, which is not decomposed into smaller constituents (e.g., morphemes): the parts that systematically recur in word forms emerge from the systematic patterns of the comparison of whole forms, but the boundaries of these parts are not necessarily clear-cut. For word-forms organized into paradigms this means that the paradigm is a primitive (and not an epiphenomenon as in morpheme-based theories). Functional and formal **analogical similarities** hold between surface forms organized into paradigms: Word and Paradigm/Pattern morphology (e.g. Blevins 2016, Blevins and Blevins 2009) is a morphological model of this type. These similarities mean partially identical surface forms. General and language-specific regularities of paradigmatic organization result from the discovery of these analogical patterns. By contrast, in atomistic models the systematic similarities between surface forms are derivative: they are explained by the (partial) identity of underlying forms stored in the lexicon (e.g., identical underlying representations of roots or affixes) and/or the (partial) identity of formal manipulations that apply to these forms, i.e., (morpho)phonological rules/constraints. One of the most important drawbacks of atomistic models is that generalizations which cannot be expressed by these representational and/or operational tools cannot be captured at all, even when they are clearly manifest in the relevant patterns of the surface forms (e.g., patterns motivated by the uniformity and contrast of whole surface forms, see Rebrus & Törkenczy 2005, Wunderlich 2012).

(ii) The requirement that generalizations should be surface true precludes the use of abstract elements or features that are unobservable at the surface: this means that any type of opacity (e.g. Kiparsky 1971, McCarthy 2007) is disallowed in (morpho)phonological analysis. This limits the abstractness of the analysis to a great extent: only those generalizations are licit that can be formulated as a subset of surface forms that is defined intensionally by properties visible on the surface. These properties may refer to the forms of words (e.g., consonant-final or monosyllabic) and/or their function/semantics defining paradigmatic patterns (syntactic categories, e.g., verbs, nominals or values of morphosyntactic features, e.g. plural and/or accusative forms or words that belong to a specific lexeme, i.e., members of the same paradigm). Atomistic approaches make extensive use of abstract features and specifications that are not surface true, and structures and formal mechanisms that manipulate them. In

holistic models these tools are absolutely disallowed, which makes these grammars more transparent and more plausible from a cognitive point of view (see Vihman 2013).

(iii) In holistic grammars the formal and functional properties of a construction are associated with the *entire* construction as a whole rather than (some of) its parts as opposed to the approach where such properties are projected from the lexicon onto a structure. This also constrains abstractness in morphophonological analysis<sup>2</sup> since this *in-situ* interpretation of patterns has far-reaching consequences: morphemes (stems or affixes) do not exist as abstract entities and properties that do not appear in their surface forms (if indeed they do occur in isolation) cannot be assigned to them. For instance, “stems” labelled with an “antiharmonic” feature (see Kiparsky & Pajusalu 2003, Nádasdy & Siptár 1994, Siptár & Törkenczy 2000, Rebrus et al. 2012, Rebrus and Törkenczy 2019) is not a valid concept in this model: the back harmonic suffixation in forms like *sírnak*, *sírról*, *sírok*, *sírunk*, *sírotok*,<sup>3</sup> etc. (‘grave-DAT, -DEL, -PL, -POSS.1PL, -POSS.2PL’) is not a property of their stem but the entire word form as a construction. The common attribute of these forms is that they contain back suffix alternants: the free form *sír* (the “stem”) in itself has no such property detectable at the surface (thus this property cannot be associated with the bare stem), it emerges from the relevant suffixed forms and thus is a property of its paradigm.<sup>4</sup> To put it in another way, these forms containing back suffix alternants are antiharmonic because we suffix them in this way (we have memorized them in this way) and not because their abstract “stem” is marked with some abstract “backness” feature in the lexicon. In this paper we will employ this view to the morphophonology of the vowel height (especially the mid~low) alternations in Hungarian (the “lowering” phenomenon).

(iv) Another important difference between the two grammatical approaches is how they treat “exceptionality” and “productivity”. Atomistic approaches concentrate mainly on patterns that they consider “productive”, that involve general (universal) rules/constraints, and stigmatize others as “exceptional” and outside the scope of the grammar. This view raises theoretical and empirical/methodological problems. On the one hand the usual definitions of productivity are too tight: there are relevant broad generalizations on large closed word-classes that cannot be productively expanded (e.g., monomorphemic verbal roots in Hungarian), and there are lower-level generalizations on closed subclass of open classes (e.g., lowering noun stems in Hungarian) that play a crucial role in grammar. On the other hand, the

---

<sup>2</sup> Naturally, this holds for syntax, too, where categories (parts of speech) are not primitives; contrary to the assumption of mainstream theories, the properties of syntactic categories may emerge from syntactic patterns, see Croft 2001, Kálmán 2021c2021c.

<sup>3</sup> We give Hungarian words in the standard orthography. The accent mark on vowels marks length. The vowels relevant in this article are: *e* [ɛ], *a* [a], *o* [o], *ö* [ø]. Some consonants are represented by digraphs: *cs* [tʃ], *sz* [s] (*s* [ʃ], *ssz* [s:]), *gy* [j], *ny* [ɲ].

<sup>4</sup> In a more radical interpretation the paradigm itself is not a primitive but emerges from functional/semantic similarity. It is an advantage of this view that it can easily account for the phenomenon that paradigms are not homogeneous and their boundaries are fuzzy. Paradigmatic behaviour is not always uniform: paradigms may include subparadigms of more closely related forms that behave more homogeneously. Some derivational forms (frequently with lexicalised meaning) containing certain derivational affixes often lie at the “boundaries” of a paradigm and sometimes pattern differently from, sometimes similarly to inflected forms; cf. the idea of extended paradigms, Steriade 2000. We adopt this idea in this paper: we take the paradigm to contain derivational forms in addition to inflectional ones and interpret it as a primitive of the theory for simplicity’s sake.

productive and exceptional behaviour is gradual, their border is not clear cut: some lower-level generalizations can be used in innovations. Atomistic theories often misuse the notion of exceptionality when they define the scope of the theory: if the theory works in the relevant “nonproductive” data then they analyse them, if it does not then they label “exceptional” phenomena as extragrammatical, often irrespectively of their frequency. Holistic theories, however, recognize the fact that linguistic patterns are at different levels of exceptionality/productivity, ranging from totally idiosyncratic behaviour through different generalizations of (variously) limited scope to the almost exceptionless linguistic patterns. In this view, frequency plays a crucial role (Bybee 2007): productive language behaviour can be predicted by analogical attraction, whose strength depends on the frequency and similarity of the relevant sets of linguistic expressions.

This paper has two aims: firstly we attempt to give a survey of the vowels that occur between what is traditionally called stem and suffix in Hungarian, and secondly we show why the patterns of the quality of these vowels that emerge require highly overcomplicated analyses in an atomistic, morpheme-based approach to morphophonology. We begin by giving a definition of the vowels under discussion (1) and introducing the factors that influence their height (§2). We next show some patterns of successive linking vowels within a word, and the assumptions a morpheme-based account of these patterns must make, together with the difficulties these assumptions face (§3). We then outline a paradigm-based account, which is based on comparing the initial and final substrings of whole words located in different cells of paradigms (§4). We show that the creation of new word forms leads to solving the paradigm cell filling problem (§4.4). The paper is concluded by a summary (§5).

## 1 Linking vowels

Some Hungarian words lend themselves to an unambiguous morphological parse<sup>5</sup>, as in (1a), while in others the assignment of phonological material to morphs is ambiguous and thus leaves behind a vowel that could be parsed in either or neither direction, as in (1b).

### (1) Morphological parsing

#### a. **unambiguous parse**

hajó-nk ‘ship-POSS.1PL’  
 erő-nk ‘force-POSS.1PL’  
 daru-k ‘crane-PL’  
 öl-tök ‘kill-2PL’  
 szó-tlan ‘word-less’  
 vér-t ‘harm-ACC’

#### b. **ambiguous parse**

baj-**u**-nk ‘trouble-POSS.1PL’  
 vér-**ü**-nk ‘blood-POSS.’  
 kar-**o**-k ‘arm-PL’  
 ölt-**ö**-tök ‘stitch-2PL’  
 pár-**a**-tlan ‘pair-less’  
 gyér-**e**-t ‘sparse-ACC’

---

<sup>5</sup> In a holistic grammar, morphological parse is an epiphenomenon, morphs can emerge from the paradigmatic patterns, but their borders might be fuzzy. We use the mainstream morpheme-based view here: a clear-cut parse means that an exponentless free form (the free “stem”) is followed by an exponent (the “suffix” allomorph) that only contains the phonological material which recurs associated with the same morphosyntactic value in the system.

The emboldened vowels in (1b) do not occur at the end of the stem when it is unsuffixed (*baj*, *vér*, *kar*, *ölt*, *pár*, *gyér*) and they do not appear before the same suffixes either after other stems, as the counterparts in (1a) show. The emboldened vowels in (1b) also occur in other forms of the paradigm of the relevant stems (e.g., the *o* of *karok* also occurs in *kar-o-m* ‘arm-POSS.1SG’, *kar-o-d* ‘-POSS.2SG’, *kar-o-s* ‘-ADJZ’, *kar-o-l* ‘-VRBZ’ etc.), so they could be analysed as part of the stem. However, they also regularly occur with the same suffixes with a large set of other stems (e.g., an *o* frequently occurs before the plural suffix in many plural nouns and adjectives), which would justify parsing them as part of the suffix. This is a problem for an atomistic approach, which intends to exhaustively and unambiguously assign phonemes to morphs. We remain agnostic of the morphological affiliation of such vowels, in fact, do not think this is relevant at all. Following the descriptive tradition, we will call them **linking vowels**.<sup>6</sup> Linking vowels will be identified by hyphens on both sides. We skip linking vowels in glosses.

The data in the last line of (1) and those in (2) show that the appearance of a linking vowel is not merely phonologically conditioned, rather it is governed by the interaction of complex morphological, phonological and lexical conditions (see Rebrus 2000a, Rebrus and Törkenczy 2008, 2009; Kálmán et al. 2012).

## (2) Linking vowels and their absence

	<b>linking vowel absent</b>	<b>linking vowel present</b>
a.	nyom-tok ‘push-2PL’	nyom-o-tok ‘trace-POSS.2PL’
b.	gond-nak ‘trouble-DAT’	mond-a-nak ‘say-NDF.3PL’
c.	fing-t-a ‘fart-PST-DEF’	fing-o-tt-a ‘id.’
d.	briliáns-t ‘brilliant-ACC’	briliáns-o-t ‘id.’
e.	fiú-k ‘boy-PL’	hiú-a-k ‘vain-PL’

A linking vowel may occur after the nominal stem but not after the homophonous verbal stem, as in (2a). However, this is not due to word class, in (2b) we see the opposite situation: the linking vowel occurs after the verbal stem but not after the near homophonous nominal stem. The next examples show that the same stem, verbal in (2c) or nominal in (2d), may exhibit both types of behaviour in free variation.<sup>7</sup> A linking vowel may occur after a vowel-final stem, too, as in (2e).

Six of the seven short vowels of Hungarian (*u*, *ü*, *o*, *ö*, *a*, *e*, but not *i*) may function as a linking vowel.<sup>8</sup> The quality of the linking vowel is partly affected by vowel harmony: we find

<sup>6</sup> One can define a linking vowel as a vowel that alternates with zero with a specific suffix, but never occurs in the exponentless form of the same paradigm (see Kálmán et al. 2012). The advantage of this definition is that it does not refer to the morphological segmentation into “stem” and “suffix”. This definition underdetermines the status of vowels in some cases, especially if there is no exponentless free form (“free stem”) in a paradigm. We think this is not a serious disadvantage, because the status of being a “linking vowel” is a gradual concept — see the discussions in §§4.2 and 4.4.

<sup>7</sup> For the vowel–zero alternation in past forms, see Trón & Rebrus 2005; for the vowel–zero hesitation in accusative, see in Kálmán et al. 2012; for vowel–zero alternation within stems, see Rácz et al. 2021.

<sup>8</sup> It is possible to analyse the *i* in the suffixed infinitive forms *ad-n-i-[j]a* ‘give-INF-3SG’, *ad-n-i-[j]uk* ‘-3PL’ as a linking vowel too, where the quality of the “linking” *i* is due to suffix-initial glide [j] (cf. the “normal” mid linking vowels in other persons of the infinitival forms: *ad-n-o-m* ‘-1SG’, *ad-n-o-d* ‘-2SG’, etc.). We leave this

*ü, ö, e* after front and *u, o, a* after back harmonic stems. The occurrence of *u* and *ü* as linking vowels is limited to the first person plural possessive suffix and the verbal nondefinite first person plural suffix.<sup>9</sup> The choice between mid *ö* and low *e*, on the one hand, and mid *o* and low *a*, on the other, is governed by an interplay of phonological and morphological factors, traditionally referred to as **lowering**.

## 2 Lowering

The term lowering is used (e.g., Vago 1980, Szentgyörgyi 1999, Siptár and Törkenczy 2000) to identify and account for a set of morphophonological *patterns* and their distribution. The patterns are phonologically rather disparate. Lowering manifests itself in the height of linking vowels: the alternation of mid *o* vs. low *a* and mid *ö* vs. low *e* [ɛ].<sup>10</sup>

We have seen in §1 above that linking vowels are all the short vowels other than *i*. In the lowering pattern in the reference accent the linking vowel is *a* with a back harmonic base (*ház-a-k* ‘house-PL’ vs. *gáz-o-k* ‘gas-PL’) or *e* with a front base where the harmonic trigger is a rounded vowel (*őz-e-k* ‘roe-PL’ vs. *gőz-ö-k* ‘vapour-PL’). As a result, when lowering occurs, the harmonically alternating linking vowels are *a~e* (*ház-a-k, őz-e-k*) and not the other possible harmonic alternations involving linking vowels: *o~e~ö* (*gáz-o-k, géz-e-k* ‘gauze-PL’, *gőz-ö-k, ház-o-n* ‘house-SUE’, *géz-e-n* ‘gauze-SUE’, *őz-ö-n* ‘deer-SUE’) or *u~ü* (*gáz-u-nk, géz-ü-nk, gőz-ü-nk, ház-u-nk, őz-ü-nk* ‘-POSS.1PL’).<sup>11</sup>

With a few exceptions any stem may be followed by a low, mid, or high vowel, be it nominal, verbal, or adjectival.<sup>12</sup> Suffixes come in four flavours with respect to lowering on their left side.

