The Hungarian Language
A Short Descriptive Grammar

Beáta Megyesi

Hungarian, also called Magyar, traditionally belongs to the Ob-Ugric languages (e.g. Khanty and Mansi) of the Finno-Ugric branch of Uralic. Hungarian is the official language of the Republic of Hungary, and has approximately fifteen million speakers, of which four million reside outside of Hungary.

In this paper a description of Hungarian phonology, morphology and syntax follows. The sections are based on Benkö & Imre (1972), Rácz (1968), Olsson (1992) and Abondolo (1992).

1.1 Phonology

Hungarian has a rich system of vowels and consonants.

The vowel inventory consists of 14 phonemes of which one can distinguish 5 pairs, consisting of short and long counterparts; these are i - í, o - ó, ö - ö, u - ú, ü - ü. The remaining four are e - é and a - á. Short vowels, if they are marked, take an umlaut (¨), while long vowels are indicated by an acute (´) or with a double acute accent (´´) which is a diacritic unique to Hungarian. Long vowels are usually somewhat tenser than their short counterparts with two exceptions; *e* is low while *é* is higher mid and *á* is low whereas *a* is lower mid and slightly rounded (Abondolo, 1992). Vowel length is independent of prosodic factors such as stress.

The vowels may be interconnected through the laws of vowel harmony which means that suffixes, which may assume two or three different forms, usually agree in backness with the last vowel of the stem. In other words, front vs. back alternatives of suffixes are selected according to which vowel(s) the stem contain(s) (Benkö & Imre, 1972). The vocalism of stems, classified by Abondolo (1992), is inherently back for all stems containing at least one back vowel and for most verbs with the sole vowel *i* or *í*. For all other stems the vocalism is front. In regard to vowel harmony, *i* and *í* are neutral and can be used with either front (high) or back (low) vowels. Harmony causes the following alternations among suffix combinations: a/e (-ban/-ben - 'in'), á/é (-nál/-nél - 'at'), ó/ö (-ból/-böl - 'from'), u/ü (-ul/-ül - 'for, by') and o/e/ö (-hoz/-hez/-höz - 'to').

The vowels may show even paradigmatic alternations as long and short vowels (é vs. e and á vs. a, see the example below) alternate in some stems (Benkö & Imre, 1972).

Example 1

teh*é*n	teh*e*n-et	f*a*	f*á*-t
cow:NOM	cow-ACC	tree:NOM	tree-ACC
'cow'		'tree'	

When building a Hungarian corpus it is usual to delete the accent and the umlaut from the vowel and mark vowel length as well as the umlaut by numerals which follow the vowel: 1 denotes the acute accent (e.g. ó -> o1), 2 the umlaut (ü -> u2) and 3 the double acute accent (e.g. ö -> o3). This notation can be useful when automatically deriving rules from a corpus because of the paradigmatic alternations of the long vs. short vowels.

There are totally 25 consonants, which can be determined according to the manner and the place of articulation, voicing and quantity. The consonants are shown in the table below.

Table 1. The Hungarian consonant chart, given with regular orthographic symbols. Phonetic values are given in square brackets.

	Labial/labiodental		dental/alveolar		palatal		velar		glottal
	-voice	+voice	-voice	+voice	-voice	+voice	-voice	+voice	-voice
Stops	p	b	t	d	ty [t']	gy [d']	k	g	
Affricates			c [ts]	dz	cs [c']	dzs [j']			
Fricatives	f	v	sz [s]	z	s [s']	zs [z']			h
Nasals		m		n		ny [n']			
Laterals				l					
Tremulants				r					
Glides						j, ly[1]			

Consonant length is distinctive and is independent of vowel length and of prosodic factors such as stress. Each consonant can be pronounced short or long, where the last mentioned has almost double the length of short consonants and is written by doubling the letter (gg), or the first element of a digraph (ggy) (Benkö & Imre, 1972). Many of the long consonants occur at morpheme boundaries or root-finally in foreign vocabulary (Abondolo, 1992).

Assimilation is either full or partial and can be indicated by orthography. In the next section, the most important morpheme specific assimilation rules will be presented.

1.2 Morphology

Hungarian is basically agglutinative, i.e. grammatical relations are expressed by means of affixes. For understanding the function of different affixes and how they interact the following section will give an overview of these for different parts of speech. The types of homography will also be described under the main categories.

1.2.1 Articles

The articles [DET][2] are invariable for number, person, gender and case. The *indefinite article* is *egy*, while the *definite article* has two forms *a* and *az*, where the first is used before consonants and the latter before vowels, similar to English indefinite articles.

With regard to homography, the form of the definite article *az* can also be a demonstrative pronoun and the indefinite article can be homonymous to the numeral 'one' (Pajzs, 1996).

