

1

Introduction

1.1 Genealogy, areal distribution

Hungarian is a Finno-Ugric language. The Finno-Ugric languages and the practically extinct Samoyed languages of Siberia constitute the Uralic language family. Within the Finno-Ugric family, Hungarian belongs to the Ugric branch, together with Mansy, or Vogul, and Khanty, or Ostyak, spoken by a few thousand people in western Siberia. The family also has:

- a Finnic branch, including Finnish (5 million speakers) and Estonian (1 million speakers);
- a Sami or Lappish branch (35 000 speakers); as well as
- a Mordvin branch, consisting of Erzya (500 000 speakers) and Moksha (250 000 speakers);
- a Mari or Cheremis branch (550 000 speakers); and
- a Permi branch, consisting of Udmurt or Votyak (500 000 speakers) and Komi or Zuryen (350 000 speakers).

Hungarian, Finnish and Estonian are state languages; Sami is spoken in northern Norway, Sweden and Finland, whereas the Mordvin, Mari, and Permi languages are spoken in the European territories of Russia.

Hungarian is spoken in Central Europe. It is the state language of Hungary, but the area where it is a native language also extends to the neighboring countries. In Hungary it has 10 million speakers, in Romania 2 million speakers, in Slovakia 700 000 speakers, in Yugoslavia 300 000 speakers, in Ukraine 150 000 speakers. There is also a Hungarian minority in Croatia, Slovenia, and Austria, and a considerable diaspora in Western Europe, North America, South America, Israel, and Australia.

The period in which the Finno-Ugric peoples represented a kind of linguistic and areal unity is believed to have lasted until 2000 BC. On the basis of linguistic evidence – e.g. the habitat of the plants and animals whose names are shared by the Finno-Ugric languages – the Finno-Ugric homeland is located on the south-western slopes of the Ural mountains. The Hungarian language must have

2 The syntax of Hungarian

emerged here from among the Ugric dialects after 1000 BC. The Hungarian tribes left the Finno-Ugric homeland in the fifth century AD, and occupied the territory surrounded by the Carpathian mountains in 895. In the period of migration, the language had been subject to heavy Turkic influence. After the Hungarian tribes settled in Central Europe, the Slavic languages and German had a noticeable impact on Hungarian.

The first written Hungarian records are Hungarian fragments in a Greek and a Latin text, dating from 950 and 1055, respectively. The first two surviving coherent written Hungarian texts originated in 1192–95, and in 1300. Interestingly, they are still to a large extent comprehensible to the present-day reader. Hungarian is also fairly homogeneous areally; the only dialect displaying substantial lexical, phonological, and syntactic differences from standard Hungarian is the easternmost, archaic Csángó dialect spoken in Romania.

1.2 A general overview of the syntactic and morphosyntactic features of Hungarian

Hungarian is often referred to as a free word-order language, because the grammatical functions of subject, object, etc. are not linked to invariant structural positions in the sentence. Thus, a transitive verb and its two arguments, e.g. *keresi* ‘seeks’ *János* ‘John’ *Marit* ‘Mary-ACC’ can form a sentence in any of the theoretically possible SVO, SOV, OVS, OSV, VSO, and VOS combinations:

- | | | |
|-----|----------------------------|----------------------------|
| (1) | <i>János keresi Marit.</i> | <i>Marit János keresi.</i> |
| | <i>János Marit keresi.</i> | <i>Keresi János Marit.</i> |
| | <i>Marit keresi János.</i> | <i>Keresi Marit János.</i> |

A closer scrutiny, however, makes it clear that the order of major sentence constituents is just as strictly constrained in Hungarian as it is, for example, in English or French – merely the functions associated with the different structural positions are logical functions instead of the grammatical functions subject, object, etc.

The Hungarian sentence can be divided primarily into a topic part and a predicate part. The topic, functioning as the logical subject of predication, names the individual that will be predicated about in the sentence. The topic role is independent of the function ‘grammatical subject’; in other words, an action or state can be predicated about any of its participants. Thus, in (2a) the agent, or grammatical subject, and in (2b) the theme, or grammatical object, occupies the position associated with the topic function. (Since in English the topic and the grammatical

subject roles have to coincide, (2b) is translated by a passive sentence. Hungarian (2b), however, differs from (2a) only in its word order.)