---

issue open here, but see Siptár 2011 for a morpheme-based analysis and Kálmán & Rebrus 2009 for a frequency-based one.

<sup>9</sup> The status of the *u/ü* occurring with the *third* person plural possessive affix and in *definite* first person plural verbal suffix is debatable. This high vowel occurs after many consonant-final stems (*dal-uk* ‘song-POSS.3PL’), but not in the possessive form of plural nouns where the stem is vowel final (*dal-a-i-k* ‘-POSS-PL-3PL’), and also for verbs (cf. *für-t-uk* ‘drill-PST-DEF.1PL’ and *%für-nó-k* ‘-COND-DEF.1PL’). We do not take sides on the morphological parsing of these forms.

<sup>10</sup> We focus on vowel quality here and refer the reader interested in other phenomena associated with lowering to Siptár & Törkenczy 2000, Kálmán et al. 2012, Rebrus 2019, Rebrus et al. 2017, Rebrus & Szigetvári 2022.

<sup>11</sup> The term lowering is a partial misnomer in the reference accent: *a* is indeed lower than *o* (*ház-a-k* PL’ vs. *gáz-o-k*) and *e* [ɛ] is lower than *ö* [ø] (*őz-e-k* vs. *gőz-ö-k*), but when the harmonic trigger is a front unrounded vowel, the low *e* indistinguishably occurs both in otherwise lowering and nonlowering environments (*het-e-d-e-t* ‘seven-FRAC-ACC’ and ‘-POSS.2SG-ACC’ vs. *hat-o-d-o-t* ‘six-FRAC-ACC’, but *hat-o-d-a-t* ‘-POSS.2SG-ACC’: with lowering after the 2sg possessive, but not after the fractional suffix with back stems), since an *\*e/é/i/í...ö* sequence is not possible in this environment due to labial harmony (see Rebrus & Törkenczy 2008). Since in the case of front unrounded stems, labial harmony obscures the distinction that would be the effect of lowering, we temporarily ignore such morphemes in the discussion. We will return to them in §4.1.

<sup>12</sup> The exceptions include the adjective *nagy* ‘big’, which is never followed by a lowered vowel and almost all adpositions which can only be followed by person suffixes (which are indifferent about the lowering status of the linking vowel that precedes them) and are thus followed by a lowered vowel: *nál-a-m* ‘at me, lit. ADE-1SG’, *től-e-m* ‘from me, lit. ABL-1SG’.

(i) The linking vowel is always low before the conditional (*-nA*),<sup>13</sup> the infinitive (*-n(i)*), the third person plural nondefinite (*-nAk*), the second person singular nondefinite (*-sz*), and the first person singular second person object (*-lAk*) verbal suffixes, the privative (*-tLA*n), as well as the comparative suffix (*-bb*).<sup>14</sup>

(ii) The superessive and the nondefinite third person singular verbal suffix (which are incidentally homophonous: *-n*), on the other hand, are never preceded by a low linking vowel (recall that at this point we disregard front unrounded stems). This is also true of the past tense suffix (*-t(t)*), the second person singular nondefinite suffix (*-l*),<sup>15</sup> and the frequentative *-gAt*, because the linking vowel only occurs before them when they are only added to uninflected verb stems, which never force lowering.

(iii) The third group of suffixes may be preceded either by a low or a mid linking vowel, as determined by the preceding stem. This group includes the plural (*-k*), the accusative (*-t*), various possessive, verbal (1SG *-m*, *-k*, 2SG *-d*, 2PL *-tOk*) and derivational suffixes (verbalizers *-z*, *-l*, adjectivizer *-s*, distributives *-nként*, sociative *-stUl*, diminutive *-cskA*, adverbializer *-n*).

(iv) Only the verbal and the possessive first person plural suffixes, and debatably the third person possessive, as well as the first and third person definite verbal suffixes belong to the fourth type, these are preceded by a high linking vowel: *u* or *ü*.

With respect to their right side, the past tense suffix *-t(t)*, and the adjectivizer suffixes *-i*, *-Ú*, *-Ó* always have a low linking vowel (the adpositions mentioned in footnote 12 also belong here). No suffix is always followed by a mid or high linking vowel. The rest of the suffixes and roots may be followed by low or mid linking vowels. We summarize these possibilities in (3). Suffixes are emboldened when they can be preceded *and* followed by a linking vowel. We shall refer to group (iii) as *conformist* suffixes and to the other groups, (i), (ii), and (iv), as **nonconformist** ones (and the linking vowel before these suffixes as resistant).

---

<sup>13</sup> Following the tradition, we use capitals as shorthand to indicate harmonically alternating vowels: *A*=*a/e*, *O*=*o/ö/e*, *U*=*u/ü*, *Ó*=*ó/ő*, *Ú*=*ú/ü*.

<sup>14</sup> With the exception of the stem *nagy* mentioned in the previous footnote: *nagy-o-bb* ‘big-CMPR’. To avoid overcomplicating the table we choose to ignore this word, since its inclusion in (3) would entail splitting the “variable” group in two with the comparative suffix forming a separate group, which is preceded by a lowered linking vowel after all stems except *nagy* (which itself would be located as the only stem in the “mid after” cell).

<sup>15</sup> The two second person singular nondefinite suffixes, *-sz* and *-l*, are by and large phonologically conditioned allomorphs exhibiting a great deal of variability in different dialects.

(3) Linking vowels before and after suffixes and after stems

	(i) low	(iii) variable	(ii) mid	(iv) high
b e f o r e	<b>CMPR</b> ↓ - <i>bb</i> <b>PRIV</b> - <i>tlAn</i>	<b>PL</b> ↓ - <i>k</i> , <b>ACC</b> - <i>t</i> , <b>ADVZ</b> - <i>n</i> , <b>POSS</b> : <b>1SG</b> ↓ - <i>m</i> , <b>2SG</b> ↓ - <i>d</i> , <b>2PL</b> ↓ - <i>tOk</i> <b>ADJZ</b> ↓ - <i>s</i> , <b>FRAC</b> - <i>d</i> , <b>VBZ</b> - <i>z</i> , - <i>l</i> <b>DISTR</b> - <i>nként</i> , <b>SOC</b> - <i>stUl</i> , <b>DIM</b> - <i>cskA</i>	<b>SUE</b> - <i>n</i>	<b>POSS.1PL</b> ↓ - <i>nk</i> ? <b>POSS.3PL</b> ↓ - <i>k</i>
	<b>INF</b> - <i>n(i)</i> <b>NDF.2SG</b> - <i>sz</i> <b>NDF.3PL</b> - <i>nAk</i> <b>2&lt;1SG</b> - <i>lAk</i> <b>COND</b> - <i>nA</i>	<b>NMZ</b> ↓ - <i>dAl(O)m</i> <b>1SG</b> - <i>m</i> <b>NDF.1SG</b> - <i>k</i> <b>DEF.2SG</b> - <i>d</i> <b>2PL</b> - <i>tOk</i>	<b>NDF.3SG</b> - <i>n</i> <b>NDF.2SG</b> - <i>l</i> <b>PST</b> ↓ - <i>t(t)</i> <b>FREQ</b> - <i>gAt</i>	<b>1PL</b> - <i>nk</i> ? <b>DEF.1PL</b> - <i>k</i> ? <b>DEF.3PL</b> - <i>k</i>
a f t e r	<b>ADJZ</b> ↓ - <i>i</i> <b>ADJZ</b> ↓ - <i>Ú</i> <b>ADJZ</b> ↓ - <i>Ó</i> adpositional roots↓	<b>PL</b> ↓ - <i>k</i> , <b>ADJZ</b> ↓ - <i>s</i> , <b>POSS</b> : <b>1SG</b> ↓ - <i>m</i> , <b>2SG</b> ↓ - <i>d</i> , <b>1PL</b> ↓ - <i>nk</i> , <b>2PL</b> ↓ - <i>tOk</i> , <b>CMPR</b> ↓ - <i>bb</i> , <b>PRIV</b> - <i>tlAn</i> , <b>FRAC</b> - <i>d</i> <b>VBZ</b> - <i>z</i> , - <i>l</i> , all nominal/adjectival roots(↓)		
	<b>PST</b> ↓ - <i>t(t)</i>	<b>SBJV</b> ↓ - <i>j</i> -, <b>NMZ</b> ↓ - <i>dAl(O)m</i> % <b>INF</b> - <i>n</i> -, <b>FREQ</b> - <i>gAt</i> , <b>POT</b> - <i>hAt</i> all verbal roots		

Some stems and suffixes in the “variable” column in (3) have a greater inclination for being followed by a low vowel than others (this is indicated by a downwards arrow). Therefore, although they are all labelled as variable in (3), they have to be distinguished. In the literature they are referred to as “lowering stems” and “lowering suffixes” (Vago 1980, Szentgyörgyi 1999, Siptár & Törkenczy 2000). However, they are *not* followed by a low vowel in *all* situations (compare *ház-a-t* ‘house-ACC’ vs. *ház-o-n* ‘house-SUE’, *áld-j-a-k* ‘bless-SBJV-NDF.1SG’ vs. *áld-j-o-n* ‘bless-SBJV-NDF.3SG’). And vice versa, “nonlowering stems” are followed by a low vowel in *some* situations (compare *gazdag-o-k* ‘rich-PL’ vs. *gazdag-a-bb* ‘rich-CMPR’; *bont-o-k* ‘unwrap-NDF.1SG’ vs. *bont-a-sz* ‘unwrap-NDF.2SG’).

Although we do assign morphemes to the “low after” cell in (3) on descriptive grounds, none of these have to be *analysed* as morphemes that carry a specification for *obligatorily* having a low linking vowel to their right at the surface, that is, morphs that force lowering after them. On the one hand, neither the past tense suffix, nor adpositions are ever followed by a further suffix that requires a preceding mid linking vowel, thus these suffixes are followed by a low linking vowel in all their presuffixal occurrences. On the other, the vowel-final adjectivizers -*i*, -*Ú*, -*Ó* are only followed by a linking vowel before conformist suffixes, that are indifferent to lowering, but they do not have a linking vowel before the superessive, which would inhibit lowering (cf. *ház-i-a-k* ‘house-ADJZ-PL’, *vonz-ó-a-k* ‘attract-ADJZ-PL’, *kis-láb-ú-a-k* ‘small-foot-ADJZ-PL’ vs. *ház-i-n*, *vonz-ó-n*, *kis-láb-ú-n* all three ‘-SUE’). This means that all the morphemes in the “low after” cell of (3) could be categorized as “variable after” that just happen to be never attested with a following mid linking vowel due to independent morphotactic or phonological reasons. To put it in another way, the fact that these suffixes are always followed by a low linking vowel at the surface is due to a conspiracy between their preference for a right flanking low linking vowel and the morphotactic constraint on suffixes



that could overwrite this preference. Furthermore, there are suffixes whose right preference is *in principle* not knowable for morphotactic reasons: these are suffixes that cannot be followed by another suffix (e.g., the case suffixes, including the accusative<sup>16</sup> and superessive, and all verbal person/number suffixes). That is, suffixes may force the height of the linking vowel before them, but suffixes (and also roots – with the exception of *nagy*, see footnote 12) may only have a bias about the height of the linking vowel after them. Whether this bias prevails or not depends on the suffix that follows them.

### 3 Linking vowels — a morpheme-based view

In an atomistic approach, where the low/mid character of a linking vowel is associated with morphs, a stem falls in one of two categories: it may be nonlowering, meaning that before *some* suffixes that immediately follow it and require a linking vowel this vowel is mid (*o~ö*), or lowering, meaning that before the same suffixes this vowel is low (*a~e*). The latter status is indicated by the label “(low)” in (4b, e–g). Since, as we have just seen, this specification turns out to be the underdog (i.e., it does not always prevail in the surface form), as proven in (4b, e), it is parenthesized. Suffixes, on the other hand, are of four types: some are invariably preceded by a high vowel (“**high**” in (4a, b)), some by a mid vowel (“**mid**” in (4d, e)), some by a low vowel (“**low**” in (4g, h)), with all three specifications in boldface because they always prevail in case of a conflict). Most suffixes with a linking vowel belong to the fourth type: they are conformists (“—” in (4c, f)), that is, they obey what the preceding stem dictates. The eight resulting possibilities are all illustrated in (4).

---

<sup>16</sup> In the “repetitively suffixed” form %*ő-t-e-t* ‘he/she-ACC-ACC’ the accusative suffix has a low vowel to its right. Note though that this word can also be interpreted as %*ő-t-e-t* ‘he/she-ACC’, i.e., a lowering root, *őt*, and a single suffix.