1.2.2 Nouns

Every Hungarian noun [FN] may be analysed as a stem followed by three positions in which inflectional suffixes can occur. Thus, nouns are inflected for number, person (possessor) and case, with the relevant suffixes attached in that order (Abondolo, 1992). Any or all of the three inflectional suffixes may be occupied by a zero suffix which denotes either singular number (first position), absence of possessor (second position) or nominative case (third position) (Abondolo, 1987). Thus we have:

Example 2

gyereke-Ø-m-en	gyereke-k-Ø-en	gyereke-i-m-Ø
child-Ø-1POSS-SUPESS	child-PL-Ø-SUPESS	child-PL-1POSS-Ø
'on my child'	'on children'	'my children'

There is no grammatical gender. The personal pronoun ö means both 'he' and 'she'.

---

[1] j and ly are pronounced alike.
[2] Within brackets [] are those PoS tags that are used in the Hungarian corpus.

Describing Hungarian, many authors, not without reason, 'forget' to mention something about the complicated system of noun stem alternation. Here, an outline of this system based on the summary by Abondolo (1987) is presented.

Noun stems may end in either a consonant or a vowel. All stems with final *a* or *e* are lengthened (*á* vs. *é*) before most suffixes, whether derivational or declensional.

Example 3

lámp*a*	=>	lámp*á*-m	kef*e*	=>	kef*é*-m
lamp:NOM		lamp-1POSS	brush:NOM		brush-1POSS
'lamp'		'my lamp'	'brush'		'my brush'

There also exists a very special stem form, called 'oblique stem', which occurs with several nouns. This oblique stem form differs from the nominative singular in a way that there is present a stem-final *a* or *e*, or there is absent the stem-penultimate *o*, *ö* or *e*, or both. The example below illustrates this where *fal* 'wall', *dal* 'song', *gyomor* 'stomach', *nyomor* 'misery', *sátor* 'tent' and *mámor* 'rapture' (Abondolo, 1987).

Example 4 Different nouns have different inflectional patterns based on their oblique stems

Stem	present *a*	normal	Absent *o*	normal	absent *o* and present *a*	normal
Nominative	fal	dal	gyomor	nyomor	sátor	mámor
Oblique	fal-a	dal-	gyomr	nyomor-	sátr-a	mámor-
Accusative	fala-t	dal-t	gyomr-ot	nyomor-t	sátra-t	mámor-t

There are words whose oblique stems not only have final *a* or *e* but also *v* instead of *u*, e.g. *falu* 'village' whose oblique stem is *falva-*, as in the form *falva-k* 'villages'.

Nominative stems with the long vowels *á* and *é* also change to short *a* vs. *e* in the penultimate position when becoming oblique stem, e.g. *madár* 'bird', oblique stem *madara-*, thus *madara-k* 'birds'. This phenomenon may cause problem when automatically identifying stems.

1.2.2.1 Number

The category number is realised as singular and plural. There are two plural suffixes. The suffix *-k* is preceded by an epenthetic vowel after a consonant final stem (Olsson, 1992).

Example 5

	SINGULAR	PLURAL
'university'	egyetem	egyetemek
'student'	diák	diákok

The other plural suffix is *-i*, which is used only when person suffixes are present.

Example 6

gyereke-i-nke-t
child-PL-1PL-ACC
'our children'

1.2.2.2 Person/Possession

Possession is usually indicated with a personal suffix on the possessed noun. The forms vary for number and person, as shown in Appendix A under possessives. If there is a chain of possessors, the last possessor, closest to the head, takes a dative case marker *-nak/-nek* in addition to the possessive suffix. Consider the following example.

Example 7

az apá-m	barát-*já-nak*	a könyv-*e*
the father-1SG.POSS	friend-3SG.POSS-DAT/GEN	the book-3.SG.POSS
'my father's friend's book'		

The suffix *-já-* marks possession by *apám* 'my father' and *-nak* signals the pending possessive ending *-e* (Campbell, 1991).

Nouns consisting of a stem with a possessive ending followed by a cases suffix are in most cases homonymous. For example, the word fej+é+nek with the suffix *-é-* as a possessive suffix means 'to the possession of his head', while the same word with *-é-* as a paradigmatic alternation to the vowel *e* means 'to his head' (Pajzs, 1996).

1.2.2.3 Case

Hungarian has a complex case system involving 16 to 24 distinct forms to mark that an NP bears some identifiable grammatical or semantic relation to the rest of the sentence. The case suffixes may be classified into two groups, non-local and local. The non-local cases express primary syntactic or adverbial functions, such as subject, direct and indirect object, possessor or instrument. The local cases show concrete spatial and kinetic conditions such as interior vs. exterior, stationary vs. moving (Abondolo, 1987).

There are different assumptions about the exact number of case suffixes. I count 19 and the names of the cases and the forms (allomorph) of the suffixes are given in the Appendix with examples (*ház* 'house', *öt* 'five'). Unfortunately, there is no space for explaining the function of each case but hopefully the examples illustrate their functions.