- (2) a. [Top János] [Pred fel hívta Marit]¹
 John up called Mary-ACC²
 ‘John called up Mary.’
- b. [Top Marit] [Pred fel hívta János]
 Mary-ACC up called John-NOM
 ‘Mary was called up by John.’

As for the predicate of the sentence, its central element is the verb (V). The postverbal positions are argument positions. The verb usually has a so-called verb modifier (VM), i.e., a non-referential complement incorporated into it, acting as an aspectual operator, as in (3):

- (3) [Top János] [Pred **fel** ásta a kertet]
 John up dug the garden-ACC
 ‘John dug up the garden.’

The preverbal section of the predicate phrase contains operator positions. The verb may be immediately preceded by a focus constituent, expressing exhaustive identification (N.B. small capitals are used in the Hungarian text here and throughout the book to mark focus, indicating its phonological prominence):

- (4) a. [Top János] [Pred MARIT kérte fel]
 John Mary-ACC asked VM
 ‘As for John, it was Mary that he asked for a dance.’
- b. [Top Marit] [Pred JÁNOS kérte fel]
 Mary-ACC John-NOM asked VM
 ‘As for Mary, it was John who asked her for a dance.’

The focus position is preceded by a distributive quantifier position, the locus of universal quantifiers and *is* ‘also’ phrases; for example:

¹ When the verbal prefix precedes the verb, they are traditionally spelled as one word. Nevertheless, I often spell them as two separate words in this book, in order to express the fact that they represent two separate syntactic constituents.

² Hungarian is an agglutinating language, in which morphosyntactic elements are right-adjoined to the lexical root. In cases when the morphosyntactic elements play no role in the discussion, the order of the stem and the suffix may be reversed in the glosses in order to facilitate understanding; i.e., example (i) may be glossed as (iii), instead of the more precise (ii).

- (i) Siettek iskolába.
 (ii) hurried-they school-to
 (iii) they.hurried to.school

4 The syntax of Hungarian

- (5) a. [Pred **Marit is** JÁNOS kérte fel]
 Mary-ACC too John-NOM asked VM
 ‘(In the case of) Mary, too, it was John who asked her for a dance.’
- b. [Pred **Mindenki** MARIT kérte fel]
 everybody Mary-ACC asked VM
 ‘(For) everybody, it was Mary that he asked for a dance.’

As is clear from the examples, the preverbal operator positions, too, are filled with no regard to the grammatical function of the filler.

The preverbal operator field of the predicate phrase can also contain negation in addition to the identificational focus and the distributive quantifiers. The negative particle occupies either the immediately preverbal position or the immediately prefocus position, or both simultaneously. The lower negative particle negates the VP – i.e., essentially the propositional content of the clause – whereas the higher negative particle negates the identification expressed by the focus.

- (6) a. János nem hívta fel Marit.
 John not called up Mary-ACC
 ‘John did not call up Mary.’
- b. János nem MARIT hívta fel.
 ‘As for John, it was not Mary that he called up.’
- c. János nem MARIT nem hívta fel.
 ‘As for John, it was not Mary that he did not call up.’

The negative particle triggers negative concord among universal quantifiers and indefinites, i.e., its universally quantified clause-mates, and the indefinites in its scope have a special negative form:

- (7) János senkinek nem mondott semmit.³
 John nobody-DAT not said nothing-ACC
 ‘John did not say anything to anybody.’

Examples (6a) and (6b), displaying preverbal negation and prefocus negation respectively, also call attention to a very characteristic property of Hungarian: operators precede and c-command their scope, i.e., Hungarian sentences are disambiguated scopally.

In sum: the Hungarian sentence structure to be argued for in this book is a hierarchical structure falling into a topic part and a predicate part, with the predicate part containing a V-initial propositional kernel as well as preverbal operators.⁴

³ For arguments that the negative particle + focus string in (6b,c) is not an instance of constituent negation, and for details of negative concord, see Chapter 6.

⁴ For the first formulations of this theory, see Brassai (1860, 1863–65), and É. Kiss (1977).