(4) Biases after stems and requirements before suffixes

	Examples	morpheme specifications		LINKING VOWEL
		STEM	SUFFIX	
a.	<i>tag-u-nk</i> ‘member-POSS.1PL’ <i>áld-u-nk</i> ‘bless-NDF.1PL’	—	<b>high</b>	high
b.	<i>fog-u-nk</i> ‘tooth-POSS.1PL’ <i>áld-j-u-nk</i> ‘bless-SBJV-NDF.1PL’	(low)	<b>high</b>	high
c.	<i>tag-o-d</i> ‘member-POSS.2SG’ <i>hat-o-d</i> ‘six-FRAC’ <i>áld-o-d</i> ‘bless-DEF.2SG’	—	—	mid
d.	<i>tag-o-n</i> ‘member-SUE’ <i>hat-o-n</i> ‘six-SUE’ (% <i>vagy-o-n</i> ) ‘be-NDF.3SG’	—	<b>mid</b>	mid
e.	<i>fog-o-n</i> ‘tooth-SUE’ <i>nyolc-o-n</i> ‘eight-SUE’ <i>áld-j-o-n</i> ‘bless-SBJV-NDF.3SG’	(low)	<b>mid</b>	mid
f.	<i>fog-a-d</i> ‘tooth-POSS.2SG’ <i>nyolc-a-d</i> ‘eight-FRAC’ <i>áld-j-a-d</i> ‘bless-SBJV-DEF.2SG’	(low)	—	low
g.	<i>tág-a-bb</i> ‘broad-CMPR’ <i>áld-j-a-nak</i> ‘bless-SBJV-NDF.3PL’	(low)	<b>low</b>	low
h.	<i>vak-a-bb</i> ‘blind-CMPR’ <i>áld-a-nak</i> ‘bless-NDF.3PL’	—	<b>low</b>	low

We can see that the linking vowel is mid if neither the stem, nor the suffix sets up any specific requirement about its height, as in (4c) (the hypothesis that the default height of the linking vowel is mid is mainly based on this), or if the suffix explicitly requires a mid linking vowel to precede it, as in (4d, e). The linking vowel is low if the stem prefers this, as in (4f, g), and/or the suffix requires it, as in (4g, h). The linking vowel is high if the suffix requires this, as in (4a, b).<sup>17</sup>

The same strategies can be witnessed in doubly suffixed word forms. In this case the first suffix creates a stem which is either lowering or nonlowering depending on the category of this suffix. The relevant patterns are summarized in (5). The plural, the second person singular, the first person plural possessive, and the past tense suffixes, for example, prefer lowering on their right flank, but this only prevails if the second suffix does not require some other linking vowel. This is the case in the accusative (*tag-u-nk-a-t*, *tag-o-k-a-t*, *hat-o-d-a-t*,

<sup>17</sup> Alternative analyses are also possible. For instance, we can specify nonlowering stems as having a (mid) feature, and leave lowering stems unspecified. In this case the default linking vowel is low, and it is mid (or high) only when it takes mid specification from the stem (from the left) or from the suffix (from the right) — see Rebrus & Polgárdi 1997 for this type of approach. These two analyses can be treated as formally equivalent, thus all their theoretical consequences are the same. Carried to the extreme, a mixture of the two approaches is also possible. In such an analysis, both the lowering and nonlowering morphemes are specified. Then, no default is needed but the price to pay is that both the lowering and the nonlowering character of a stem must be allowed to be overridden by the mid or the low specification of the following morpheme.

*fog-a-k-a-t*, or *nyolc-a-d-a-t*) or in the first person singular suffix (*áld-o-tt-a-m*). A high linking vowel is required by the possessive (*hat-o-d-u-nk*) or verbal 1pl suffix (*áld-o-tt-u-nk*). The second linking vowel is mid if this is what the second suffix requires, as the superessive suffix (*tag-u-nk-o-n*, *tag-o-k-o-n* or *fog-a-k-o-n*), but also if the first suffix does not prefer lowering, as the fractional (*hat-o-d-o-t*, *nyolc-a-d-o-t*), the frequentative (*áld-o-z-o-m*), or the infinitive (*áld-a-n-o-m*) and the second suffix does not prescribe a low vowel on its left. The second linking vowel is low if this is what the second suffix requires even if the first suffix would otherwise not be followed by a low vowel (as the first person singular second person object suffix in the partially accepted form %*áld-a-n-a-lak*, cf. *áld-a-n-o-m*, where the 1sg suffix has no left requirement).

(5) Patterns of linking vowels in doubly suffixed forms<sup>18</sup>

linking V: 2nd → 1st ↓	mid -o-	low -a-	high -u-
high -u-	<i>tag-u-nk-o-n</i>	<i>tag-u-nk-a-t</i>	—
mid -o-	<i>tag-o-k-o-n</i> <i>hat-o-d-o-t</i> <i>áld-o-z-o-m</i>	<i>tag-o-k-a-t</i> <i>hat-o-d-a-t</i> <i>áld-o-tt-a-m</i>	<i>tag-o-z-u-nk</i> <i>hat-o-d-u-nk</i> <i>áld-o-tt-u-nk</i>
low -a-	<i>fog-a-k-o-n</i> <i>nyolc-a-d-o-t</i> <i>áld-a-n-o-m</i>	<i>fog-a-k-a-t</i> <i>nyolc-a-d-a-t</i> % <i>áld-a-n-a-lak</i>	<i>fog-a-z-u-nk</i> <i>nyolc-a-d-u-nk</i> <i>áld-a-n-u-nk</i>

In an atomistic account morphophonological patterns derive from properties associated with or encoded into the representations of parts (morphemes) concatenated into complex forms (words). Let us collect the ingredients needed for such an account to produce the patterns just described. The charts presented in (3) and (4) suggest that morphs must be specified for lowering in three ways: (i) for the height of the linking vowel: low, mid, or high; (ii) for the direction of the linking vowel with respect to the morpheme that “carries” it (there is only one possible direction in the case of roots, of course); and (iii) the “strength” of this specification: we have seen that if the specification of adjacent morphemes clash (one requires a low, the other a mid or high linking vowel), the one on the right wins.

Let us illustrate the distinctions that must be made in a morpheme-based analysis. There are many ways of encoding these in underlying representations (full specification, underspecification, floating features, etc.) employing various operational tools (rules, constraints, general or parochial conditions on rule application, etc.). Suppose that we represent the height of linking vowels autosegmentally, by floating features (cf. Kálmán p.c., Kornai 1989). Then, we need a floating low (lexically part of, e.g., the infinitive or the 2sg verbal suffixes), a floating mid (for, e.g., the superessive, the nondefinite 2sg, or 3sg verbal suffixes, *-l* and *-n*, respectively), as well as a floating high feature (for the 1pl suffixes). The

<sup>18</sup> Glosses, line 1: ‘member-POSS.1PL-SUE’, ‘-ACC’; line 2: ‘member-PL-SUE’, ‘-ACC’, ‘member-VRBZ-NDF.1PL’; line 3: ‘six-FRAC-ACC’, ‘-POSS.2SG-ACC’, ‘-FRAC-POSS.PL’; line 4: ‘bless-FREQ-1SG’, ‘-PST-1SG’, ‘-NDF.1PL’; line 5: ‘tooth-PL-SUE’, ‘-ACC’, ‘-VRBZ-NDF.1PL’; line 6: ‘eight-FRAC-ACC’, ‘-POSS.2SG-ACC’, ‘-FRAC-POSS.1PL’; line 7: ‘bless-INF-1SG’, ‘-2<1SG’, ‘-1PL’.

problem is that a floating feature is by definition not anchored, which means that the direction of lowering cannot be encoded by standard autosegmental means. For example, if we assign a floating low feature to the plural or the infinitive suffix we cannot control that lowering affects the linking vowel *after* the plural and not before it (*tag-o-k-a-t*, not *\*tag-a-k-o-t*), but the one *before* the infinitive suffix and not after it (*áld-a-n-o-m*, not *\*áld-o-n-a-m*). In fact, the same difficulty arises in the case of stem-internal vowel~zero alternations. The relevant stems show this kind of alternation internally in their final syllable, e.g., *lator* ‘rogue’~*latr-o-k* ‘rogue-PL’. Examples exhibiting some of the vowel patterns in (5) are shown in (6).

(6) The height of unstable vowels within the stem and after the stem before conformist suffixes

unstable V: in suffix → in stem ↓	mid linking vowel -o-	low linking vowel -a-
<b>high -u-</b>	<i>bajusz</i> – <i>bajsz-o-k</i> ‘moustache, -PL’	—
<b>mid -o-</b>	<i>lator</i> – <i>latr-o-k</i> ‘rogue, -PL’	<i>farok</i> – <i>fark-a-k</i> ‘tail, -PL’
<b>low -a-</b>	<i>vacak</i> – <i>vack-o-k</i> ‘junk, -PL’	<i>kazal</i> – <i>kazl-a-k</i> ‘stack, -PL’

The vowel height in the singular and plural forms of *lator* and *kazal* can be accounted for by a stem-final floating mid or low feature, respectively, but when the alternating vowel within the stem and the linking vowel after it are different, as in the rest of the examples, we are facing the same problem as in the case of successive suffixes above: if we posit a floating low specification to get the alternating low vowel in *vacak*, we cannot explain the mid linking vowel in *vack-o-k*, and vice versa, if we posit a floating mid specification for the alternating stem internal vowel in *farok*, we cannot explain the low linking vowel in *fark-a-k*. If we were to posit *two* floating autosegments (a sequence of low–mid for *vacak* and mid–low for *farok*) there is still no autosegmentally canonical way (barring completely ad hoc stipulations) to make sure that the first one docks on the stem-internal alternating vowel and the second one on the following linking vowel. Thus, the number of different representations and the stipulations necessary to handle the direction of docking makes this type of analysis overly complicated.

Furthermore, an explanation of why it is the right-hand side floater that prevails in cases of distinct, and therefore incompatible floating features also has to be stipulated. We need a low feature to float at the end of a lowering stem, e.g., a plural stem, to explain why the linking vowel after it is low (*tag-o-k-a-t*). However, this floating low feature is suppressed when a suffix follows with a floating mid feature (*tag-o-k-o-n*). If we identify the lowering stem or the lowering suffix as the source of lowering, then we have to account for why this effect is not observed before a suffix that inhibits lowering before itself. There must also be a way to inhibit the low feature associated with the plural suffix from docking before it when it cannot dock after it (*\*tag-a-k-o-n*). Thus, in addition to the appropriate representations, some operational device is also needed to account for the interaction (the dominance of the specification of suffix on the right in the case of incompatibility). It is also curious that while

a mid linking vowel is the default (this is what appears if neither the left, nor the right morph has any specification), the same mid quality must also be specified on some morphs, in which case it prevails over the low feature, which was assumed not to be the default.

In footnote 11 we have already mentioned that labial harmony obscures lowering patterns. The fact that a linking vowel can only be front rounded if the immediately preceding vowel is also front rounded entails that the linking vowel is low *e* [ɛ] whenever the last vowel of the stem is front unrounded.<sup>19</sup> The linking vowel is mid in *gáz-o-k* ‘gas-PL’ and *gőz-ö-k* ‘vapour-PL’, but it can only be low in *géz-e-k* (\**géz-ö-k*) ‘gauze-PL’. This effect does not only lower the default mid vowel, but also the “obligatorily” mid vowel enforced by the superessive or 3sg suffixes: *géz-e-n* ‘-SUE’ (\**géz-ö-n* vs. *gőz-ö-n* ‘vapour-SUE’), *kér-j-e-n* ‘ask-SBJV-3SG’ (\**kér-j-ö-n* vs. *tör-j-ö-n* ‘break-SBJV-3SG’). Morpheme-based representational analyses of this pattern typically involve some form of absolute neutralization (e.g. Vago 1980, Siptár & Törkenczy 2000), resulting in a high degree of opacity, that is, they are not surface true, cf. footnote 27.<sup>20</sup>

Due to their design, morpheme-based approaches also miss a number of generalizations about aspects of lowering that (in these analyses) involve underlying representations which are unrelated in the sense that they do not exhaustively contain the same morphemes. These include generalizations that can be made between lowering and the absence/presence of yod in 3sg possessive suffixes *-(j)a~(j)e*, cf. Rebrus et al. 2017, 2022, 2023b (see a short discussion in §4.44 below); and other paradigmatic effects unrelated to lowering that cannot be explained on the basis of individual morphemes, e.g., the role of Harmonic Uniformity in front/back harmony (cf. Rebrus & Törkenczy 2019, 2021, Rebrus et al. 2020, 2023ab), paradigm conservatism and its interaction with phonotactics conditioning the nonoccurrence of certain forms of “defective” verbs (cf. Rebrus & Törkenczy 2009), homophony avoidance, i.e., the preservation of paradigmatic contrast and uniformity motivating some apparent exceptions in the verbal paradigm (cf. Rebrus 2000b, Rebrus & Törkenczy 2005) and the allomorphy (specifically, the occurrence of the linking vowel) of the past suffix *-t(t)*, (cf. Trón & Rebrus 2005), and the accusative suffix *-t* (Rebrus 2019). These topics extend beyond the scope of the present paper.

<sup>19</sup> To be precise, it can also be *back* mid after an antiharmonic stem (e.g., *sír-o-k* ‘grave-PL’, *cél-o-k* ‘goal-PL’) or a vowel transparent to front~back harmony (e.g., *tapír-o-k* ‘tapir-PL’, *acél-o-k* ‘steel-PL’). This is independent of and irrelevant for the phenomenon we discuss here.

<sup>20</sup> The same criticism applies to versions of this analysis which avoid positing an abstract underlying vowel by underspecifying one of the two *e*’s (e.g., Siptár and Törkenczy 2000). Some varieties of Hungarian indeed distinguish a mid and a low front unrounded vowel, the distribution of which is very similar to what follows from Siptár’s analysis. That is, the analysis may be imitating the historical development of the reference accent, but this in itself is no justification for it.

## 4 A paradigm-based account

### 4.1 Basic tools

A holistic, paradigm-based analysis has the following tools available to account for the patterns observed: (i) the paradigms that related full word-forms are organised into (a paradigm may be — as in our case — an extended one, which includes all forms of the inflection–derivation continuum), (ii) the (lexical) paradigm classes partitioning the paradigms, (iii) the paradigm cells defined by morphosyntactic categories, and (iv) the analogical similarities (partial identities) between the word forms assigned to the cells. These are the fundamental concepts stipulated by the theory. This approach entails that the complex patterns of the quality of linking vowels emerge from paradigm organization without need for arbitrarily stipulated objects (features, morphemes) and overcomplicated grammatical processes (direction sensitive resolution of conflicting features, complex interaction with labial harmony, multiple default qualities for vowels).

To facilitate comparison with the previous section, for the sake of simplicity we will informally keep using the labels “lowering” and “nonlowering”. There are two kinds of nominal/adjectival paradigm classes illustrated in (7): lexically nonlowering (e.g., *vak* ‘blind’, *hős* ‘hero’) and lexically lowering (e.g., *tág* ‘wide’, *zöld* ‘green’). Any stem affixed by a lowering suffix becomes a lowering stem (i.e., belongs to the lowering paradigm class): e.g., *old* ‘solve’ and *küld* ‘send’ are not lowering stems, but *old-j* ‘solve-SBJV’ *küld-j* ‘send-SBJV’ are lowering stems.<sup>21</sup> (The suffix *-n* is ‘SUE’ after nominal and adjectival stems and ‘NDF.3SG’ after verbal stems: *vagy-* ‘be’, *%mögy-* ‘go’.)