Case suffixes are the same both in singular and in plural. The plural suffix always precedes the case suffix, as the examples below show:

Example 8

    a   háza-k-ban   a háza-i-m-ban
    the house-PL-INESS   the house-PL-1SG-INESS
    'in the houses'   'in my houses'

Assimilation at juncture takes place in instrumental and translative case which may cause problems in automatic tagging systems. The initial -v- in the instrumental suffix -val/-vel and the translative suffix -vá/vé is assimilated to a preceding consonant (Olsson, 1992).

Example 9

                  Instrumental   Translative
    ház 'house'   ház-zal   ház-zá
    öt 'five'   öt-tel   öt-té

The case suffixes may also occur as stems and take personal suffixes if not postposed to a noun. They are then usually regarded as pronouns or adverbs in traditional Hungarian grammar.

Table 2. The Hungarian case system, listed with each allomorph of the case suffix in singular and plural and exemplified by the words ház 'day', and öt 'five'

Case [Tag]	Suffixes	Examples	
Nominative [NOM]	-	Ház	house
Accusative[3] [ACC]	-t, -ot, -et, -öt	Házat	house
Dative-genitive[4] DAT]	-nak, -nek	Háznak	of the house
Instrumental [INS]	-(V)al, -(V)el	Házzal	with the house
Essive-modal [SOC]	-stul, -stül	házastul	with the house and its parts
Translative [FAC]	-(V)á, (V)é	házzá	into a house
Causal-final [CAU]	-ért	házért	for the house
Illative [ILL]	-ba, -be	házba	into the house
Sublative [SUB]	-ra, -re	házra	onto the house
Allative [ALL]	-hoz, -hez, höz	házhoz	to the house

---

[3] The accusative case ending in certain constructions may be zero (-Ø-) if the object is a noun with a possessive personal endig, e.g. *eladom a házam/házamat* 'I sell my house.'.
[4] The reason for marking the genitive and the dative cases as the same is, that dative may mark not only the indirect object but also the possessor.

Inessive [INE]	-ban, -ben	házban	in the house
Superessive [SUP]	-n, -on, -en, -ön	házon	on the house
Adessive [ADE]	-nál, nél	háznál	at the house
Elative [ELA]	-ból, -böl	házból	out of the house
Delative [DEL]	-ról, röl	házról	from (top of) the house
Ablative [ABL]	-tól, töl	háztól	from (nearby) the house
Terminative [TER]	-ig	házig	as far as the house
Formal [FOR]	-ként	házként	as a house
Temporal [TEM]	-kor	ötkor	at five

1.2.3 Pronouns

The use of personal pronouns [NM] is not frequent in Hungarian because it is a pro-drop language. They basically have two cases: nominative and accusative. The singular forms in nominative are *én* (I), *te*('you'), and *ö*('he', 'she'). Third person plural in the nominative case can be derived from the corresponding singular by adding the plural suffix *-k* (the same as the plural suffix for nouns) to the singular stem, e.g. *ö+k*. For the other plural forms there is no such simple connection (*mi* 'we' and *ti* 'you').

The personal pronouns in first and second person singular in accusative occur usually without the accusative suffix *-t*. (*engem* 'me', *téged* 'you'). The first and second plural forms in accusative are, on the other hand, constructed as nominative form + corresponding possessive suffix + accusative marker -t, i.e. the same marker for possessive suffixes and for the accusative for nouns, e.g. *mi+nk+et* 'us' and *ti+tek+et* 'you'. The third person in accusative can be derived as nominative + (plural ending) + accusative ending, thus *ö+t* 'her/him' *ö+k+et* 'them'.

By adding the enclitic (personal) markers to the case endings of nouns, oblique forms are made. These correspond to prepositional phrases in English, and are often regarded as pronouns in different cases, or as adverbs because of their adverbial function in the sentence. The enclitic markers with examples are shown in Table 3.

Table 3. The enclitic markers for Hungarian pronouns

	Singular	Plural	Examples - SG	Examples - PL
1	-m	-unk, ünk	nek-em 'to me'	nek-ünk 'to us'
2	-d	-tok, -tek	nek-ed 'to you'	nek-tek 'to you'
3	-i, -e	-ik, -ük	nek-i 'to him/her'	nek-ik 'to them'

Reflexive pronouns [NM] consist of the word *mag* 'pit, nucleus' and a possessive personal suffix, listed in Appendix A. Besides the reflexive function they also have a non-reflexive function to emphasise the personal pronouns. In reflexive function the accusative ending *-at* is 'obligatory', except the first and second person where the nominative form is common.

Possessive pronouns [NM] which serve to express possession have the following forms:

enyém 'mine' miénk 'ours'
tied 'yours' tietek 'yours'
övé 'his/hers' övék 'theirs'

Possessive pronouns cannot stand together with the noun head (the possessed entity). When the noun head is present a personal pronoun is used. Thus, the use of double markers on the pronoun and the noun head at the same time is not allowed in Hungarian.