Chapters 2–6 of this book are devoted to establishing the precise configuration of the structural positions illustrated above, analyzing the operations filling them: examining their trigger, the constraints they are subject to, the semantic consequences they bring about, etc. Chapter 2 discusses the topic position, the topic function, and the operation of topicalization. Chapter 3, dealing with the core of the predicate phrase, examines, on the one hand, the properties, relative order, and hierarchical relations of postverbal arguments and, on the other hand, the properties of the verb modifier, acting as an aspectual operator. Chapter 4 describes the focus position, and the syntax and semantics of focussing/exhaustive identification. Chapter 5 is devoted to questions of quantification: the position of the distributive quantifier, the operation of quantifier raising, scope interpretation, and the like. Chapter 6 discusses negation, including negative concord. Operator movement across clause boundaries is examined in Chapter 10, dealing with subordination.

Hungarian lacks well-known manifestations of the structural prominence of the subject over the object. For instance, in English a subject pronoun and the genitive specifier of the object display disjoint reference; whereas an object pronoun and the genitive specifier of the subject can also corefer, as follows from Binding Condition C applied to an asymmetrical SVO (subject–verb–object) structure. In Hungarian, we find disjoint reference in both cases; compare:

- (8) a. * He_i loves John_i's mother.
 b. John_i's mother loves him_i.
- (9) a. * Ő_i szereti János_i anyját.
 he_i loves John_i's mother
 b. * János_i anyja szereti őt_i.
 John_i's mother loves him_i

The question whether the Hungarian VP is flat, with the subject and the object being sisters, or is configurational, with the subject asymmetrically c-commanding the object, was in the focus of interest in the 1980s and inspired a large amount of literature.⁵ The arguments for and against a flat VP are summarized in Chapter 3. The question will also be addressed as to whether the assumption of a flat VP is compatible with current assumptions about the possible format of syntactic structures.

The apparent freedom of Hungarian word order – i.e., the attested parallelism between the syntactic behavior of the subject, object, and other arguments – and their equal eligibility for operator movement must be related to the fact that Hungarian

⁵ See, among others, É. Kiss (1981, 1987a, 1987b), Horvath (1986a, 1986b, 1987), Marác (1986b, 1989).

6 The syntax of Hungarian

morphology is very rich. Hungarian is a nominative–accusative language with 18 cases, all of which appear to be lexically selected. (Hungarian has no grammatical-function-changing transformations such as passivization.) There is no evidence of the assignment/licensing of particular Cases being linked to particular sentence positions. Accordingly, the lack of a thematic subject does not give rise to an expletive.

Hungarian displays agreement in several areas of grammar. The verb agrees not only with the subject, but also with the object if it is definite. In the possessive construction the head noun bears an agreement suffix reflecting the person and number of the possessor. In postpositional constructions the postposition agrees in person and number with its pronominal complement. Hungarian also has a type of inflected infinitive, agreeing with its dative-marked subject. The phrase types displaying agreement all license pro-drop. Verb morphology is discussed briefly in Chapter 3, dealing with the VP. Nominal, postpositional, and infinitival inflection are analyzed in Chapters 7, 8, and 9, respectively. Inflectional morphology is treated as part of syntax.

The chapters analyzing the inner structure of the noun phrase (NP) and the postpositional phrase (PP) will reveal a great degree of parallelism between the extended VP, and the extended noun phrase/PP. Greenbergian typologies (e.g. Greenberg 1966) categorize Hungarian as a head-final language, in part because the structures of the noun phrase, the attributive adjective phrase, and the postpositional phrase are head final on the surface. However, the VP and the CP are clearly head initial, and the predicative adjective phrase need not be head final, either. This book derives all phrase types from a head-initial base, by subjecting the post-head complement to extraposition, incorporation, or phrase-internal topicalization. For example, in the case of the noun phrase, a post-head complement must be removed because it would block the merging of the head noun with case. Hungarian postpositions are morphosyntactic suffixes, and, as such, they must be right-adjoined to their complement; that is why we attest a DP P order instead of the underlying P DP.