(7) The mutual dependence of linking vowels (back and front rounded harmonic classes)

	<b>low</b> (nonconformist suffix)	<b>stem dependent</b> (conformist suffix)	<b>mid</b> (nonconformist suffix)
<b>nonlowering paradigm</b>	<i>vak-a-bb</i> <i>hős-e-bb</i> <i>old-a-nak</i> <i>küld-e-nek</i>	<i>vak-o-k</i> <i>hős-ö-k</i> <i>old-o-k</i> <i>küld-ö-k</i>	<i>vak-o-n</i> <i>hős-ö-n</i> <i>%vagy-o-n</i> <i>%mögy-ö-n</i>
<b>lowering paradigm</b>	<i>tág-a-bb</i> <i>zöld-e-bb</i> <i>old-j-a-nak</i> <i>küld-j-e-nek</i>	<i>tág-a-k</i> <i>zöld-e-k</i> <i>old-j-a-k</i> <i>küld-j-e-k</i>	<i>tág-o-n</i> <i>zöld-ö-n</i> <i>old-j-o-n</i> <i>küld-j-ö-n</i>

<sup>21</sup> The same holds for nominal stems, too, *só* ‘salt’, *barna* ‘brown’, *nő* ‘woman’ are nonlowering stems, but *só-k-a-t* ‘-PL-ACC’, *barná-bb-a-k* ‘-CMPR-PL’, *nő-m-e-t* ‘-POSS.1SG-ACC’ exhibit lowering. Vice versa, *láb* ‘leg’, *fal* ‘wall’ are lowering stems, but *láb-a-tlan-o-k* ‘-PRIV-PL’, *fal-a-z-o-k* ‘-VRBZ-NDF.1SG’ do not exhibit lowering, due to their penultimate suffix.

The data in (7) show what we have already seen in §3: in some cases the linking vowel is determined only by the suffix (in the first and the last columns), while in others it is mutually dependent on the stem and the suffix (in the middle column).

In a paradigm-based analysis we only have to stipulate the following distinctions:

(i) (extended) paradigms are of two types (“lowering” or “nonlowering”); (ii) the exponents of whole word-forms associated with a specific morphosyntactic value<sup>22</sup> are either sensitive (“conformist suffixes”) or resistant to this difference (“low linking vowel suffixes” and “mid linking vowel suffixes”).<sup>23</sup> We have seen above that these lexical distinctions between the different paradigm classes on the one hand and the exponents of the different morphosyntactic values on the other must be made in a morpheme-based analysis, too (as stem type and suffix type, respectively), since the distinctions cannot be derived from the morphological, categorial, or morphophonological properties of the relevant morphemes. In a morpheme-based approach these characteristics are represented by hypothesized features (phonologized or diacritical) assigned to morphemes in the lexicon. The crucial difference between this and a paradigm-based account is that in the latter the only aspect of arbitrariness consists in having to define lexically identified sets in both “dimensions” of a paradigmatic pattern (paradigm classes and morphosyntactic values).<sup>24</sup> In this way we can avoid further complex hypotheses about the interaction of features of different strength and direction of influence and there is no need to stipulate a default value that surfaces when no feature is specified as a result of the interaction.

The simplicity of a paradigm-based account is clearly an advantage, which is even more conspicuous when we examine stems that are followed by unrounded front linking vowels in all the relevant cases. In the reference accent there is no surface-detectable lowering with these stems since a nonhigh linking vowel will always be low *e* ([ɛ]), as can be seen in the behaviour of the adjective *gyér* ‘scant’ in (8) below. Note that it makes no difference for a verbal stem if it is a root, e.g. *kezd*, or a “lowering stem” with a lowering suffix, e.g. *kezd-j* (recall that verbal back harmonic roots never show lowering with conformist suffixes).

---

<sup>22</sup> We are using the term morphosyntactic value in a sense that also includes the meaning of derivational affixes. This conforms to our interpretation of the paradigm as an extended one (cf. footnote 4). This interpretation is also supported by the graduality of the inflection–derivation scale in general and in Hungarian in particular (cf. Kiefer 1999, Mártonfi 2006).

<sup>23</sup> There are other morphosyntactic values resistant to this paradigm class distinction, recall the exponents which always have a high linking vowel discussed above (e.g., verbal and nominal *-nk*) and others that never have a linking vowel. We disregard these for the sake of simplicity in the present discussion.

<sup>24</sup> This arbitrariness is not complete but it is not negligible either. Harmonic classes that define paradigm classes unambiguously and lowering cannot *in general* be predicted from the phonological, morphophonological, categorial, or morphological properties of a form but *rarely* they can and in these cases these properties often co-determine the selection of a particular paradigm class or classes. Similarly, there is a correlation between morphosyntactic values whose exponents fit into the same linking vowel pattern and the morphological, morphophonological or phonological properties of the exponents or morphosyntactic values, although this correlation is not definitive and is typically rather weak. The rare cases when the correlation is strong can be accounted for if we recognise that in an analogical model paradigms are not assumed to be homogeneous but may contain emergent subparadigms whose members are more closely connected (e.g., the possessive subparadigm, see Rebrus et al. 2017).

(8) The uniformity of linking vowels (front unrounded harmonic class)

	<b>low</b> (nonconformist suffix)	<b>low</b> (conformist suffix)	<b>low</b> (nonconformist suffix)
<b>front unrounded harmonic class</b>	gyér- <b>e</b> -bb kezd(-j)- <b>e</b> -nek	gyér- <b>e</b> -k kezd(-j)- <b>e</b> -k	gyér- <b>e</b> -n kezd-j- <b>e</b> -n

The table in (8) shows that linking vowels in the forms associated with the relevant morphosyntactic values are not different: there is only one paradigm class here to which they all belong.<sup>25</sup> This results from rounding harmony and the asymmetrical inventory of vowels (Rebrus et al. 2023a): the word forms of this class cannot have a mid linking vowel because rounding harmony requires an unrounded front vowel in the exponents of these morphosyntactic values but the reference accent has no short unrounded *mid* front vowel — so low *e* [ɛ] shows up instead (the only nonhigh unrounded short front vowel in the system). As we have pointed out in §3, in a morpheme-based account this is due to the operation of a harmony “subsystem” (of rules/constraints and representations) which is independent of linking vowels and whose interaction with the subsystem responsible for the quality of linking vowels has to be accomplished by further stipulations. It has to be ensured that rounding harmony overrides the mechanism of lowering and nonlowering in this case. However, rounding harmony is asymmetrical and this results in further complications. On the one hand, it permits the occurrence of a low unrounded *e* linking vowel to follow a front rounded *ü, ű, ö, ő* in the preceding syllable irrespective of whether this linking vowel is due to the “low” specification of the root or the suffix: *zöld-e/\*ö-k, zöld-e/\*ö-bb*, (cf. *zöld-ö-n*), *küld-j-e/\*ö-k, küld-j-e/\*ö-nek* (cf. *küld-j-ön*); cf. (7). On the other hand, rounding harmony bans a mid (rounded) linking vowel following an unrounded front vowel both as default value and also when it is required by the specification of the suffix (cf. the examples in (8): *kezd-e/\*ö-k, kezd-j-e/\*ö-n, bír-e/\*ö-k bír-e/\*ö-n*).<sup>26</sup> Thus, the complex process of lowering is manifest in the first case but it is fully covert (opaque) in the second: it is undetectable at the surface.<sup>27</sup>

## 4.2 Identical substrings

Let us examine the analogical relations between cells that a paradigm-based analysis employs. The similarity of suffixes resistant to paradigm class (the nonconformist suffixes) is

<sup>25</sup> If we consider not only the quality, but also the conditions on the appearance of the linking vowel in the accusative or the occurrence of a yod in possessive forms of nouns, then the -*e*- class must be split into further subclasses. To some extent this holds true of the other paradigm classes as well (cf. Rebrus 2019, Rebrus et al. 2017).

<sup>26</sup> In some south-eastern varieties (e.g., Szeged) these forms contain *ö*. The working of labial harmony is fundamentally different in these varieties (cf. Polgárdi & Rebrus 1998, Rebrus & Törkenczy 2019).

<sup>27</sup> A possible solution is to introduce mid /*e*/ into the system, a sound which is completely missing from the dialect examined (cf. Vago 1980, Siptár & Törkenczy 2000). In this analysis this abstract unrounded mid front vowel is active in the phonology but at some point undergoes complete neutralisation and surfaces as low [ɛ] (indistinguishable from the vowel that is underlyingly low /*ɛ*/). Absolute neutralization represents an even higher degree of opacity and introduces a further element of high operational complexity into the system. Note that it makes no difference how the two unrounded vowels are distinguished representationally (e.g., both fully specified or one of them underspecified) since in any case a distinction is made that never occurs at the surface, which further complicates an already complicated process.



manifested in the height of the linking vowel, that is, suffixes requiring a low linking vowel to precede them will always have the same vowel within a harmonic class (provided that this vowel appears in the first place). This is *a* in the back class, *e* in the front class: *vak-a-bb*, *tág-a-bb*; *old(-j)-a-nak*, and *hős-e-bb*, *zöld-e-bb*, *gyér-e-bb*; *küld(-j)-e-nek*, *kezd(-j)-e-nek* (as in the left columns of (7) and (8)). Suffixes with a linking vowel prescribed to be mid will also have the same vowel: *o* in the back class (*vak-o-n*, *tág-o-n*; *old-j-o-n*), *ö* in the front rounded class (*hős-ö-n*, *tág-ö-n*; *old-j-ö-n*), and *e* in the front unrounded class (*gyér-e-n*; *kezd-j-e-n*). Thus, in a given morphosyntactic value and harmonic class, the linking vowels are identical. In a whole-word model this analogical relationship can be captured in the identity of the final substrings of words belonging to a given morphosyntactic value across paradigm classes, since the nonalternating part of the exponent is fixed anyway: *-bb* in the comparative, *-n* in the superessive; *-sz* in 2sg nondefinite. The following analogical relations hold for back harmonic words containing a linking vowel<sup>28</sup> (we have underlined the common final substrings within the (sub)paradigms, “≈” means analogical similarity manifested in the identical linking vowel, the arrow indicates the substring emerging from the given pattern).<sup>29</sup>

(9) Common final substrings of exponents with a resistant linking vowel (back harmonic)

CMPR :	<i>vak<u>abb</u> ≈ gazdag<u>abb</u> ≈ magas<u>abb</u> ≈ tág<u>abb</u> ...</i>	→ <i>_abb</i>
NDF.3PL :	<i>old<u>anak</u> ≈ ring<u>anak</u> ≈ oldj<u>anak</u> ≈ ringj<u>anak</u> ...</i>	→ <i>_anak</i>
SUE :	<i>nap<u>on</u> ≈ tag<u>on</u> ≈ láb<u>on</u> ≈ ág<u>on</u> ...</i>	→ <i>_on</i>
SBJV.NDF.3PL :	<i>oldj<u>on</u> ≈ ringj<u>on</u> ≈ fusson ≈ tar[ʃ]<u>on</u> ...</i>	→ <i>_on</i>

The similarities in (9) can be referred to as **morphosyntactically determined** since if the linking vowel appears the uniformity of **final substrings** follows irrespective of paradigm classes (i.e., of lowering) modulo vowel harmony.

The identity of final strings also hold with suffixes conformist with respect to lowering, but it is constrained by the paradigm class: it is only identical within a paradigm class, not across them, cf. (10).

(10) Final substrings of exponents conforming to paradigm classes (back harmonic)

PL :	<i>vak<u>ok</u> ≈ nap<u>ok</u> ≈ tag<u>ok</u> ... ≠ tág<u>ak</u> ≈ láb<u>ak</u> ≈ ág<u>ak</u> ...</i>
------	---

Note, however, that as opposed to suffixes resistant to paradigm class, which are often unique, there is an abundant number of conformist suffixes. Words belonging to the same paradigm class (and harmonic class) contain the same linking vowel with a wide range of

<sup>28</sup> In (9) and displays that follow, the morphosyntactic value is followed only by the forms containing the given morphological specification and relevant for that pattern, not *all* of them. Thus CMPR in (9) means all the *back-vowelled* comparative forms. The same holds later for the *names* of paradigms: they contain only the relevant forms associated with certain (but not all) morphosyntactic values.

<sup>29</sup> This holds for high linking vowels, too, e.g., POSS.1PL: *napunk ≈ tagunk ≈ lábunk ≈ águnk*, etc. and NDF.1PL: *oldunk ≈ ringunk ≈ oldjunk ≈ ringjunk*, etc.

For the sake of simplicity, we here disregard cases of multiple agglutinative suffixation. In multiple suffixation the exponent we focus on may be followed by another exponent and as a result the relevant uniform substring is not word final, but word medial. This also can occur with some of the exponents mentioned above, e.g., *tág-a-bb-nak* ‘-CMPR-DAT’, *nap-u-nk-hoz* ‘-POSS.1PL-ALL’, etc.

morphosyntactic values. This similarity can be referred to as **paradigmatically determined**, since these words show the same behaviour with many suffixes within a paradigm. In this case the similarity is manifested in the uniformity of the **initial substrings**. We label the paradigm of specific lexemes with the English gloss in italic small capitals and underline the uniform word-initial strings in (11).

(11) Initial substrings within a paradigm in case of conformist exponents (back harmonic)

DAY :	<u>napok</u> ≈ <u>napom</u> ≈ <u>napod</u> ≈ <u>napotok</u> ≈ <u>napot</u> ≈ <u>napos</u> ...	→ napo_
FOOT :	<u>lábak</u> ≈ <u>lábam</u> ≈ <u>lábád</u> ≈ <u>lábatok</u> ≈ <u>lábat</u> ≈ <u>lábás</u> ...	→ lába_
SOLVE INDV:	<u>oldok</u> ≈ <u>oldom</u> ≈ <u>oldod</u> ≈ <u>oldotok</u> ...	→ oldo_
SOLVE SBJV :	<u>oldjak</u> ≈ <u>oldjam</u> ≈ <u>oldjad</u> ≈ <u>oldjatok</u> ...	→ oldja_

“Initial substring of a (sub)paradigm” as a description is a surface-true generalization. It contains all the information about the phonological properties of what is traditionally called the stem,<sup>30</sup> including whether it is lowering or nonlowering, in a unified manner.