Example 10

eny:ém   volt   a   könyv   a(z) (én) könyv-em
POSSPRON:1SG COP:PAST the book   the   I   book-POSS1SG
'It was my book.'   'my book'

The system of *demonstrative pronouns* [NM] consists of two categories; pronouns with front vowel mean 'near' in contrast to pronouns with back vowel which mean 'far', e.g. *ez/az* 'this/that'. Case endings are usually added to the pronouns and show both regressive (ez-nek => en-nek) and progressive (az-val

=> az-zal) assimilation. The demonstrative pronoun cannot take enclitic markers or possessive suffixes and takes the same position as the head nominal.

Example 11

> *ezek*   mögött a problémá-i-d   mögött
> this-PL behind   the problem-PL-2SG   behind
> 'behind these problems of yours'

The *interrogative pronouns* [KSZ] are based on the stems *ki* 'who', *mi* 'what' and *hol* 'where'. Their derivatives are different case suffixes as in the case of the demonstrative pronouns.

Other pronouns, e.g. relative and indefinite pronouns [NM] are compounds of the interrogative pronouns. *Relative pronouns* are derivable from interrogative pronouns which is done by the prefix *a-*. It has personal and non-personal counterparts and singular and plural forms as well as full declension. They follow the pattern of Hungarian nouns in case and number.

One can describe other type of pronouns in a similar way, as below.

Pronouns	Prefix	Pronominal stems	
Relative	a-	+ ki 'who'	+ hol 'where'
Indefinite	vala-	+ mi 'what'	+ hova 'where to'
Negative	se(n/m)	+ milyen 'what kind'	+ honnan 'where from'
Selective	bár- / akár-	+ mely 'which'	+ mikor 'when'
General	minden	+ mennyi 'how many'	+ meddig 'where to'

Each pronominal prefix can be added to the pronominal stems. Thus, the indefinite pronoun *valami* means 'something', *valaki* 'somebody', *valahol* 'somewhere', the selective pronoun *bárki* means 'anyone', *bármi* 'anything' and so on. The above mentioned pronouns also take case- and plural suffixes, e.g. *valamiben* 'in something'.

1.2.4 Adjectives

Adjectives [MN] can be used as nouns and are then declined fully, i.e. plural and case endings can be added to adjectives. There is also a special case only used by adjectives: the modal-essive with the forms *-en/-an/-on* and *-leg*, e.g. *gazdag-on* 'in a rich way'.

Comparative is formed by adding a (harmonic vowel) + -bb to the stem. Some comparative forms are suppletive, such as *sok* 'many' vs. *több* 'more'. Superlatives are formed by adding the prefix *leg-* to the comparative form, e.g. *rossz* 'bad', *rossz+abb* 'worse', *leg+rossz+abb* 'worst'

In Hungarian, adjectives in attributive position precede their head nouns and do not agree with them. On the other hand, adjectives in predicative position agree in number with the subject.

Example 12

a szép	virág-ok	a virágok	szép-ek
the beautiful	flower-PL	the flower-PL	beautiful-PL
'the beautiful flowers'		'The flowers are beautiful.'	

1.2.5 Numerals

Numerals [SZN] precede nouns and do not agree with them. When standing without the head noun they carry the same case endings as the nouns, pronouns or adjectives, e.g. *öt-ért* 'for five'.

1.2.6 Verbs

Hungarian verbs [IGE] may be analysed as a stem, followed by a tense/mood suffix, followed by a person-and-number suffix, see . The morphological manifestations of the tense/mood and person/number system are interconnected and will be discussed together. In Appendix A under the

Verb [IGE] category all allomorphs of different tense/mood and person/number categories are listed. Below just a few examples of different paradigms follow.

Example 13

ír-Ø-om	ír-t-am
write-PRS-1SG:DEF	write-PST-1SG:DEF
'I write'	'I wrote'

In verb conjugation the personal suffixes play a central role because free pronominal subjects are only present when there is emphasis on the person, as in above. Personal suffixes express first, second and third person not only with pronouns, nouns and postpositions to mark personal relations but also with verbal personal endings and on non-finite forms of the verb.

The verbal paradigm is split into two conjugations according to the definiteness of the complement of the verb - the direct object. Thus, each person suffix refers not only to the person and number of the subject, but also to the person or the definiteness (though not the number) of the object. The object is definite if it is a proper name, a noun with a definite article, a noun with a personal ending or a personal pronoun in the third person. Other pronouns in object position take verbs in indefinite conjugation (Benkö & Imre, 1972). Thus, there are two first person singular suffixes in the non-past form of the verb *ír* 'write': *-k* is used with an indefinite direct object and *-m* is used with definite objects. Note that both suffixes also refer to the first person singular noun.

Example 14

ír-o*k*		*egy*	könyv-et		ír-o*m*	*a*	könyv-et
write-1SGINDEF		*a*	book-ACC		write:1SGDEF	*the*	book-ACC
'I write a book.'					'I write the book.'		

Hungarian basically distinguishes between two tense features; past and non-past. The present tense suffix is zero (-Ø-) while past tense is marked by the *-t-* suffix. Note, that *-t* also marks the accusative case for noun, pronouns, etc.

