Grammatical encoding of referentiality in the history of Hungarian

Barbara Egedi
Hungarian Academy of Sciences

The paper demonstrates that the Old Hungarian article, although homophonous with the distal demonstrative it developed from, is a fully grammaticalized element encoding definiteness on a syntactic level. Questioning its status as a true article is the consequence of the fact that it had a more restricted use in the first records than it does in the subsequent language stages. The careful classification of the articleless noun phrases with a definite interpretation in a selected closed corpus reveals that the Old Hungarian article only appears when there is no other device to identify the referent of the nominal expression, thus it is absent with proper names, generics, nouns modified by a demonstrative, head nouns of possessive constructions, and, occasionally, with non-argument adjuncts. Spreading of the article in certain cntexts can be observed only at the end the period and in Middle Hungarian, in line with the gradual functional extension of the nominal left periphery.

1. Introduction

This paper forms part of an ongoing research project whose aim is to investigate the history of Hungarian, as well as language change in general, and to analyze the empirical results in a modern theoretical framework. Traditional historical grammars limited themselves so far to descriptive statements without intending to explain certain syntactic phenomena or propose any analysis with an explanatory force. We are interested in reconstructing the syntax of different synchronic systems as well as examining and modeling the grammaticalization processes that can be attested through the early history of this language. Studies of this type do not seem to have been made for a Finno-Ugric language before. Admittedly, these languages are poorly documented, and Hungarian is exceptional with its approximately 800 years of documentation.

Furthermore, the project aims to digitize all the Old Hungarian records and some selected texts from the Middle Hungarian period, and to build an on-line searchable historical language corpus. A considerable part of the texts will also be morpho-syntactically annotated.[1] Although a high amount of texts, more than one million tokens, has already been digitized with their original spelling, only a small part of the corpus is annotated at the moment, thus a syntactic research based on our database is still very limited. Accordingly, the present paper will put forth my first observations and research results concerning the use of the Old Hungarian article and the strategies of the referential identification, based on a traditional philological work, with no automatic query involved. Since the time of completing this manuscript, the Corpus considerably developed and the hypothesis presented here has been successfully checked against a larger amount of texts, namely in five normalized codices from the Old Hungarian Corpus.[2]

The historical languistic stages of Hungarian can be observed in Table 1, with some additional notes on the sources we have from the different periods.

Table 1. Language stages of Hungarian

Proto-Hungarian	1000 BC – 896 AD	No written documents, only reconstructed grammar
Old Hungarian	896–1526	Manuscripts, mainly codices
Middle Hungarian	1526–1772	Book printing New secular genres appear Significant increase in quantity
Modern Hungarian	1772–present day	

The first completely Hungarian *printed* book (*The Letters of Saint Paul*) appeared only in 1533, the traditional division of the stages is linked up with historical events.

1. The *Hungarian Generative Diachronic Syntax* Project is supported by the Hungarian Scientific Research Fund (OTKA No. 78074) and is located in the Research Institute for Linguistics of the Hungarian Academy of Sciences. I am also very grateful to my anonymous reviewers for the questions they raised and the corrections they suggested.

2. The results of this research have been presented on the 14th June 2012 at the conference 'Exploring Ancient Languages through Corpora' in Oslo. I complemented the earlier version of the present manuscript with some of the data used in the Oslo talk.

Table 2. Types of sources from the Old Hungarian Period

Early Old Hungarian	The age of the Árpád dynasty	Sporadic records, glosses, a few short texts, e.g. *Funeral Sermon and Prayer* (ca. 1195)
Late Old Hungarian	From around 1370 (time of compilation of Codex Jókai)	Codices containing translations of Latin religious literature + original Hungarian compositions (documents, poems and letters)

Table 2 shows a further subdivision of Old Hungarian, which is motivated by the purpose of the present study as well. The linguistic records of the Early Old Hungarian Period, which are the first written sources at the same time, are sporadic records, mostly names of people and places and other glosses embedded in Greek or Latin documents and charters, dating from the 10th century onwards. The very first charter, which survived in its original format and contains more than 50 Hungarian words and word-groups is the *Letter of Foundation of Tihany* (1055). These sources, however, provide us with no information with respect to syntactic structures of this early language stage. The first continuous texts from the same period are actually very short and few in number: we have four records dating from between the end of the 12th c. to the end of the 13th c. The first texts of considerable length and thus suitable for syntactic investigation come from the first half of the Late Old Hungarian Period (see 3.2 below).