Uniformity which is more restricted than that in (9) typically has full scope: for example, all nonpossessive nominative plural forms end in *-k*, independently even of vowel harmony and/or paradigm class. This indicates that the alternation between the exponents is morphophonological and not lexical (the latter occurs in suppletive allomorphy where no phonological generalization can be made about the alternants). In the case of forms with a linking vowel this is disadvantageous since it only affords a partial generalization (which does not include the linking vowel), but it has the advantage that it holds true of all forms irrespective of which paradigm class they belong to. No broader generalization than this can be made about forms where the linking vowel is missing — vowel-final stems typically behave in this way (e.g., *fű-k*, *kapu-k*, *hajó-k*, *fű-k*).<sup>31</sup>

(12) Minimally uniform final substring for all forms

PL :	<u>vakok</u> ≈ <u>napok</u> ≈ <u>tágak</u> ≈ <u>ágak</u> ≈ <u>rögök</u> ≈ <u>szögek</u> ...	→ _k
PL :	<u>fűk</u> ≈ <u>kapuk</u> ≈ <u>kocsik</u> ≈ <u>hajók</u> ≈ <u>fűk</u> ≈ <u>erdők</u> ...	→ _k

<sup>30</sup>A reference to initial substrings successfully captures the correlation of stem alternations with the presence or quality of linking vowels. All this information is contained within the initial substrings of the relevant forms. This holds for phonologically unpredictable stem alternations, for instance, vowel shortening (*nyár* ‘summer’, *nyár-o-n* ‘-SUE’ vs. *nyar-a-k* ‘-PL’, *nyar-a-t* ‘-ACC’, etc.), vowel–zero alternation (*lator* ‘rogue’ vs. *latr-o-k* ‘-PL’, *latr-o-t* ‘-ACC’, *latr-o-n* ‘-SUE’, etc.) or palatalization in the subjunctive (*oszt* ‘distribute’, *oszt-o-k* ‘-NDF.1SG’, *oszt-o-m* ‘-DEF.1SG’, etc. vs. *osz-sz-a-k* ‘-SBJV-NDF.1SG’, *osz-sz-a-m* ‘-SBJV-DEF.1SG’, etc.):

SUMMER :	<u>nyarak</u> ≈ <u>nyarat</u> ≈ <u>nyaram</u> ≈ <u>nyaras</u> ≈ <u>nyaranta</u> ...	→ nyara_
BUSH :	<u>bokrok</u> ≈ <u>bokrot</u> ≈ <u>bokrom</u> ≈ <u>bokros</u> ≈ <u>bokronként</u> ...	→ bokro_

DISTRIBUTE INDV :	<u>o[st]ok</u> ≈ <u>o[st]om</u> ≈ <u>o[st]od</u> ≈ <u>o[st]otok</u> ...	→ o[st]o_
DISTRIBUTE SBJV :	<u>o[ss]ak</u> ≈ <u>o[ss]am</u> ≈ <u>o[ss]ad</u> ≈ <u>o[ss]atok</u> ...	→ o[ss]a_

Thus initial substrings simultaneously contain all the information traditionally distributed over the lexicon (the underlying form of the stem) and the grammar (the morphophonological rules).

<sup>31</sup> It is interesting to note that the segment before the final substring *\_k* is always a vowel but the stem-final vowels (i.e., those that occur at the end of the unsuffixed form) are in complementary distribution with the non-high linking vowels (*a*, *e*, *o*, *ö*) here, cf. Novák 1999. Therefore this is a valid generalization about the final substring of plural forms: *\_Vk*, where *V* is any vowel.

Segmentation into morphemes is not important in an analysis based on similarity: the boundaries of morphs do not have to be circumscribed (an unavoidable decision that must be made in an atomistic theory). Thus, it is possible for a form to be similar to some forms in its initial substring and to other forms in its final substring at the same time so that the substrings overlap (in our case the relevant overlap is the linking vowel). A linking vowel in this case is neither a final nor an initial substring but can be interpreted as the substring they share. We illustrate this with the linking vowel in the form *nap-o-k* ‘day-PL’. The recurrent constant final substring associated with the morphosyntactic value PL in the relevant paradigm class is *\_ok* (emboldened in (13a)). The initial substring of forms associated with several morphosyntactic values in the lexical paradigm of *DAY* is *napo\_* (underlined in (13b)). The paradigm cell for the morphosyntactic value PL in this lexical paradigm  $\langle DAY, PL \rangle$  contains the form *napok*, which realizes the initial substring of the lexical paradigm and the final substring associated with the morphosyntactic value PL in a single form. The generalization that holds true of all the forms in the relevant part of the lexical paradigm of *DAY* and/or the relevant forms associated with the morphosyntactic value PL is that they all contain the overlapping part of the substrings in (13a) and (13b), i.e. the segment *o*, cf. (13c).

(13) The overlap of initial and final substrings

- a. PL :            *napok*  $\approx$  *tagok*  $\approx$  *vakok* ...         $\rightarrow$  *\_ok*
- b. *DAY* :        *napok*  $\approx$  *napod*  $\approx$  *napos* ...         $\rightarrow$  *napo\_*
- c. all forms associated with PL and/or *DAY*  $\rightarrow$  *\_o\_*

The linking vowel *o* above<sup>32</sup> occurs in a specific form (*napok*) such that it is contained in the initial and final substrings of several other forms: the form *napok* is thus in the “intersection” of the relevant lexical paradigm and morphosyntactic value. If we consider more paradigms and more morphosyntactic values, then we can obtain the linking vowel(s) (if they do indeed occur) and observe their quality in all the cells of the paradigms in the system by taking for each specific form the shared part of (i) the initial substring of forms in the lexical paradigm associated with the specific form and (ii) final substring associated with the morphosyntactic value of the specific form. Therefore, we can consider the linking vowels of various qualities (identical or different) as a pattern emergent from the paradigmatic system. In an atomistic approach, such patterns have a number of isolated sources: the underlying representations of morphemes and the various morphophonological processes applying to them. That is, the emergent patterns appear to result from a conspiracy of unrelated factors, so generalizations on surface patterns cannot be directly captured.

The linking vowel arrived at in the way described above is the overlapping part of the recurrent initial and final substrings, which implies ambiguity in parsing (cf. the examples in (1b)). There is no need to distinguish a “stem morph” and an “affix morph” to define what a

---

<sup>32</sup> In a more detailed analysis we can also make a generalization about the environment of the linking vowel *o* must be preceded and followed by a consonant, so the shared part is *-CoC-*. This is a sign of phonological motivation: a consonant cluster would occur in the absence of a linking vowel, which is a phonotactically marked state of affairs (especially word-finally).

linking vowel is since it is a derived element in a paradigm-based analysis.<sup>33</sup> The question underlying the fuzziness of this interpretation of linking vowels is this: what counts as “recurrent”? Can a substring be considered a linking vowel if it only occurs in a handful of forms? Or only one? This fuzziness, however, is an advantage: here the boundaries of the linking vowel category are not expected to be clear-cut but are a function of the generality of the relevant paradigmatic pattern. If a suffix that is the sole representative of its class occurs in a paradigm, the relevant initial substring will only occur in a single form, e.g., there will be no other forms with an initial substring whose final segment is *a*, *o*, or *u* in the lexical paradigms that include *lát-a-tlan* ‘see-PRIV’, *tág-o-n* ‘broad-SUE’, or *nap-u-nk* ‘day-POSS.1PL’, respectively. There are very few other forms with the initial substring *láta\_* (because the verb *lát* is not a lowering stem), *tágo\_* (because *tág* is a lowering stem), or *napu\_* (because *-nk* is unique in taking a high vowel on its left). However, there will be many forms with the final substrings *\_atlan*, *\_on*, or *\_unk* that have the appropriate morphosyntactic value, practically as many as appropriate stems exist — in this case all the consonant-final back-harmonic stems. Thus, the status of these vowels as linking vowels receives less support from the lexical paradigms (as the type frequency of the initial substrings is low) but strong support from the final substrings associated with the relevant morphosyntactic values (since the type frequency of the final substrings is high).

### 4.3 Paradigmatic patterns of the linking vowel

The emergence of linking vowels from the system of paradigms allows us to establish paradigmatic patterns based on the quality of these vowels. Paradigm classes can be set up based on the linking vowel occurring in word forms. According to harmonic classes and “lowering”, there are five such classes, which we arbitrarily label with their *most common* linking vowel(s): (i) back nonlowering -o-, (ii) back lowering -a-, (iii) front rounded nonlowering -ö-, (iv) front rounded lowering -e|ö-, and (v) front unrounded -e|e- class, which is unspecified for lowering. The label -other- indicates paradigms in which there is some other segment (vowel or consonant) preceding the final substring, traditionally these are analysed as not having a linking vowel. The behaviour of the -e|ö- and -e|e- classes differ only in those cases where the linking vowel is constantly mid: *zöldön* (vs. *gyéren*) and *küldjön* (vs. *kezdjen*), both have the linking vowel *e* or *ü* in all other cases. The linking vowels are different in case of conformist suffixes in all the other classes. This is shown in (14), where the rows represent the six paradigm classes (five with, one without a linking vowel). The columns represent four groups of morphosyntactic features, the linking vowels of which are different in one or several of the other groups in some paradigm class. Areas of the paradigm system where the linking vowels are different are separated by a line (the low

<sup>33</sup> Furthermore, other elements that are similar to linking vowels but have a limited distribution can be derived in the same way, e.g. elements that rarely or never alternate with zero, but have the same quality as linking vowels, like the fractional (*hat-o-d*, *nyolc-a-d*, *öt-ö-d*), the 3rd person possessive and person/number endings *-a~e*, *-uk~ük~k* (*láb-a*, *láb-uk*, *für-t-a*, *für-t-uk*), the NDF.3SG *-n* alternant (*für-j-o-n*), the NDF.3PL *-k* alternant (*für-t-a-k*) and the NDF.2SG *-l* alternant (*mos-o-l*). These data show that the linking vowel as a category is fuzzy too. In a paradigm-based analysis, nothing hinges on whether a given element is a linking vowel or not — a clear example is the analogical connection between the linking vowels of the possessive paradigm and the vowel of the yodless third person alternants (*-a~e*, *-uk~ük*), see Rebrus et al. 2017.

linking vowels *a* and *e* are emboldened). The areas thus delineated define the portions of the paradigm space in which strong analogical similarities operate horizontally (identical initial substrings) and vertically (identical final substrings, both containing the linking vowel). The varying width of columns is related to the numerosity of the given suffix class: conformist suffixes are by far the most numerous type, followed by the low-before group (containing a couple of nominal and several verbal suffixes), the least numerous “mid” and high groups contain one or two denominal or deverbal suffixes and nominal and verbal inflections. We provide sample suffixes in the last row, denominal in the first, deverbal in the second line. Suffixes whose distribution is limited morphotactically or otherwise are parenthesized.

(14) Paradigmatic patterns by the quality of linking vowels

Paradigm classes	Morphosyntactic values by the linking vowel				Examples
	low	variable	“mid”	high	
-o-	<b>a</b>	o	o	u	<i>tag vak sír (fúr-n-)</i> <i>old (fúr, só-z)</i>
-a-	<b>a</b>	<b>a</b>	o	u	<i>ág tág díj; nál- után</i> <i>(fúr-j fúr-t)</i>
-e e-	<b>e</b>	<b>e</b>	<b>e</b>	ü	<i>heg gyér hír (ér-n-); ért- kezd (ér-j ér-t sí-z-)</i>
-e ö-	<b>e</b>	<b>e</b>	ö	ü	<i>szög zöld; töl- előtt</i> <i>(tör-j tör-t)</i>
-ö-	<b>e</b>	ö	ö	ü	<i>rög hős (tör-n-)</i> <i>küld (tör nő-z-)</i>
-other-	—	—	—	—	<i>sí kapu apró; rá- alá</i> <i>(fúr-ná- fúr-já- ér-i)</i>
Examples	<i>-bb -tlan</i> <i>-lak -nak (-n- ...)</i>	<i>-k -m -d -tOk (-t -z -l -n -s ...)</i> <i>-k -m -d -tOk (-k -tt -z -l -gAt ...)</i>	<i>-n</i> <i>(-n)</i>	<i>-nk (-k)</i> <i>-nk (-k)</i>	

The definition of paradigm classes is based on the quality of linking vowels. None of these paradigm classes can be defined exclusively by phonological properties.<sup>34</sup> The difference between paradigm classes may be (i) completely lexical, (ii) partly, or (iii) completely phonologically predictable.

(i) The difference between lowering and nonlowering classes (-o- vs. -a- and -ö- vs. -e|ö-) is lexical, it correlates with phonological properties only very weakly (see Trón & Rebrus 2001). Functional/semantic correlations are stronger, but not definitive (for example, all verbal roots are nonlowering, but suffixes can change this, e.g., *fúr-j-a-m* ‘drill-SBJV-DEF.1SG’, *fúr-n-o-m* ‘-INF-1SG’; see lowering and nonlowering suffixes in §2). This is why these paradigm classes are not separated by a line in the leftmost column of (14).

<sup>34</sup>The phonological properties meant here are the phonologically “visible” (surface true) properties of the forms in the paradigm, not the abstract (invisible) features of some theory of phonology. The exponentless elements (free “stems”) in the -other- paradigm class, which lacks a linking vowel, generally end in a vowel, but even this is not exceptionless: many adjectives and some nouns contain a low linking vowel in this context (e.g., *férfi-a-s* ‘man-ADJZ’, *fi-a-m* ‘son-POSS.1SG’, *hosszú-a-k* ‘long-PL’, *méltó-a-n* ‘worthy-ADVZ’, *pest-i-e-k* ‘Pest-ADJZ-PL’, etc.), and with some pronominal forms other vowel qualities also occur (e.g., *mind-annyi-o-tok* ‘all of youse, lit. all-so\_much-2PL’, *vala-mennyi-ü-nk* ‘all of us, lit. some-how\_much-1PL’). And, vice versa, a linking vowel may be missing between a consonant-final stem and some suffixes, see data in (1) and (2); this is not shown in (14).