Although Hungarian differs from the best-known Indo-European languages, particularly English, in obvious ways, its basic syntactic properties are not at all unique. Sámuel Brassai, the first linguist to identify the outlines of Hungarian sentence structure, realized already in the middle of the nineteenth century that the topic–predicate articulation relates to the subject–predicate structure attested in some Indo-European languages as the general relates to the specific. That is, the subject–predicate articulation is a topic–predicate structure with the target of topicalization restricted to the grammatical subject (compare Brassai 1860, 1863–65). The generative framework also provides a clue as to the reason for this restriction: in subject–predicate languages the subject can assume nominative case

only if it is raised to Spec,IP. In Spec,IP it is closer to the topic position than the other arguments, hence it is topicalizable more economically.⁶

The immediately preverbal identificational focus of Hungarian is also a property shared by a great many languages. For example, in an examination of 35 European languages, 27 languages were found to have an invariant structural position associated with the function of identificational focus, and in 18 of these languages the focus position was found to be immediately preverbal.⁷

The distributive quantifier position at the head of the predicate phrase may not be a feature unique to Hungarian, either. Observations concerning Arabic (compare Khalaily 1995), KiLega (Kinyalolo 1990), Chinese (Bartos 2000b), etc. suggest that universal quantifiers – irrespective of their grammatical function – gravitate to a fixed position in other languages, as well.

An interesting open question is whether there are also other languages besides Hungarian in which the subject and the object behave in a parallel way in so many areas of syntax. The possibility of a head bearing a sister relation to all of its complements is not an option directly predicted by current theoretical assumptions. According to the Linear Correspondence Axiom of Kayne (1994) the hierarchical asymmetry of structural relations ensures their unambiguous mapping onto linear order. Kayne claims that the symmetry of a structural relation, i.e., the impossibility of its linearization, leads to the crashing of the derivation. Perhaps under appropriate conditions, however, the lack of asymmetry might result in a partially linearized structure of the type represented by the Hungarian VP, in which the head – asymmetrically c-commanding its complements – occupies initial position; the post-head arguments, mutually c-commanding each other, on the other hand, can stand in any order. There may also be other ways of reconciling a flat VP with Kayne's antisymmetry theory; for example one might argue that the flat VP represents an intermediate stage of the derivation, with the traces of the initial, asymmetric stage deleted.

The description of Hungarian syntax presented in this book adopts the basic theoretical assumptions and the basic methodology of generative linguistics. However, the approach is empirical rather than technical; the goal of the analyses is to present the theoretically relevant facts of Hungarian explicitly, but without necessarily providing accounts in terms of the most recent theoretical innovations.

⁶ The claim that the subject in a subject–predicate language is, in fact, moved to topic position can be checked by examining if subjects unsuitable for the topic role – e.g. non-specific indefinites – occupy the same position that topicalizable subjects occupy. A large amount of evidence indicates that non-specific subjects stay in a predicate-internal subject position in subject–predicate languages, as well (compare É. Kiss 1996, 1998c, Diesing 1992).

⁷ The details of the investigation, performed in the framework of the EURO TYP project of the European Science Foundation, appeared in É. Kiss (1998c).

2

The topic–predicate articulation of the sentence

2.1 The topic function

Although in the Hungarian sentence the complements of the verb, including the subject, can appear in various permutations (1), the syntactic structure of the sentence is constant, with its units expressing invariant logical-semantic functions.

- (1) a. A védők sokáig tartották a várat a törökök ellen.
 the defenders long held the fort-ACC the Turks against
 ‘The defenders held the fort against the Turks for a long time.’
- b. A védők sokáig tartották a törökök ellen a várat.
 c. A várat sokáig tartották a védők a törökök ellen.
 d. A várat sokáig tartották a törökök ellen a védők.
 e. A törökök ellen sokáig tartották a védők a várat.
 f. A törökök ellen sokáig tartották a várat a védők.

The sentences in (1a–f) all express predication about the referent of their initial constituent. Although they describe the same event (the event of the defenders holding the fort against the Turks for a long time), they formulate it as statements about different participants of the event. (1a,b) make a statement about the defenders (that they held the fort against the Turks for a long time), (1c,d) make a statement about the fort (that the defenders held it against the Turks for a long time), whereas (1e,f) make a statement about the Turks (that the defenders held the fort against them for a long time). That is, the sentences all instantiate a similar predication relation, with the initial constituent functioning as the logical subject of predication, and the rest of the sentence functioning as the logical predicate. In order to avoid confusion of the notions of logical subject of predication, and grammatical subject, the logical subject of predication is referred to here by the term ‘topic’. So, the Hungarian sentence primarily divides into a topic and a predicate, with the topic expressing the following function:

(2) *The topic function*

The topic foregrounds an individual (a person, an object, or a group of them) from among those present in the universe of discourse as the subject of the subsequent predication.