2. The first attestations of the definite article: Is it an article at all?

Modern Hungarian makes extensive use of the definite article (Examples 1–3). The article occurs together with demonstratives, possessives, and optionally even with proper names.

Modern Hungarian

(1) *ez a könyv*
this *the* book
'this book'

(2) *az én könyv-em*
the I book-POSS.1SG
'my book'

(3) *a szerzetesek-nek a könyv-e*
the monks-DAT *the* book-POSS
'the book of the monks'

The corresponding Old Hungarian examples look rather differently with respect to the article use: in most of the above contexts definite articles are absent, as illustrated in the following examples:[3]

Old Hungarian

(4) *ez könyv*
this book
'this book'

(5) *én könyv-em*
I book-POSS.1SG
'my book'

(6) *a szerzetesek-nek könyv-e*
the monks-DAT book-POSS
'the book of the monks'

Not only does the distribution differ, but the first attestations of the article are by no means obvious either. As in so many other languages, in Hungarian the definite article developed from a demonstrative modifier, but in this early period they cannot be distinguished from each other merely on formal grounds – at least for today's reader. There is one morpheme that apparently corresponds to two different functions: sometimes it seems to behave as a demonstrative, sometimes as a definite article, and often it is just impossible to decide between the two options. Since the functions of the demonstratives and the definite article may also overlap in several cases, there is a strong tendency in the literature on Old Hungarian not to consider these early articles as true articles, but rather as 'pre-articles' that belong to a kind of transitional word-class with dual nature, namely they are 'pronoun-articles' (Bakró-Nagy 1999: 7; I. Galassy 1992: 721–722). The same problem arises in other languages too, as long as the emergence of the article falls in their written history and thus can be well observed in the records. (For an extensive summary of the problem for Old English with an exhaustive list of references, see Sommerer 2011).

3. Note that the dative marked elements in these examples are always the possessor expressions, and the gloss 'POSS' stands for a suffix on the head noun that encodes its being possessed. The possessor expression may have either a dative or a nominative case as it will be seen at a later point of the discussion. The article in (6) belongs to the possessor only rather than to the construction as a whole (cf. 4.4.2). At last, it must be noted that spelling out the personal pronoun (e.g. 'I' in (2)) has a contrastive or emphatic effect in Modern Hungarian, but seems to be neutral, or even preferred in the codices (5).

In many Old Hungarian instances the modern reader remains uncertain how to interpret a phrase such as (7) due to the formal and positional equivalence and functional overlap (e.g. anaphoric use) between the article and the demonstrative modifier, which the article developed from.

(7) a. *az kapu*
the gate
b. *az kapu*
that gate

Nevertheless, it does not mean that the category itself is transitional or that the Old Hungarian speakers could not make a distinction between a deictic and a purely referential use of determiners. There might have been a difference in the intonational properties of the two phrases, but as stress is not marked in writing, this assumption is impossible to test.

The lack of consensus about the existence of articles in systems like Old Hungarian or Old English usually derives from the fact that opinions differ on what makes a determiner feature as an article and which are the more applicable criteria to distinguish between a demonstrative and an article.

Lotte Sommerer (2011: 183–209), employs seven criteria in her dissertation for establishing 'articlehood', but she herself admits that many of these criteria fail to apply to Old English, or the results turned out to be inconclusive. The source of her problem is similar to ours: in this language stage the one and the same form is used to express various functions. Paola Crisma (2011) attempts to solve the problem by introducing two completely formal requirements. These refer to the overt marking of the [+ definite] feature and to the relation between the category D and argumenthood. If these requirements are met, definiteness grammaticalized and the article can be identified as a compulsory element of the given grammar.