(ii) The front and back harmonic classes, which distinguish the paradigm classes -e- vs. -o- and -a-, respectively, can only partly be identified by the phonological properties of the free stem. The reason is the strongly morphological nature of front/back harmony (cf. Rebrus et al. 2023ab). Antiharmonic roots (classes -a- or -o-) contain a vowel that normally occurs in front harmonic roots (class -e|e-): e.g., *díj-a-k* ‘prize-PL’, *sír-o-k* ‘grave-PL’, *hír-e-k* ‘news-PL’. Moreover, in a large set of words there is a systematic vacillation in the harmony of the linking vowel (e.g., *fotel-o-k/fotel-e-k* ‘armchair-PL’), but for some roots this vacillation is found in the suffix vowel, but not in the linking vowel (e.g., *haver-nak/nek* ‘pal-DAT’ vs. *haver-o/\*e-k* ‘-PL’; *partner-nak/nek* ‘partner-DAT’ vs. *partner-e/\*o-k* ‘-PL’; cf. Forró 2013, Kálmán & Forró 2014, Rebrus et al. 2022). Consequently, the distinction between the three classes is only partly phonological (represented by the dashed line between them in (14)).

(iii) The only harmonic behaviour that can be predicted for the phonological form is labial harmony, which is limited to front vowels. Stems whose last vowel is front rounded can only select the front rounded variant of suffixes exhibiting a three way (*o~e~ö*) alternation (e.g., *-hoz~hez~höz* ‘ALL’, *-tok~tek~tök* ‘2PL’). This constraint only limits the quality of the linking vowel before nonconformist mid suffixes (e.g., *szög-ö-n* ‘nail-SUE’, *nő-k-ö-n* ‘woman-PL-SUE’, *rög-ö-n* ‘clod-SUE’, *tör-j-ö-n* ‘break-SBJV-NDF.3SG’), but not with conformist suffixes, since such stems may be lowering -e|ö- class and nonlowering -ö- class (e.g., *szög-e-t*, *nő-k-e-t*, *rög-ö-t* ‘-ACC’, *nő-z-ö-tt* ‘woman-VRBZ-PST’, *tör-j-e-k* ‘break-SBJV-NDF.1SG’ vs. *tör-n-ö-m* ‘-INF-1SG’, *tör-ö-m* ‘-DEF.1SG’). The two classes are separated by a solid line in the first column of (14).

The last column of (14) contains some stems of the relevant paradigms, both suffixed and unsuffixed. We first give items taking nominal suffixes (monomorphemic noun and adjective roots, e.g., *tag* ‘member’, *vak* ‘blind’, and infinitive stems, e.g., *fúr-n-* ‘drill-INF’), then items taking verbal suffixes (morphologically complex, e.g., *só-z* ‘salt-VRBZ’, *fúr-j* ‘drill-SBJV’, and unsuffixed roots where available, e.g., *áld* ‘bless’, *fúr* ‘drill’); lastly we give adpositional stems: suffixable inflections (e.g., *nál-* ‘ADE’) and postpositions (e.g., *után* ‘after’). In many examples there is a morphotactic or morphophonological limitation on the suffixes that can follow or that have a linking vowel before them; these items are parenthesized.

Table (14) conspicuously shows that analogical connections hold across paradigms and types of morphosyntactic values. On the one hand, the same linking vowel can occur associated with a specific morphosyntactic value (type) in more than one paradigm class (the vertical dimension in the table). Nonconformist suffixes behave in this way — they are only sensitive to harmonic classes but not to paradigm classes. Accordingly, the linking vowels *a*, *o*, *u* each occur in two paradigm classes (the -a- class and the -o- class). Similarly, the linking *ö* also occurs in two paradigm classes (the -ö- class and the -e|ö- class) with the suffixes labelled as “mid”. The linking vowels *e* and *ü* each occur in three paradigm classes (classes -e|e-, -e|ö- and -ö-) with high and low nonconformist suffixes. The linking vowel *e* is special because it occurs in two paradigm classes (classes -e|e and -e|ö-) even with nonconformist suffixes; cf. (15b) below.

On the other hand, in a specific paradigm (class), the same linking vowel (except high *u/ü*) can occur with more than one different type of morphosyntactic values (the horizontal dimension in the table). The linking vowels *o* and *ö* occur with conformist and “mid” nonconformist suffixes (in the -*o*- and -*ö*- paradigm classes — the traditional nonlowering stems). The linking vowels *a* and *e* occur with conformist and low nonconformist suffixes (in the -*a*- and -*e|ö*- paradigm classes — the lowering stems). Finally, the linking vowel *e* is shared by three different types of morphosyntactic values in the -*e*- paradigm class *e* occurs in low, the “mid” and the variable type too, cf. (15a) below.

(15) Expanse of the linking vowel patterns

a. paradigmatic patterns (same initial substrings)					b. morphosyntactic patterns (same final substrings)				
		low	variable	m.   h.		↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓	↓	↓	
-o-	⇒	a	o	u	a	o	o	u	
-a-	⇒	a		o u	a	a			
-e e-	⇒	e		ü	e	e	e		
-e ö-	⇒	e		ö ü	e	e	ö	ü	
-ö-	⇒	e	ö	ü	e	ö			

The linking vowel patterns observed in the two dimensions discussed in the previous paragraphs are depicted in (15), which shows the patterns in (14) separately for paradigm classes (15a), indicated by horizontal arrows, and for morphosyntactic types (15b), indicated by vertical arrows. Cells that are shaded grey stretch over more than one type “horizontally” or “vertically”. The complex pattern in (14) can be considered as the result of the patterns in (15), showing which cells can participate in analogical inference in the given dimension.

#### 4.4 The paradigm cell filling problem

The patterns of linking vowels in the paradigmatic system described above has important repercussions on our view of the creative use of language (the so-called linguistic productivity). In a holistic model, it is a viable assumption that the some word forms are memorized. This cannot hold for all word forms though, since there exist open word classes (e.g., nominal roots, as well as nominal and verbal stems created by productive derivational affixes) and closed classes also contain a vast array of elements (e.g., there are many verbal roots, each with hundreds of potential forms in their paradigm, some with practically zero frequency). It is very likely that less frequent word forms, including loans at the early stages of entering the language, are not stored.<sup>35</sup> What is and what is not stored may vary by individual, and it is theoretically impossible to draw the line between the two sets. So how can the speaker determine which pattern an unstored item will fit into? This is the task known as the *paradigm cell filling problem* (PCFP): in a given paradigmatic system with a set of

<sup>35</sup> A similar view of storage (listedness vs. unlistedness) does not only characterize a holistic models, but is accepted by some morpheme-based ones, too, cf. Steriade 1999, Breiss 2022.

analogical similarities, how and how well can it be predicted which similarities will hold for a specific paradigm cell (cf. Ackerman et al. 2009, Blevins 2016, Blevins et al. 2017).

We have seen in §4.2 that the analogical identities of linking vowels are of two basic types: morphosyntactically and/or paradigmatically determined. The former are linked by the identity of the final, the latter by the identity of the initial substrings of word forms. These similarities are the clues to PCFP. When predicting the quality of the linking vowel in a form in a cell of the paradigmatic system, there are two main possibilities: (i) the given cell is associated with a morphosyntactic value that is resistant to paradigm class distinctions or (ii) with one that is sensitive to them, i.e., the given suffix is conformist.

(i) In the first case the amount of information needed to predict the linking vowel (and thus the whole word form) is minimal: the harmonic class of the form. This can be extracted from any cell of the word's paradigm whose word form contains a harmonic exponent (the harmonizing vowel need not even be a linking vowel). There is no more we need to know about the paradigm class of the word. For instance, the comparative form of an infrequent adjective that contains a neutral vowel (e.g., *mísz* 'indisposed') can be successfully inferred after deciding whether it belongs to the front ("harmonic") or back ("antiharmonic") class (the former in our case, as revealed by any of the forms *mísz-e-k*, *mísz-nek*, *mísz-től*, etc.). Since in the case of front harmonicity the final sequence belonging to the value CMPR is *\_ebb* (see (14)), the relevant cell will contain *mísz-e-bb* (and not *\*mísz-a-bb*).<sup>36</sup> Likewise, the superessive form can be predicted from any other form in the paradigm that contains a harmonically variable exponent. The form we expect in the SUE cell of the paradigm of the rare, harmonically mixed word *bóher* 'bohur' will depend on other forms (in this case typically containing a front exponent: *bóher-e-k*, *bóher-nek*, *bóher-ről*, *bóher-ü-nk*, with some vacillation). Since the final sequence is *\_en* here (see (14)), the cell will contain the form *bóher-e-n* (?%*bóher-o-n* is less likely, though it may appear for speakers who also use other back harmonic forms, e.g., ?%*bóher-o-k*, ?%*bóher-nak*, ?%*bóher-ról*, ?%*bóher-u-nk*). In a similar vein, to predict the subjunctive third person singular nondefinite form of the rare verb *murdel* 'die', we have to check the harmony of known forms. If we are aware of any of the forms containing a back-vowelled exponent, *murdel-o-k*, *murdel-nak*, *murdel-hat*, *murdel-j-a-k*, etc., then the form *murdel-j-o-n* is predictable with full certainty. If, however, some of the forms %*murdel-e-k*, %*murdel-nek*, %*murdel-het*, %*murdel-j-e-k*, etc. are stored, then the relevant form will be *murdel-j-e-n*.<sup>37</sup> The same considerations apply to high linking vowels, e.g., *bóher-u-nk*, ?%*bóher-ü-nk*, *murdel-u-nk*, %*murdel-ü-nk*.<sup>38</sup>

<sup>36</sup> If the last vowel in the root is not neutral, even its harmony is unambiguous, so there is no need to check other paradigm cells, e.g., the CMPR of *kül* can only be *kül-a-bb*, and that of *prüd* can only be *prüd-e-bb*. But in these cases, speakers must have access to information about the form of the exponent (the final substring(s)) associated to the specific morphosyntactic value, which knowledge presupposes that they have memorized a certain number of forms with this morphosyntactic value.

<sup>37</sup> Labial harmony does not have to be considered in such words because it is automatically triggered by the vowel immediately preceding the linking vowel: if that vowel is front rounded, the linking vowel will be so too in all cases, i.e., *-ön*, if it is unrounded, it will be *-en*.

<sup>38</sup> The description of the mechanism is somewhat simplified. Front/back harmony is sensitive to multiple phonological, morphological, and semantic analogies (cf. Rebrus et al. 2023b), and it is often subject to vacillation and/or hesitation even in the same speaker. Hesitation occurs in two situations: either contradictory sources (with both front and back harmonic exponents) memorized in the paradigm, or no forms are memorized



(ii) For the prediction of the linking vowel before conformist suffixes, information about front/back harmony is insufficient, since lowering or its absence is independent of harmony. In this case we have to know the paradigm class the given word belongs to, which is revealed only by another form in a cell of the same paradigm containing a conformist exponent. So in order to get the adverbial form of the adjective *konkrét* ‘factual’, we have to find a form that contains the linking vowel before a conformist suffix (e.g., *konkrétak* ‘-PL’ or *konkrétat* ‘-ACC’, not *\*konkrétok/t* or *?%konkréték/t*), which discloses that the initial substring is *konkréta\_*, that is, the word belongs to the -a- paradigm class. Accordingly, the adverbial form will be *konkrétan*, not *\*konkrétón* or *?%konkrétén*.<sup>39</sup> The condition for finding such a form with great probability is that there should be a large enough set of morphosyntactic values that are sensitive to the paradigm class and that there should be a cell with a form in it whose token frequency is large enough to guarantee its previous memorization.

In some cases the analogical deduction is possible in both directions. Many verbal suffixes, for example, are preceded by a low linking vowel, thus the linking-vowelled 2<1SG form of a verb is always *\_alak* or *\_elek*. The initial substrings of forms with a low linking vowel will also be identical. In case of the recently coined verb (*le*)*csippant* ‘cause to beep’, this substring is *csippanta\_* (cf. the other forms *csippant-a-nak* ‘-NDF.3PL’, *csippant-a-ni* ‘-INF’, *csippant-a-na* ‘-COND’, *csippant-a-sz* ‘-NDF.2PL’). The final substring is *\_alak* (due to the last, back vowel of the stem), accordingly the cell will contain the combination of these initial and final substrings, *csippant-a-lak*.

Sometimes, knowing the contents of a cell associated with some other type of morphosyntactic feature is sufficient for the inference of a form. Conformist exponents are generally good predictors about their resistant counterparts, since the harmonic class is directly indicated by the quality of the linking vowel, which is determined by the paradigm class. As we have shown, the harmonic class is sufficient to predict the vowel of nonconformist suffixes. There exists yet another case: if a resistant linking vowel is specified as “mid”, the linking vowel of conformist suffixes will always be *e*, e.g., if the exponent of the superessive is *-en* in a word, that of the plural will be *-ek*.

The analogical relations in the two dimensions (horizontal and vertical) are represented in (16). A verbal or nominal paradigm may contain multiple forms with the same type of morphosyntactic value; this is shown in (16a). In this scenario any of the forms infers any other, since their linking vowel (if present) will be identical. This logical relation is indicated by “ $\Leftrightarrow$ ”. This kind of inference is very important for conformist exponents, which have no other information for identifying the linking vowel associated with their cell. But it may have a role in other cases where multiple types of exponents are available, for instance, in the case of the nonconformist verbal suffixes mentioned above, the linking vowel before which is either uniformly low or uniformly high.

---

that could provide a clue about harmony. In the latter case the speaker has to search for further analogies, which if inconclusive leads to hesitation.

<sup>39</sup> The form *†konkrét-e-n*, which formerly prevailed, is now obsolete.