2.2 The formal features of the topic constituent

As is clear from the examples in (1a–e), the topic role is not linked to a particular grammatical function. Subjects are in fact more frequent topics than objects – but the link between subjecthood and topichood is only indirect. We tend to describe events from a human perspective, as statements about their human participants – and subjects are more often [+human] than objects are. In the case of verbs with a [–human] subject and a [+human] accusative or oblique complement, the most common permutation is that in which the accusative or oblique complement occupies the topic position (3a,b). When the possessor is the only human involved in an action or state, the possessor is usually topicalized (3c).

- (3) a. [Topic Jánost] [Predicate elütötte egy autót]
 John-ACC hit a car
 ‘A car hit John. [John was hit by a car.]’
- b. [Topic Jánosból] [Predicate hiányzik a becsület]
 from.John is.missing the honesty
 ‘Honesty is missing from John. [John lacks honesty.]’
- c. [Topic Jánosnak] [Predicate összetörték az autóját]
 John-DAT they.broke the car-POSS-ACC
 ‘They broke John’s car. [John had his car broken.]’

The topic, naming a participant of the action or state to be described in the sentence, represents an argument of the verb, i.e., it binds an argument position in the predicate part. (This formal criterion helps us to distinguish topics from sentence adverbials, which are predicate-external constituents not functioning as logical subjects of predication.) Since the topic denotes an individual (a person, a thing, or a group of persons or things), it is represented by a referring expression.¹ Thus,

¹ The notion of referentiality must be somewhat relaxed, though, for this statement to be tenable. Thus generics can also be topicalized – see (i) and (ii) – but they can be said to refer to kinds.

- (i) A nők szeretik megosztani a problémáikat.
 the women like to.share their problems
 ‘Women like to share their problems.’
- (ii) Egy nő szereti megosztani a problémáit.
 a woman likes to.share her problems

(continued overleaf)

10 The syntax of Hungarian

arguments with an operator feature, e.g. quantifiers, cannot be topicalized:

- (4) a. * [_{Topic} Kik] [_{Predicate} meg védtek a várat a törökök ellen]?
 who-PL VM defended the fort-ACC the Turks against
 ‘Who defended the fort against the Turks?’
- b. * [_{Topic} Kevés várat] [_{Predicate} meg védtek a zsoldosok a törökök ellen]
 few fort-ACC VM defended the mercenaries the Turks against
 ‘Few forts were defended against the Turks by the mercenaries.’

The topic picks an individual as the subject of the subsequent predication from among those already present in the universe of discourse; therefore, it must be represented by a constituent associated with an existential presupposition. Thus, a proper name or a definite noun phrase (or a postpositional phrase containing a proper name or a definite noun phrase) is a possible topic; an indefinite noun phrase, however, can only be topicalized if it is specific in the sense of Enç (1991), i.e., if its referent is the subset of a referent already present in the universe of discourse. From a different perspective: a topicalized indefinite noun phrase always requires a specific (i.e., partitive) interpretation. Consequently, the topicalization of a predicate-internal indefinite noun phrase brings about an interpretational difference:

- (5) a. [_{Predicate} Meg állt egy autó a házunk előtt]
 VM stopped a car our house before
 ‘A car has stopped in front of our house.’
- b. [_{Predicate} Egy autó állt meg a házunk előtt]
 ‘A car has stopped in front of our house.’
- c. [_{Topic} Egy autó] [_{Predicate} meg állt a házunk előtt]
 ‘One of the cars has stopped in front of our house.’

In (5a,b), the predicate-internal *egy autó* ‘a car’ is likely to be non-specific, serving to introduce a car into the domain of discourse. The topicalized *egy autó* in (5c), on the other hand, can only be understood specifically, meaning ‘one of the aforementioned cars’. The [+specific] feature required of topicalized indefinite noun phrases is in fact also shared by definite noun phrases. (Since a definite noun phrase is identical with a previously introduced referent, and since the identity relation is a sub-type of the subset relation, its referent always represents a subset of a

As is shown below, existential quantifiers of the *valaki, valami* ‘somebody, something’ type can also appear in topic position; however, they can be said to behave like indefinites, which function either as referential expressions or as bound variables. *Valaki*-type expressions are topicalizable only when used referentially. Apparently universal quantifiers can also be topicalized if they are contrasted. However, as will be argued in Section 2.7, a contrasted quantifier functions as the name of a set property.