Example 15

ír-Ø-ja	a könyvet	ír-t-a	a könyv-et
write-PRS-3SGDEF	the book-ACC	write-PST-3SGDEF	the book-ACC
'He/she is writing the book.'		'He/she was writing the book.'	

The future tense is made with the auxiliary '*fog*' with different personal endings + infinitive (*-ni*), as in *fogok/fogom adni* 'I'll give' depending on whether the complement is indefinite or definite (Campbell, 1991). The present tense used with the verb particle *meg-* can also be used in future tense, e.g. *megírom a könyvet* 'I'll write the book'.

There are three mood categories: indicative, subjunctive, which also functions as imperative, and conditional. In *indicative* mood, as it was mentioned above, the present and past tenses are made by means of personal endings for every person and number, and for definite and indefinite complement.

In imperative/subjunctive mood the marker is *-j*. In verbs in which the root ends in *t*, the *j* suffix of the resulting '*tj*' is phonetically realised as '*ss*' and this is sometimes indicated in the orthography: *más+j => mássz* 'Climb!' and *mos+ja => mossa* 'She washes it' (Benkö & Imre, 1972).

The conditional form is made by the marker *-n-* followed by a harmonic vowel, e.g. *adnám* 'I'd give' (definite object) and *kérnék* 'I would ask' (indefinite object) (Campbell, 1991).

The rich marking system is complicated because of allomorphic variation where some of the allomorphs are not phonologically indicated, but rather to avoid homonymy. For example, the first person singular past tense form *lát-ta-m* 'I saw' is used both with indefinite 3rd person object, no object and with the definite object. The reason for this, according to Olsson (1992) is that the indefinite form would otherwise be identical with the third person plural in past tense with indefinite object *lát-ta-k* 'they saw' because of the personal suffix *-k* with indefinite objects.

Selection of the personal suffix is governed by phonological, lexical and tense/mood factors. There are over twenty distinct suffixes of person because of the two conjugations and in order to avoid homonymy.

There are also words which are homonymous between the past tense and the present tense of different verbs, such as *vált* 'he became' and 'he changes' (Pajzs, 1996).

Furthermore, verbs in third person singular are often homonymous with nouns in nominative case, e.g. *vár* 'he waits vs. 'castle'. Sometimes even a conjugated form of a verb and an inflected form of a noun are homonymous, as in *várnak* 'they wait' vs. 'to the castle' (Pajzs, 1996).

Copula

In Hungarian, the copula [IGE] expresses that something exists. It can signal the existence of something, someone, place, and also signifies time, weather, a material, an origin, a cause, or a purpose. The use of the copula expresses even possession, e.g. 'to have something' where many other languages have a transitive verb for this type of construction.

The copula like other verbs is conjugated according to person/number, tense and mood. In the case of present indicative third person singular/plural the copula is realised as zero (Ø) if the predicate is nominal, e.g. pronoun, noun or adjective and the predicate expresses a profession, thing, state, quality or characteristic, etc. This phenomenon may cause problems in automatic tagging systems because there is no verb in the sentence and the elements may be considered as an example of a single NP.

If the predicate is an adverbial and signifies place, time, purpose, etc. or expresses possession the form in third person singular is *van* and in third person plural is *vannak*.

The negation of the singular form *van* is *nincs* 'is not' and of the plural form *vannak* is *nincsenek* 'are not'. All other forms of the copula are negated by a preposed *nem* 'no/not', e.g. *nem vagyok* 'I am not'.

1.2.6.1 Infinitive

Hungarian infinitives [INF], unlike most European languages, may be inflected for person. The reason for this is functional because often there is no other element in the sentence to mark the person. The suffixes are almost identical to the nominal paradigm, except in the third person both in singular and plural, where there is an *i* instead of the epenthetic *j* (see Appendix A).

1.2.6.2 Participles

The present, past and 'future' suffixes of the participles [MN] are -ó/-ö, -t/-tt, and -andó/-endö, respectively. The suffixes are added to the verb stem, and follow the rules of vowel harmony. The present participle form (-ó/ö) is quite productive and is often used as a noun (Campbell, 1991).

Example 16

a dolgoz-ó ember	a dolgozó-k
the work-PART man	the worker-PL
'the working man'	'the workers'

The main function of the past participle is to express an antecedent action and the states which result from it. It syntactically often behaves as adjective. Its form is identical with the third person singular past indefinite verb form (-*t*).

The participles are not used as predicates. Instead, there is a structure consisting of the copula and the verbal adverb with the suffix -*va/-ve*. The difference between participles and verbal adverbs is that the latter are more closely connected with the finite verb of the sentence than the participle is in time, state or mood. Participles modify the noun head while verbal adverbs modify the verb. The use of the verbal adverb is more limited than the use of the participle. The following examples show the difference between them.