The relevance of paradigmatically determined analogical relations (initial substrings) is highlighted by the patterns of high nonconformist suffixes. The two exponents in the nominal possessive subparadigm are the 1PL *-nk* and the *-k* allomorph of the 3PL suffix (the others being *-juk~jök*), the vowel before which is always high (unless it is part of the stem). Accordingly, the alternants under consideration are *-unk~ünk* and *-uk~ük*, respectively. Rebrus & al. (2017) demonstrate that the morphological condition for the appearance of the yodless alternants of third person possessive exponents is analogical support within the possessive subparadigm, that is, the occurrence of the same linking vowel in a possessive form of some other person.<sup>40</sup> The analogical constraint results in extensive and multiple asymmetries. Two yodful 3sg possessive forms are available for the harmonically vacillating root *fotel* ‘armchair’ (*fotel-je/fotel-ja*), but the back yodless 3sg possessive form is not (*\*fotel-a*) because the linking vowel in its subparadigm is not identical: *o*, not *a* (as shown by the 1sg, 2sg, 2pl possessives *fotel-o-m*, *fotel-o-d*, *fotel-o-tok*, respectively). However, the front yodless 3sg form *fotel-e* is available since the front linking vowel is *e* elsewhere in its subparadigm (*fotel-e-m*, *fotel-e-d*, *fotel-e-tek*). The back yodless possessive form is available in 3pl though (*fotel-u-k*) since this is supported by the back 1pl form (*fotel-u-nk*): here the linking vowel is high irrespective of the bias of the stem. Thus in 3sg we have two yodful but only one yodless alternant: *fotel-je/fotel-ja/fotel-e*; while in 3pl all four alternants are available: *fotel-jök/fotel-juk/fotel-ük/fotel-uk* (needless to say, the availability of alternants is speaker-dependent, not all speakers have all of them).<sup>41</sup> The need for analogical support within the possessive subparadigm also indicates that it makes no difference whether the relevant vowel is a linking vowel or not. In the third person possessive exponents, *-je~ja~e~a* and *-jök~juk~ük~uk*, the vowel is not a linking vowel, since it is always present, yet it is the vowel of the yodless alternants that receives analogical support from the linking vowel of the forms of other persons simply by containing identical initial substrings (*fotele\_* and *fotelü\_/fotelu\_*).<sup>42</sup> A morpheme-based theory can handle this only in a very complicated manner, if at all, because linking vowels are generally not present underlyingly, but stable suffix vowels are, so the pattern can only be captured by the comparison of surface forms.

The other dimension of analogical relations are shown in (16b). There is usually no need for paradigm-internal analogies for nonconformist suffixes (except for the case of the possessive subparadigm discussed above), the linking vowel can be inferred from any harmonizing suffix. In this case the analogy is manifested in the identity of final subsequences within a harmonic class (indicated by “ $\Phi$ ” in the chart).

<sup>40</sup> This basic and productive pattern can be overridden by phonological factors or frequency, which we disregard here, cf. Rebrus 2013, Rácz and Rebrus 2012.

<sup>41</sup> In the paradigms of so-called cultural loans (cf. Forró 2013, Forró & Kálmán 2014, Rebrus & al. 2023ab), like *partner* ‘id.’ back linking vowels are inhibited, so *\*partner-u-nk* cannot lend support for *\*partner-u-k*, although both front and back harmonic suffix alternants exist where the vowel is not a linking vowel, even in 3pl possessives: *partner-jök/partner-juk*. Thus only the front yodless 3pl form *partner-ük* is available.

<sup>42</sup> Furthermore, in 3sg the initial substring coincides with the full form (*fotele*), like it regularly does in lowering stems, e.g., *ág-a* ‘branch-POSS.3SG’, beside *ág-a-m* ‘-POSS.1SG’ and in front unrounded stems, e.g., *vég-e* ‘end-POSS.3SG’, beside *vég-e-m* ‘-POSS.1SG’ (Kálmán p.c.).

(16) Analogical inferences between linking vowels

a. recurrent linking vowels: (same initial substrings of exponent(s) of the same type)					b. resistant linking vowels: (same final substrings of harmonic exponent(s))								
		low	variable	high									
-o-	→	a ⇔ a	o ⇔ o ⇔ o		↓ ↓ ↓ ↓	a ⇕ a	o	o ⇕ o	u ⇕ u	⇔B			
-a-	→		a ⇔ a ⇔ a									a	u
<hr/>													
-e e-	→	e ⇔ e	e ⇔ e ⇔ e		↓	e ⇕ e	e	e ⇕ ö	ü ⇕ ü	⇔F			
-e ö-	→												
-ö-	→		ö ⇔ ö ⇔ ö									ö	ö
				u ⇔ u									
				ü ⇔ ü									

The above considerations suggest that the paradigmatic patterns of linking vowels are characterized by a low degree of uncertainty, due to the bias on the contents of paradigm cells resulting from the type frequency of morphosyntactic types. Where the initial substring is needed for predicting the linking vowel (in the case of conformist suffixes), paradigms have abundant cells, what is more, some of these have word forms of exceptionally large token frequency (typically plural nouns and adjectives and present indicative 1sg nondefinite verbs). In the case of nonconformist suffixes no such other cells are needed, since the linking vowel before them can be predicted from any cell in which the exponent is harmonically alternating. Such suffixes are much fewer in number, too.

So what happens when a new loan or a nonsense word has to be suffixed? A widely held view is that there always is a productive pattern for such purposes (e.g., Yang 2016). This, however, is not the case in Hungarian: only in the case of nouns is there an obvious productive pattern, nonlowering, that is, the -e-, the -o-, and the -ö- classes are open, which can be attributed to their great type frequency. But even here there are harmonically mixed forms (e.g., *haver*, *partner*, *fotel* mentioned above) whose paradigm class (and hence the quality of the linking vowel) cannot be predicted with full certainty. Although these words typically vacillate with exponents without a linking vowel, their paradigm class is -o-, -e-, or -o/e-, respectively. Several factors may play a role in determining the paradigm class: morphological, phonological beyond their vocalism, frequency, even semantic/pragmatic (cf. Hayes et al. 2009, Forró 2013, Forró & Kálmán 2014, Rebrus et al. 2023ab); it is often difficult or impossible to clearly distinguish the different factors. Uncertainty is even greater in adjectives: some of them are unambiguously lowering, others unambiguously nonlowering, but the borderline is often fuzzy, and semantic factors play an even more significant role (for instance, ethnonyms and language names never lower, while prototypical adjectives do). It is not surprising then that in want of an obvious productive pattern new loanwords often vacillate, they are assigned to the -a/o- or -e/ö- paradigm classes depending on their harmony. (If their harmony is front unrounded, there is no vacillation, only the -e- class can host such words.) Furthermore, the choice of the linking vowel may depend on the frequency of the usage of morphosyntactic categories, which is a sign of the nonhomogeneity of paradigms (cf. Kálmán et al. 2012, Rebrus & Szigetvári 2022). The notion of productivity is not directly

applicable to verbs at all: newly acquired loan verbs are obligatory suffixed, most commonly with the conformist verbalizers *-l* or *-z*. The “linking vowel” of these suffixes often defies the harmonic requirements expected of the loan verbal “root”, we often find antiharmony (e.g., *brillír-o-z* ‘excel’, *fixír-o-z* ‘stare’, *síb-o-l* ‘smuggle’, *slissz-o-l* ‘sneak’) or the even more marked counterharmony (e.g., *csekk-o-l* ‘check’, *gründ-o-l* ‘establish’, *curükk-o-l* ‘reverse’, *stír-ö-l* ‘stare’), both strongly speaker-dependent (cf. Rebrus et al. 2023b).

In sum, we can state that there is no obvious productive pattern (or even patterns) referring to the basic word categories. The quality of the linking vowel in novel words depends on many factors and is subject to a large degree of uncertainty, which is manifested in vacillation, that is, filling paradigm cells with multiple forms: overabundance. The paradigm-based model predicts that if the linking vowel of a conformist exponent is nativized and attains large enough token frequency, it will be determinative for its paradigm class: the linking vowel of other exponents of its type will follow suit, including systematic vacillation.

## 5 Summary

This paper argues that atomistic models of morphophonology, which are based on segmentation into morphemes, assigning specifications to them, and manipulating these specifications, are often overly complicated, in addition, simply unable to capture some correlations that can be explicitly documented. We believe that this is due to the fact that such theories interpret linguistic patterns occurring on the surface, such as the quality of linking vowels and other suffix-initial vowels, as consequences of abstract features and configurations stipulated in the underlying representation. The basis of this mechanism is a discrete system (composed of a finite set of distinct elements) that functions in an algebraic and logical way, defining the objects to be distinguished atomistically. As a result, it can only capture clear-cut categories. On the contrary, a holistic view takes frequent patterns to be emergent phenomena, which are constituted stochastically, taking into account their frequency. As a result, the boundaries of the emergent categories are generally not clear cut.

We have introduced a model that describes the surface patterns of linking vowels *as is*, using analogical relations between paradigms and paradigm cells, without hidden abstractions. The system compares whole word forms in paradigm cells with the help of their initial and final substrings and the emerging patterns of linking vowels. This allows the expression of phenomena and correspondences which a morpheme-based theory cannot or does not aim to examine. They include generalizations across morphemic status and suffix type, the nonhomogeneity of paradigms, the role of frequency, or the appearance and suspension of systematic vacillation.

## References

- Ackerman, Farrell and Robert Malouf. 2016. Implicative relations on word-based morphological systems. In: Andrew Hippisley and Gregory Stump (eds.), *The Cambridge handbook of morphology*. Cambridge: CUP. 297–328. <https://doi.org/10.1017/9781139814720.012>

- Ackerman, Farrell, James P. Blevins, and Robert Malouf. 2009. Parts and wholes: Implicative patterns in inflectional paradigms. In: James P. Blevins and Juliette Blevins (eds.), *Analogy in grammar: Form and acquisition*. Oxford: OUP. 54–82. <https://doi.org/10.1093/acprof:oso/9780199547548.003.0003>
- Ackerman, Farrell, Robert Malouf and James P. Blevins. 2016. Patterns and discriminability in language analysis. *Word Structure* 9(2): 132–155. <https://doi.org/10.3366/word.2016.0091>
- Aronoff, Mark. 1994. *Morphology by itself*. Cambridge: MIT Press.
- Blevins, James P. 2016. *Word and paradigm morphology*. Oxford: OUP.  
<https://doi.org/10.1093/acprof:oso/9780199593545.001.0001>
- Blevins, James P. and Juliette Blevins (eds.). 2009. *Analogy in grammar: Form and acquisition*. Oxford: OUP.  
<https://doi.org/10.1093/acprof:oso/9780199547548.001.0001>
- Blevins, James P., Petar Milin, and Michael Ramscar. 2017. The Zipfian Paradigm Cell Filling Problem. In: Ferenc Kiefer, James P. Blevins, and Huba Bartos (eds.), *Perspectives on Morphological Organization: Data and Analyses*. Leiden: Brill. 139–158. [https://doi.org/10.1163/9789004342934\\_008](https://doi.org/10.1163/9789004342934_008)
- Breiss, Canaan. 2022. When Bases compete: a voting model of Lexical Conservatism. Ms., MIT.
- Bybee, Joan. 2007. *Frequency of use and the organization of language*. Oxford: OUP.  
<https://doi.org/10.1093/acprof:oso/9780195301571.001.0001>
- Croft, William. 2001. *Radical Construction Grammar: Syntactic theory in typological perspective*. Oxford: OUP.  
<https://doi.org/10.1093/acprof:oso/9780198299554.001.0001>
- Crysmann, Berthold and Olivier Bonami. 2017. Atomistic and holistic exponence in Information-based Morphology. In: Stefan Müller (ed.), *Proceedings of the 24th International Conference on Head-Driven Phrase Structure Grammar*, CSLI Publications. 141–161. <https://doi.org/10.21248/hpsg.2017.8>
- Forró, Orsolya. 2013. *Ingadozás a magyar előségi harmóniában: szempontok a variabilitás szinkroniájának és diakroniájának feltárásához és értelmezéséhez* [Variation in Hungarian frontness harmony: Aspects for the exploration and interpretation of the synchrony and diachrony of variability]. PhD dissertation, Pázmány Péter Catholic University.
- Forró, Orsolya and László Kálmán. 2014. “Lökött” korlátok nyomában [In search of “loony” constraints]. Talk given at MTA NYTI, March 13.
- Gafos, Adamantios I. and Amanda Dye. 2011. Vowel harmony: Opaque and transparent vowels. In: Marc van Oostendorp, Colin Ewen, Elizabeth Hume, and Keren Rice (eds.), *Blackwell Companion to Phonology*. Wiley-Blackwell. 2164–2189. <https://doi.org/10.1002/9781444335262.wbctp0091>
- Hayes, Bruce, Kie Zuraw, Péter Siptár, and Zsuzsa Londe. 2009. Natural and unnatural constraints in Hungarian vowel harmony. *Language* 85: 822–863. <https://doi.org/10.1353/lan.0.0169>
- Hockett, Charles F. 1954. Two models of grammatical description. *Word* 10: 210–231.  
<https://doi.org/10.1080/00437956.1954.11659524>
- Kálmán, László. 2007. *Holisztikus szemlélet a nyelvészetben* [The holistic view in linguistics]. Szabad Válogatások.
- Kálmán, László. 2008. *A mögöttes és ami mögöttes van* [Underlying forms and what underlies them]. Szabad Válogatások 5.
- Kálmán, László. 2012. Abstractness meets explanation. Paper presented at MaSzAT (Hungarian Round Table on Semantics), 20 December.
- Kálmán, László. 2013. *Az elvontság nem boldogít!* [Abstractness doesn’t make you happy!]. Paper presented at MTA NYTI, 15 January.
- Kálmán, László. 2014a. What can constructions do? *Theoretical Linguistics* 40/1–2: 137–147.  
<https://doi.org/10.1515/tl-2014-0005>
- Kálmán, László. 2014b. Making sense of sense. Presentation given at MTA NYTI, 26 June.
- Kálmán, László. 2015. Some analogical phenomena in Hungarian syntax. Presentation given at MTA NYTI, 3 December.
- Kálmán, László. 2018a. *Ahány szó, annyi szófaj* [Each word its own part of speech].  
<https://qubit.hu/2018/02/17/ahany-szo-annyi-szofaj>
- Kálmán, László. 2018b. *A régens és a csápjai: mit jelent az a szó, hogy figyel* [The governor and its tentacles: what does the word figyel mean?]  
<https://qubit.hu/2018/08/04/a-regens-es-a-csapjai-mit-jelent-az-a-szo-hogy-figyel>