This paper will argue that the definite article, i.e. a fully grammaticalized category encoding definiteness on a syntactic level, already existed in Old Hungarian.[4] This early article had a more restricted use, however, than it does in the subsequent language phases, as it will be demonstrated in the following sections. To identify this early determiner as a true article, formal and distributional criteria (similar to those suggested by the authors cited above) will not be sufficient. One should rather appeal to semantics and find out whether the determiner in question appears in contexts which demonstratives are banned from. As Nikolaus Himmelmann (2001: 833–834) sums it up, demonstratives must not be used in certain semantic and pragmatic contexts in

4. The paper does not aim to discuss the causes of the change, i.e. why the article developed from the demonstrative and what factors conditioned its emergence.

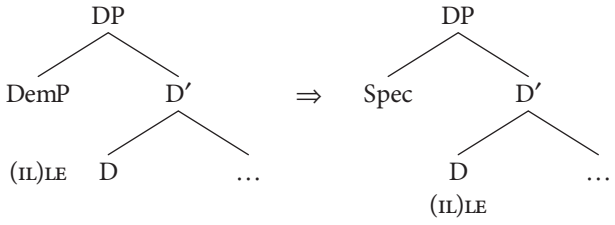
which articles consistently appear. Such contexts are *the larger situation use* ("the first mention of entities that are considered to be unique, hence generally identifiable, in a given speech community") and *the associative-anaphoric use* ("the first mention of an entity that is not unique per se but with respect to a previously mentioned referent").[5] Observing the Old Hungarian data, these uses turned out to be well attested. (8) stands here to exemplify the associative-anaphoric use of the determiner *az* 'the', from the earliest longer narrative text we have.[6]

(8) *Tertenek hogÿ nemÿ ygen zepp hews (…) ÿewue az kappu-ra*
happened that a certain very nice hero came *the gate-SUB*
es czergete [Jókai Codex 14:12–15]
and knocked
'It happened that a very nice hero came to the gate and knocked'

In the context, there was no mention of any gate or entrance actually, but in the preceding lines the reader has been informed about Saint Francis and his fellows' arrival in Spoletano and their dwelling in a certain place.

Despite the fact that the distal demonstrative and the article are homophonous elements (cf. (7)), I assume that they occupy distinct structural positions. The situation may be similar to what Giuliana Giusti proposes (2001: 167) for the structural reanalysis of Latin *ille* in Romance languages from a demonstrative into the definite article:

(9) a. b.



5. For a more detailed pragmatic account of the use of the definite article, see Hawkins (1991), who claims that existence and uniqueness of a definite description must hold within the universe of discourse or a subset of it, which can be inferred by the hearer through specific pragmatic parameters. For similar approaches, see Westerståhl (1985) context sets; Roberts (2003) informational uniqueness relative to the discourse situation.

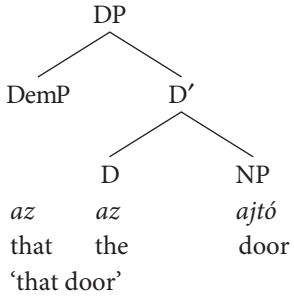
6. Glosses in this paper generally follow the Leipzig Glossing Rules, but details of verbal morphology, which is rather complex in Hungarian, are omitted. Therefore I simplified the glossing by giving the details only within the noun phrases that are the only relevant here.

What (9) illustrates is that the Latin element *ille,* originally located in the specifier of the Determiner Phrase, loses its first syllable and is reinterpreted as an element in D. The two constructions, of course, may also have coexisted in the language for several generations.

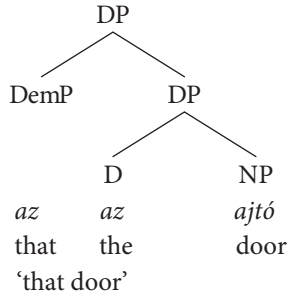
A similar process may be assumed for how the Old Hungarian definite article developed. This kind of reanalysis perfectly fits to one of the universal economical principles of Minimalist syntax, which is considered to be responsible for many language change phenomena. According to the so called 'Head Preference Principle' (van Gelderen 2008), speakers prefer to build structures where an element is merged directly into the head position instead of moving it to the specifier from below.