Example 17

a   szobá-ban   ül-ö   gyerek-ek   játsza-nak
the room-INESS sit-PRSPART child-PL   play-3PL
'The children sitting in the room are playing.'

a   gyerek-ek a   szobá-ban   ül-ve   játsza-nak
the child-PL   the room-INESS sit-VERBADV play-3PL
'The children are playing sitting in the room.'

1.2.6.3 Verbal particles

Hungarian has a very rich system of verbal particles or verb prefixes [IK] which are separable from the stem (Campbell, 1991). They serve to mark direction (*le* 'down', *ki* 'out', etc.), aspect (*meg* 'completed') and to make verbs transitive. They can be combined with many verbs. In many cases there is a different meaning of the verb depending on what particle is attached to it, e.g. *átad* 'hand over, pass', *elad* 'sell', etc.

Verbal particles may have two positions depending on the emphasis within the sentence. In the case of a neutral sentence or in yes or no questions the particle is a prefix attached to the verb.

Example 18

> Péter *ki*-megy   a   szobá-ból ./?
> Peter out-go:3SG the room-ELAT
> 'Peter leaves the room. / Does Peter leave the room?'

The particle follows the verb in questions, in negatives and when any part of the sentence is emphasised.

Example 19

> Péter nem megy   *ki* a   szobá-ból ./?
> Peter not go:3SG out the room-ELAT
> 'Peter doesn't leave the room./Doesn't Peter leave the room?'

Note that in and the local relation is marked twice within each structure: once with verbal prefix and once with the case (elative) of the noun.

1.2.7 Postpositions

Postpositions [NU] follow the head they refer to and express principally local relations and show a three-way opposition for motion relative to the speaker or other referents. Postpositions may even represent temporal or abstract meaning, e.g. *után* 'after' or *ellen* 'against'. They are reduplicated with demonstratives (DEM) as shown the example below.

Example 20

> az   *alatt* az asztal *alatt*
> DEM under the   table   under
> 'under that table'

The postpositions, like case markers, may occur as stems and take possessive endings, (see Appendix A) and in this form according to traditional Hungarian grammar they are considered as adverbs. This can be problematic for an automatic tagging system because the stem of these type of adverbs is a postposition hence will be annotated as a postposition [NU], rather than an adverb [HA].

1.2.8 Adverbs

One type of adverb [HA] is the form of case_marker/postposition + personal ending. In some corpus, words consisting of case_marker + personal ending are considered as pronouns, while in another corpus they may be considered as adverbs, as was mentioned in section .

There are also adverbs which are derived from verbal particles expressing local relations though without possessive endings, such as *be* 'in' vs. *bent* 'inside', *ki* 'out' vs. *kint* 'outside'. The local relation is marked twice within a structure: the adverb and the case marker on the noun.

Example 21

> *Bent* van a szobá-*ban*.
> inside COP:3SG the room-INESS
> 'He/she/it is inside the room.'

1.2.9 Word formation

This section will give a brief overview on composition and derivation of words in Hungarian which can cause problems when automatically analysing texts. This part is based on Benkö & Imre (1972:145-156) where the interested reader may find more information with examples on this topic.

1.2.9.1 Word composition

Compound nouns are very frequent in Hungarian. Two nouns can simply be combined without any formal means, e.g. *kávé+ház* 'coffee'+'house'. Another type of compound nouns is where the first constituent is a participle, e.g. *mosó+nö* 'washing'+'woman'. The participle can also occur as a second member but the word is already substantivised and considered to be a noun, e.g. *adó+szedö* 'tax'+'collector'. These types of words may be confused with adjectives by an automatic tagging system because of their ending, typical to adjectives.

Compound adjectives consist of a noun + adjective. The compound may express similarities, e.g. *jég+hideg* 'ice'+'cold' or the adjective member limits a certain range of meaning, e.g. *adó+mentes* 'tax'+ 'free'.

There are also compounds which consist of i) a noun with adverbial ending or adverb + noun, ii) a noun with adverbial ending + non-finite verb form, iii) a noun with possessive suffix + participle, and iv) a noun with adverbial ending or adverb + verb etc. These may also cause problems in automatic morphological analysis.

1.2.9.2 Derivation of words

Unlike most European languages, Hungarian has a very regular system for derivational suffixes where a single suffix corresponds to a significant meaning and its use is regular. One suffix is tied to each important suffix function. Unfortunately, there is no space for describing the whole system with each derivational suffix, so only a few examples of those suffixes which change the category of a word are given.

Derivational suffixes which change the part of speech of a word are very common and productive. Verbs, nouns, adjectives and even adverbs can be further derived.

Deverbal verb suffixes, for example express frequentative-iterative, causative and reflexive meaning. Thus, the suffix *-hat/-het* means 'can, is capable' and 'may, is possible', e.g. *kime-het* 'he/she may go out'. Denominal verb suffixes are *-l*, *-z, -kodik/-kedik*, *-lkodik/-lkedik*, and *-skodik/-skedik*, e.g. *szolga* 'servent' => *szolgá-l* 'serve', *társ* 'fellow' => *társa-lkodik* 'converse'. Deadjectival verb suffixes are also very productive, such as the suffix *-kodik* which coincides with the denominal verb suffixes, e.g. *ügyes* 'skilful' => *ügyes-kedik* 'behave skilfully'.