- Kálmán László. 2018c. Alfászor vagy alfászor – mindegy? Paradigmaváltás a nyelvészetben [Alfaszor or alfászor ‘alpha times’ – does it matter? Paradigm change in linguistics].  
<https://qubit.hu/2018/09/03/alfaszor-vagy-alfaszor-mindegy-paradigmavaltas-a-nyelveszetben>
- Kálmán László. 2018d. A gyakoriság szerepe a nyelvben: balettből vagy balettből? [The role of frequency in language: balettből or balettből ‘from the ballet’].  
<https://qubit.hu/2018/11/24/a-gyakorisag-szerepe-a-nyelvben-balettbol-vagy-balettbol>
- Kálmán László 2019a. Kopje koffie, avagy a hollandok esete a kicsinyítéssel [Kopje koffie, the case of the Dutch with diminutives]. <https://qubit.hu/2019/02/24/kopje-koffie-avagy-a-hollandok-esete-a-kicsinyitessel>
- Kálmán László. 2020a. -ban/-ben vagy -ba/-be? Megoldjuk okosba [-ban/-ben or -ba/-be ‘in’? We’ll solve it smart]. <https://qubit.hu/2020/08/09/ban-ben-vagy-ba-be-megoldjuk-okosba>
- Kálmán László. 2020b. Mondatolj apádnak évekenként: a képző, az a furcsa és sokszínű állatfajta [Sentencify for your father yearsly: Derivation, that weird and diverse beast].  
<https://qubit.hu/2020/08/22/mondatolj-apadnak-evekenkent-a-kepzo-az-a-furcsa-es-sokszinu-allatfajta>
- Kálmán László. 2020c. Ha van apa ~ anya és apu ~ anyu, akkor miért nincs puli ~ nyuli és paripa ~ nyaripa? [If we have mom ~ dad and mommy ~ daddy, then why don’t we have monkey ~ dankey and mongrel ~ dangrel?]  
<https://qubit.hu/2020/10/25/ha-van-apa-anya-es-apu-anyu-akkor-miert-nincs-puli-nyuli-es-paripa-nyaripa>
- Kálmán László. 2020d. Adoda a kenyeret! Jó, de mi az a kenyér? [Gimme the bread? Okay, but what is bread?]  
<https://qubit.hu/2020/11/08/adoda-a-kenyeret-jo-de-mi-az-a-kenyer>
- Kálmán László. 2020e. A nyelv leírására létezik jobb módszer, mint szabályok tízezreinek felállítása [There is a better way of describing language than setting up myriads of rules].  
<https://qubit.hu/2020/10/03/a-nyelv-leirasara-letezik-jobb-modszer-mint-szabalyok-tizezreinek-felallitasa>
- Kálmán László. 2021a. Miért bántja a fület az a szó, hogy oltakozik? [Why does the word oltakozik ‘get vaccinated’ sound odd?]  
<https://qubit.hu/2021/04/18/miert-bantja-a-fulet-az-a-szo-hogy-oltakozik>
- Kálmán László. 2021b. Oltakozunk, oltakozgatunk? [Do we get vaccinated, do we regularly get vaccinated?]  
<https://qubit.hu/2021/05/01/oltakozunk-oltakozgatunk>
- Kálmán, László. 2021c. Defectivity in syntax and the Hungarian zero copula. Poster presented at the 15th International Conference on the Structure of Hungarian, Pécs, 25 August.
- Kálmán, László and Rebrus Péter. 2009. Valóban megmagyarázhatatlanok a magyar infinitívusz toldalékai? [Are infinitival suffixes really unexplainable?]. In: Maleczki, Márta; Németh, T. Enikő (eds.) A mai magyar nyelv leírásának újabb módszerei 7: Modellek, elméletek és elvek érvényessége nyelvi adatok tükrében. Szeged: SZTE Általános Nyelvészeti Tanszék, 153–156.
- Kálmán, László, Péter Rebrus, and Miklós Törkenczy. 2012. Possible and impossible variation. In: Ferenc Kiefer, Mária Ladányi, and Péter Siptár (eds.), Current issues in morphological theory. (Ir)regularity, analogy and frequency. Papers from the 14th International Morphology Meeting. Amsterdam/Philadelphia: John Benjamins. 23–49. <https://doi.org/10.1075/cilt.322.02kal>
- Kiefer, Ferenc. 1999. Alaktan [Morphology]. In: Katalin É. Kiss, Ferenc Kiefer, and Péter Siptár (eds.), Új magyar nyelvtan. Budapest: Osiris.
- Kiparsky, Paul. 1971. Historical linguistics. In: Paul Kiparsky (1982) Explanation in Phonology. Dordrecht: Foris, 57–80. <https://doi.org/10.1515/9783111666242>
- Kiparsky, Paul and Karl Pajusalu. 2003. Towards a typology of disharmony. The Linguistic Review 20: 217–241. <https://doi.org/10.1515/tlir.2003.009>
- Mártonfi, Attila. 2006. A magyar toldalék-morfémák rendszere [The system of Hungarian suffixes]. PhD dissertation. Eötvös Loránd University, Budapest.
- McCarthy, John J. 2007. Hidden generalizations: Phonological opacity in Optimality Theory. London: Equinox.
- Nádasdy, Ádám and Péter Siptár. 1994. A magánhangzók [Vowels]. In: Kiefer Ferenc (ed.), Strukturális magyar nyelvtan 2: Fonológia. Budapest: Akadémiai Kiadó. 42–182.
- Novák, Attila 1999. Inflectional paradigms in Hungarian. The conditioning of suffix- and stem-alternations. MA Thesis. Eötvös Loránd University, Budapest.
- Polgárdi, Krisztina and Péter Rebrus 1998. There is no labial harmony in Hungarian: a Government Phonology analysis. In: Casper de Groot and István Kenesei (eds.), Approaches to Hungarian 6: Papers from the Amsterdam conference. Szeged: JATEPress. 3–20.

- Rácz, Péter and Péter Rebrus. 2012. Variation in the possessive allomorphy of Hungarian. In: Ferenc Kiefer, Mária Ladányi and Péter Siptár (eds.), *Current Issues in Morphological Theory: (Ir)regularity, analogy and frequency. Selected papers from the 14th International Morphology Meeting, Budapest, 13–16 May 2010*. Amsterdam/Philadelphia: John Benjamins. 51–64.
- Rácz, Péter, Péter Rebrus, and Miklós Törkenczy. 2021. Attractors of variation in Hungarian inflectional morphology. *Corpus Linguistics and Linguistic Theory* 17(2): 287–317.  
<https://doi.org/10.1515/cllt-2018-0014>
- Rebrus, Péter. 2000a. Morfofonológiai jelenségek [Morphophonological phenomena]. In: Ferenc Kiefer (ed.), *Strukturális magyar nyelvtan 3: Morfológia*. Budapest: Akadémiai Kiadó. 763–947.
- Rebrus Péter. 2000b. A magyar igei paradigma kivételeiről [On the exceptions to Hungarian verbal paradigms]. In: Marianne Bakró-Nagy, Zoltán Bánréti, and Katalin É. Kiss (eds.), *Újabb tanulmányok a strukturális magyar nyelvtan és nyelvtörténet köréből*. Budapest: Osiris. 249–273.
- Rebrus Péter. 2013. Miért nincs j? [Why is there no j?]. In: Attila Benő, Emese Fazakas, and Edit Kádár (eds.), *„...hogy legyen a víznek lefolyása...” Köszöntő kötet Szilágyi N. Sándor tiszteletére*. Kolozsvár: Erdélyi Múzeum Egyesület Kiadó. 383–401.
- Rebrus, Péter. 2019. A magyar névszói inflexiók rendszer a szó és paradigma modellben [The Hungarian nominal inflectional system in the Word and Paradigm model]. *Nyelvtudományi Közlemények* 115: 109–147. <https://doi.org/10.15776/NyK/2019.115.4>
- Rebrus, Péter and Krisztina Polgárdi. 1997. Two default vowels in Hungarian? In Geert Booij and Jeroen van de Weijer (eds.), *Phonology in progress – Progress in phonology (HIL Phonology Papers III)*. Hague: Holland Academic Graphics. 257–275.
- Rebrus, Péter and Péter Szigetvári. 2022. Between adjective and noun. *Acta Linguistica Academica* 69: 188–205. <https://doi.org/10.1556/2062.2022.00536>
- Rebrus, Péter and Miklós Törkenczy 2005. Uniformity and contrast in the Hungarian verbal paradigm. In: Laura J. Downing, Tracy A. Hall, and Renate Raffelsiefen (eds.), *Paradigms in phonological theory*. Oxford: OUP. 262–295. <https://doi.org/10.1093/acprof:oso/9780199267712.003.0010>
- Rebrus, Péter and Miklós Törkenczy. 2008. Morfofonológia és a lexikon [Morphology and the lexicon]. In: Ferenc Kiefer (ed.), *Strukturális magyar nyelvtan 4: A lexikon*. Budapest: Akadémiai Kiadó. 683–786.
- Rebrus, Péter and Miklós Törkenczy. 2009. Covert and overt defectiveness in paradigms. In: Sylvia Blaho and Curt Rice (eds.), *Modeling ungrammaticality in optimality theory*, London: Equinox Publishing. 195–234 .
- Rebrus, Péter and Miklós Törkenczy. 2019. Magyar harmónia: a dolgok állása [Harmony in Hungarian: the state of affairs]. *Általános Nyelvészeti Tanulmányok XXXI*: 233–333.
- Rebrus, Péter and Miklós Törkenczy. 2021. Harmonic Uniformity and Hungarian front/back harmony. *Acta Linguistica Academica* 68/1–2:175–206 <https://doi.org/10.1556/2062.2021.00475>
- Rebrus, Péter, Péter Szigetvári, and Miklós Törkenczy. 2012. Dark secrets of Hungarian vowel harmony. In: Eugeniusz Cyran, Henryk Kardela, and Bogdan Szymanek (eds.), *Sound, structure and sense: Studies in memory of Edmund Gussmann*. Lublin: Wydawnictwo KUL. 491–508.
- Rebrus, Péter, Péter Szigetvári, and Miklós Törkenczy. 2017. Asymmetric variation. In: Geoff Lindsey and Andrew Nevins (eds.), *Sonic signatures*. Amsterdam/Philadelphia: John Benjamins Publishing Company. 163–187. <https://doi.org/10.1075/lfab.14.c10>
- Rebrus, Péter, Péter Szigetvári, and Miklós Törkenczy. 2020. Variation, the Height Effect, and Disharmony in Hungarian Front/Back Harmony. In *Proceedings of the 2019 Annual Meeting on Phonology*. Edited by Hyunah Baek, Chikako Takahashi, and Alex Hong-Lun Yeung. Washington, DC.: Linguistic Society of America. DOI: <https://doi.org/10.3765/amp.v8i0.4750>
- Rebrus, Péter, Péter Szigetvári, and Miklós Törkenczy. 2022. Aszimmetrikus mintázatok a magyar alaktanban [Asymmetric patterns in Hungarian morphology]. Paper presented at the Össz nyelvész conference, 16–17 December.
- Rebrus, Péter, Péter Szigetvári, and Miklós Törkenczy. 2023a. How morphological is Hungarian vowel harmony? Supplemental Proceedings of the 2022 Annual Meeting on Phonology. Washington, DC: Linguistic Society of America.

- Rebrus, Péter, Péter Szigetvári, and Miklós Törkenczy. 2023b. Morphological restrictions on vowel harmony: The case of Hungarian. To appear in Peter Ackema, Sabrina Bendjaballah, Eulàlia Bonet, and Antonio Fábregas (eds.), *The Wiley Blackwell Companion to Morphology*. John Wiley & Sons.
- Siptár, Péter. 2011. Alakváltozatok, allomorfok, alternációk [Allomorphs and alternations]. *Magyar Nyelv* 107: 147–160.
- Siptár, Péter and Miklós Törkenczy. 2000. *The phonology of Hungarian*. Oxford & New York, NY: OUP.
- Steriade, Donca. 1999. Lexical Conservatism. in *Linguistics in the Morning Calm, Selected Papers from SICOL 1997*, Linguistic Society of Korea, Hanshin Publishing House. 157–179.
- Steriade, Donca. 2000. Paradigm uniformity and the phonetics/phonology boundary. In: Janet B. Pierrehumbert and Michael B. Broe (eds.), *Papers in Laboratory Phonology 6*. Cambridge: CUP. 313–334.
- Szilágyi N., Sándor. 2004. *Elmélet és módszer a nyelvészetben: Különös tekintettel a fonológiára* [Theory and method in linguistics: With special regard to phonology]. (Erdélyi tudományos füzetek). Kolozsvár: Erdélyi Múzeum Egyesület Kiadó.
- Szentgyörgyi, Szilárd. 1999. Lowering: The interaction of phonology and morphology in Hungarian. PhD dissertation, University of Szeged.
- Trón, Viktor and Péter Rebrus. 2001. Morphophonology and the hierarchical lexicon. *Acta Linguistica Hungarica* 48. 2001: 101–135.
- Trón, Viktor and Péter Rebrus. 2005. Re-presenting the past: Contrast and uniformity in Hungarian Past Tense Suffixation. In: Christopher Piñón and Péter Siptár (eds.), *Approaches to Hungarian 9: Papers from the Düsseldorf Conference*. Budapest: Akadémiai Kiadó, 305–327.
- Vihman, Marilyn M. and Tamar Keren-Portnoy (eds.) 2013. *The emergence of phonology: Whole-word approaches and cross-linguistic evidence*. Cambridge: CUP. <https://doi.org/10.1017/CBO9780511980503>
- Vago, Robert M. 1980. *The sound pattern of Hungarian*. Washington, DC: Georgetown University Press.
- Wunderlich, Dieter. 2012. Polarity and constraints on paradigmatic distinctness. In: Jochen Trommer (ed.), *The morphology and phonology of exponence*. Oxford: OUP. 160–194. <https://doi.org/10.1093/acprof:oso/9780199573721.003.0006>
- Yang, Charles D. 2016. *The Price of Linguistic Productivity: How Children Learn to Break the Rules of Language*. Cambridge, MA: The MIT Press. <https://doi.org/10.7551/mitpress/9780262035323.001.0001>