After the reanalysis illustrated in (9) had happened in Romance, new lexical items developed to fulfill the demonstrative function: an adverbial reinforcer has been added to the phonologically weakened demonstrative head resulting in a new series of deictic elements. In Hungarian, the superficially homophonous structures split in a different way. During the 16th century a new pattern arises for the noun phrases containing a demonstrative. In this structure a sort of *determiner doubling* can be observed, which may be analyzed as a double filling of the functional projection for definiteness. The demonstrative is spelled out in the specifier position, while the determiner in article function is the head of the DP (10a). For independent reasons, not to be explicated here, I believe that an adjunction analysis would be more plausible (10b), where the newly added demonstrative is adjoined more loosely to the noun phrase.

(10) a.



b.



This new pattern, however, is characteristic for the Middle Hungarian Period and later stages, thus not relevant at this point in our investigation.

To summarize this section, the uncertainty in the literature concerning the Old Hungarian definite article derives, on the one hand, from the fact that the written sources show two coexisting and homophonous structures that are sometimes difficult to distinguish functionally until the pattern in (10) appears. On the other

hand, what might mislead some of the researchers is that the Old Hungarian system of determination differed significantly from the present one, with a more restricted use of the definite article with respect to subsequent language phases. This investigation will concentrate on early closed corpora to demonstrate that the use of the article is logical and predictable at this synchronic stage of the language as well. Moreover, studying various texts that follow each other in a diachronic order, also the way of gradual spreading in article use can be detected.

3. Definite descriptions in the first half of the Late Old Hungarian Period

3.1 What is definiteness?

The semantic and pragmatic notion of definiteness seems to be universal, only its grammatical realization is a language specific property. For its definition one can appeal to various semantic and logical-philosophical approaches, but admitting the relevance of concepts such as uniqueness, inclusiveness, and familiarity, this paper relies on the basic pragmatic notion of *identifiability* as the crucial property of definiteness. According to this property, the speaker signals that the hearer is able to assign a referent to a certain noun phrase, that is, the hearer can identify the referent of the DP.[7] The semantic-pragmatic category of *identifiability* is prototypically realized in all languages, but not necessarily by a definite article; many languages lack this type of grammatical element. The definite article is actually the grammaticalization proper of the semantic and pragmatic concept of definiteness. The grammatical encoding of this property may segment the semantic field at different points in different languages, i.e. there is a great variation as to how extensive the ground the category of definiteness covers is within the semantic/pragmatic field in a language. Accordingly, this semantic range may even change in time (Lyons 1999: 336–337). In a language documented long enough for a diachronic investigation, one can observe how the grammatical encoding of referentiality changes gradually and what factors influence the process.

Gradualness is an important feature in the case of Hungarian too. When the definite article emerged, it did not automatically fill in every possible D-position, but first appeared only in the constructions where the referent of the noun phrase was not anchored in another way.

7. For this discourse, consult first of all Chapters 1 and 7 of Lyons (1999), and Alexiadou et al. (2007: 51–157). For the different approaches and concepts to characterize definite descriptions, see also Abbott (2004).

3.2 The time-span of the research

In order to reconstruct an early but coherent synchronic system for the grammar of reference, the present study will focus on the first half of the Late Old Hungarian Period (end of 14th c. – first half of the 15th c.). Texts from earlier periods (cf. Table 2 above) are short and not continuous. Word-lists and glosses are evidently inadequate to a syntactic analysis. The first continuous texts listed in (11), still from the Early Old Hungarian Period, are not only short, but do not form a uniform corpus either in time or in space. Furthermore, one of these texts is fragmentary, two of them are not narrative and also their exact dating is somewhat problematic.

(11) Funeral Sermon and Prayer (ca.1195, Pray Codex)
The Königsberg Fragment and Ribbons (end of 12th c – beginning of 13th c.)
The Old Hungarian Lament of Mary (religious poetry; beginning of 13th c.)
Gyulafehérvár Lines (second half of 13th c.)