Deverbal noun suffixes are also very common, especially the suffix *-ás/-és* as in *temet* 'bury' => *temet-és* 'burial, funeral', and the suffix *-at/-et* as in *talál* 'find, discover, hit' => *talál-at* 'hit, win'. The suffix *-ó/-ö* for the name of a profession which is identical with the participial belongs also to this category. The deadjectival noun suffix *-ság/-ség* is a very productive one, e.g. *alázatos* 'humble' => *alázatos-ság* 'humbleness'.

Adjectives may also be derived from verbs and nouns. Deverbal adjective suffixes among others are *-tlan/-tlen*, e.g. *árt* 'harm' => *árt-atlan* 'harmless', and *-ós/s*, e.g. *nyúl* 'stretch' => *nyúl-ós* 'stretchy'. Some examples for denominal adjective suffixes are *-s*, *-talan/-tlan*, *-i*, *-ú/-ü*, *-jú/-jü*. Note that several derivational suffixes may follow each other within one form, e.g. l*e-ír-hat-atlan* 'unwritable'.

1.3 Syntax

In the area of Hungarian syntax, the word order, main types of sentence structure, and agreement will be briefly described.

1.3.1 Word order

In most of the literature, Hungarian is described as a basically free word order language. But the fairly free word order can be said to be on the sentence level where the order of the major constituents is only free with respect to grammatical functions and their cases. The word order is pragmatically oriented with a special position for focused or emphasised constituents before the finite verb. The basic order of the sentence constituents is topic + focus + finite verb + any other items (Abondolo, 1992). Both topicalised and focused elements receive sentence stress. If a sentence contains both then the focused element is more prominent. shows some possible word orders for the sentence *Péter vesz egy könyvet a boltban* 'Peter buys a book at the store'[5]. The constituents of the sentence below (the subject, the verb, the direct object and the adverbial) can be ordered in 24 ways (4!).

Example 22

> Péter vesz egy könyvet a boltban.
> Péter vesz a boltban egy könyvet.
> Péter egy könyvet vesz a boltban.
> Péter egy könyvet a boltban vesz.
> Péter a boltban egy könyvet vesz.
> Péter a boltban vesz egy könyvet.
>   ?

However, the free word order description is inadequate on the phrase level and as soon as interrogative pronouns and/or negated complements are involved. Below, just a few examples are given.

There is a strict word order within the NP. Determiners and demonstrative pronouns always precede the noun head, e.g. in , the determiner *a* precedes the noun head *könyvet* 'book:ACC'. Qualifiers, like adjectives and participles within an NP occur after the determiners and demonstratives, but precede the noun, as in example below.

Example 23

> az-ok az elsö csodálatos karibiai nyári nap-ok
> that-PL the first wonderful Caribbean summer day-PL
> 'those first wonderful Caribbean summer days'

Furthermore, qualifiers may themselves be qualified by preceding adverbial complements.

Example 24

> a zöld ház-nál dohányz-ó fiatal nyelvész-ek
> the green house-ADESS smoke-PRSPART young linguist-PL
> 'the young linguists smoking at the green house'

Since Hungarian shares some typological characteristics of SOV languages (e.g. it is postpositional, the attribute precedes the noun, etc.) it is often described as a SOV language. On the other hand, some researchers maintain that Hungarian is partly SOV, partly SVO. However, these word orders are canonical and only represent the dominant order of simple declarative sentences, containing a nominal subject and a nominal object. In the next section, the main types of sentence structure with canonical word order are given. Note that the order of the constituents can be changed according to the information structure.

---

[5] Péter vesz egy könyvet a boltban
  Peter buys a book the store:INESS
  'Peter buys a book at the store'

1.3.2 The main types of sentence structure

The rules for the most basic syntactical forms of sentence structure are taken from Benkö & Imre, (1972:86-87) and given below. Grammatical categories which begin with a capital letter represent stem and affix and can be expanded by a subclass with the same name.

1. [Noun]$_{subject}$ + [Adjective + auxiliary_verb]$_{predicate}$ where the auxiliary may be zero.
2. [Noun]$_{subject}$ + [Noun + auxiliary_verb]$_{predicate}$ where the auxiliary may be zero.
3. [Noun]$_{subject}$ + [Copula]$_{predicate}$
4. [Noun]$_{subject}$ + [noun:case + Copula]$_{predicate}$
5. [noun:dative + Copula]$_{predicate}$ + [noun:possessive_suffix]$_{subject}$
6. [Noun]$_{subject}$ + [Verb]$_{predicate}$
7. [Noun]$_{subject}$ + [Verb + noun:case]$_{predicate}$
8. [Noun]$_{subject}$ + [Verb + noun:case + noun:case]$_{predicate}$
9. [Noun]$_{subject}$ + [Verb + Noun + postposition]$_{predicate}$

[Noun]

1.article + noun_stem:base_suffix:case_suffix
where article may be zero and base_suffix is either a single plural suffix or a possessive personal ending (including the special plural suffix).
2.article + noun_stem:base_suffix
where article may be zero and base_suffix is either a single plural suffix or a possessive personal ending (including the special plural suffix).
3.article + noun_stem:case_suffix
where article may be zero.
4.noun_stem
5.article + noun_stem:base_suffix + postposition
where article may be zero and base_suffix is either a plural suffix or a possessive personal ending (including the special plural suffix).
6.article + noun_stem + postposition
where article may be zero and base_suffix is either a plural suffix or a possessive personal ending (including the special plural suffix).

[Verb]

1.verb_stem:suffix
where suffix is personal ending, suffix of time and/or suffix of mood.
2.verb_stem
3.preverb + verb_stem:suffix
where suffix is personal ending, suffix of time and/or suffix of mood.
4.preverb + verb_stem
5.verb_stem:suffix + auxiliary_verb
where suffix is personal ending, suffix of time and/or suffix of mood.
6.verb_stem:suffix + auxiliary_verb
where suffix is personal ending, suffix of time and/or suffix of mood.

Additionally, Hungarian is a pro-drop language, which means that the subject position of the verb can be left empty. The subject is implicit as the personal endings of the verb express the first and the second person, and the third person if the context makes clear who or what the subject is.

1.3.3 Agreement

Generally, syntagmatic relations are marked, if possible, on both members of the construction which results in redundancy. The following rules illustrate the agreement patterns (Benkö & Imre, 1972).

1. The congruence of noun and substantival pronoun as genitive attribute and qualified noun.

Example 25

a	*te*	könyv-*ed*		a	fiú(*-nak a)*		könyv-*e*
DEM	you	book:2SGPOSS		the	boy(-DAT DET)		book-3SGPOSS
'your book'				'the boy's book'			

2. The demonstrative pronominal dependent *ez/az* 'this/that' agrees in case and in number with the noun which it qualifies.

Example 26

eb-böl		a könyv-böl		azo-k-at		a könyv-ek-et
DEM.PRON-ELAT		the book-ELAT		DEM.PRON-PL-ACC		the book-PL-ACC
'from this book'		'those books'				

3. Adjectival or numerical attributes are however, not in congruence with the modified word.

Example 27

nagy város-ok	öt város
big city-PL	five city:SG
'big cities'	'five cities'

4. As we have seen in the previous section, the verb agrees with the subject in number, person and even with the object, so called object-agreement - the relationship between the person of the subject and the person of the object is marked.

Example 28

Péter	ír-Ø		*egy*	könyv-et.		Péter	ír-*ja*		*a*	könyv-et.
Peter	write-3SGINDEF		a	book-ACC		Peter	write-3SGDEF		the	book-ACC
'Peter writes a book.'						'Peter is writing the book.'				

5. Agreement in number between subject and nominal predicate with zero copula.

Example 29

A	diák-o*k*	szorgalmas-a*k*.
the	student-PL	diligent-PL
'The students are diligent.'		

6. The personal endings of the infinitive refers to that part of the sentence which is in the dative and which is the logical subject.

Example 30

Péter-nek	kell	tanulni-a		(Én)	nek-em	kell	tanuln-om
Peter-DAT	must	study-INF3SG		(I)	me-1SG	must	study-INF1SG
'Peter must study.'				'I must study.'			

7. Verbal particles often agree with the case suffix of the nominal complement of the verb.

Example 31

Rá-megy-ek	a	fü-*re*.
PART-go-1SG	the	grass-ILL
'I step onto the grass.'		

References

Abondolo, D. 1987. Hungarian. In *The Major Languages of Eastern Europe.* Comrie, B. et al. Routledge, London.

Abondolo, D. 1992. Hungarian. In *International Encyclopaedia of Linguistics*. Volume 2. Bright, W. Oxford University Press.

Bencédy, J., et al. 1968. *A mai magyar nyelv*. Nemzeti tankönyvkiadó, Budapest.

Benkö, L & Imre, S. 1972. *The Hungarian Language.* Akadémiai Kiadó, Budapest.

Campbell, G. L. 1991. *Compendium of the Worlds Languages*. Volume 1. Routledge, London and New York.

Kiss, É. K. 1987. *Configurationality in Hungarian*. Akadémiai Kiadó. Budapest.

Pajzs, J. 1996. *Disambiguation of suffixal structure of Hungarian words using information about part of speech and suffixal structure of words in the context*. COPERNICUS Project 621 GRAMLEX, Work package 3 - Task 3E2. Research Institute for Linguistics, Hungarian Academy of Sciences.

Tompa, J., et al. 1970. *A mai magyar nyelv rendszere*. Akadémiai Kiadó, Budapest.