Relevant information about article-use can be gained only from the first codices dating to the first half of the Late Old Hungarian Period. These are long enough, uniform, narrative texts, each of them forming a closed corpus of their own. The sources that might provide us with satisfactory texts for a syntactic investigation therefore are Jókai Codex and two further codices containing translations of the so called Hussite Bible.

(12) Jókai Codex:
The first extant, hand-written book in Hungarian about the life and deeds of Saint Francis of Assisi. A 15th century copy of the original translation from around 1370.

The Hussite Bible (three codices, two of them cited here)
The first Bible translations made between 1416 and 1441.
Vienna Codex (survived copy from 1450): books from the Old Testament and the twelve smaller prophets
Munich Codex (survived copy from 1466): the four gospels of the New Testament

As I sought to draw my linguistic data from a limited and really closed corpus, I chose the Gospel of Matthew from the Munich Codex for the present purposes.[8] My choice has been motivated by the fact that the text of the gospels is suitable for a contrastive diachronic investigation, since they have several translations in the subsequent periods and show a strong traditionalism in content, but still, give way to micro-variation and syntactic change.

---

8. For sake of simplicity the text *loci* will always be given according to the Bible rather than referring to the pages and lines of the codex.

3.3 Search for the absence

To explore the Old Hungarian article-use the following method has been employed: I started to examine systematically the *regular absence* of definite articles, that is, I was searching for nominal expressions where, contrary to the Modern Hungarian usage, the semantic and pragmatic criteria of definiteness seem to hold, but the definite article is absent.

Semantic definiteness (i.e. referential identification) of an entity may be anchored by various ways. The results of a careful classification of the articleless noun phrases with definite interpretation will reveal that the Old Hungarian article appears only when there is no other device to identify the noun phrase reference. It appears very regularly otherwise. Thus overt marking of definiteness is obligatory in the grammar, but does not necessarily entail the presence of the definite article.

4. The Absence of the article: Major types

4.1 Proper names

The definite article is entirely banned with proper names, which are nouns with special lexical properties, being rigid designators in the sense of Kripke (1972). Unlike common nouns, proper names are inherently referential, they refer directly to single individuals, not in virtue of an associated descriptive content. As for their structural location, determinerless proper names end up in D-position, as a consequence of an (overt or covert) N-to-D raising (Longobardi 1994, 2001). What belongs to the class of proper names in a given language is not obvious, there being no independent lexico-semantic definition for proper names. Giuseppe Longobardi (2001: 589) appeals to a syntactic formulation for a generalization in the Romance languages: if N overtly moves to a phonetically empty D then it will be object-referring. Based on English and Germanic data in general, he also assumes that this movement may take place covertly depending on the parameter of D-strength: a lexically empty D is strong in Romance, but not in Germanic. Hungarian seems to be on a par with the Germanic-type languages where proper nouns do not raise overtly: Old Hungarian modifiers always precede the head noun and, unlike Italian, this surface word order does not vary with determinerless proper names either.[9]

---

9. Note that apparently (13) is the only example in the Gospel text for a proper name modified by an adjective. There are further examples, however, containing nouns with a unique referent with the modifier still preceding.

(13) *Az idő-ben hallá negyedlő Heródes Jézus-nak hír-é-t*
that time-INE heard quartering Herod Jesus-DAT fame-POSS-ACC
'In that time Herod the tetrarch heard of the fame of Jesus' [Matt 14:1]

What is worth mentioning in addition is that while Modern Hungarian (first of all, the standard dialect of Budapest) tends to use an article before names of people or institutes, we hardly find a definite article with proper names in Old Hungarian. Moreover, there is a group of lexemes, common nouns in present day language use, which seem to behave as proper names in Old Hungarian. These lexemes, of course, describe entities with a prototypically unique referent, thus can be identified easily by the hearer/reader, at least in a biblical context, e.g. *god, lord, father* (referring to God), *devil, king, queen, prophet, virgin, heaven,* etc. (See the Examples (14) and (15)). Some of them also show a moderate oscillation between being a proper name or a common noun, and in course of time, they get completely reanalyzed as ordinary common nouns.[10] In fact, when modified by an adjective, they tend to have a definite article more frequently than canonical proper names such as person and place names (Imre 1953: 357).

(14) *Mendez kedig lőtt, hogy betelyesednék, mely mondatott*
all.this CONJ became that be.fulfilled which was.said
Ø Úr-tól Ø próféta miatt, mondván: Íme Ø szűz
lord-ABL Prophet through saying behold virgin
vall fiat ő méh-é-ben, és szül
has son-ACC s/he womb- POSS.3SG-INE and brings.forth
'Now all this happened to fulfill what had been said by the Lord through the prophet, saying: Behold, the virgin has a son in her womb, and will give birth…' [Matt 1:22–23]

(15) *Az nap-ot kedig és az idő-t senki nem tudja,*
that day-ACC CONJ and that time-ACC nobody not knows
sem Ø menny-nek angyal-i, sem Ø Fiú, hanemcsak Ø Atya
neither heaven-DAT angel-POSS.PL nor son but.only father
'But nobody knows the day and the time, not even the angels of heaven nor the Son, but only the Father. [Matt 24:36]

10. A preliminary search in the Old Hungarian Corpus shows that the use of the article before these nouns slightly increases overall, but the numbers vary from one manuscript to the other with no predictable tendency. Some of these special lexemes (e.g. *ördög* 'devil', *hold* 'moon', *pokol* 'hell', etc.) do not seem to consistently resist a determiner.

4.2 Modified by a demonstrative

Although in Modern Hungarian the article co-occurs with demonstratives (as in the structures under (10)), the Old Hungarian definite article is missing with nouns modified by a demonstrative element. This can be due to the fact that demonstratives encode directly accessible reference.[11] By their deictic or anaphoric function they imply referentiality and are necessarily definite. Two examples follow, one with a proximal (16) and another with a distal demonstrative (17) in Old Hungarian. To avoid ambiguity, the Latin text has been checked in the second case.

(16) *mondj, hogy e kövek legyenek kenyerek-ké*
Tell that *this* stones become.they breads-FAC
'Tell these stones to become bread' [Matt 4:3]

(17) *Az napok-ban jövö János baptista prédikálván*
that days-INE came John Baptiste preaching
Júdeá-nak kietlen-é-ben
Judea-DAT desolation-POSS-INE
'In those days John the Baptist came, preaching in the wilderness of Judea' [Matt 3:1] (In Latin: *in diebus illis*)

4.3 Generics

More surprisingly, the definite article seems to be absent in case of a generic reading, while in Modern Hungarian it is obligatory, the same way it is in the case of an individual reading. In a generic noun phrase, reference is made to the entire class of entities of which the denotatum of the noun is a member.[12]

Simple generic nominals are often claimed to be rigid designators of some sort, hence they may behave and be interpreted similarly to proper names. In Longobardi's syntactic framework (2001: 595 and 1994: 659–662), if N does not have to move overtly to the D-head because of the weak-parameter setting for D, then it follows for Old Hungarian to have kind-referring (generic) bare nouns.

---

11. Cf. Lyons (1999: 20–21), who refers to Hawkins (1978) and his 'matching constraint': the hearer is instructed to match the referent of the DP with some object which is either identifiable/visible in the context, or which is known on the basis of previous discourse. See also Alexiadou et al. (2007: 93–130) about the syntax and semantics of demonstratives.

12. Alexiadou et al. (2007: 175), Lyons (1999: 179–198). For a general discussion of genericity, see Carlson & Pelletier (1995).

Among the examples below, (18) may be of particular interest, as the first attestations of the two noun phrases 'the winds' and 'the sea' have individual readings and the definite article does appear in both cases. In what follows, the same nouns are mentioned generically and this kind-referring use lacks the article. (This latter has been highlighted by adding a zero sign to the corresponding places).

(18) *Tahát felkelvén parancsola az szelek-nek és az tenger-nek,*
so up.getting commanded the winds-DAT and the sea-DAT
és lőtt vala nagy csendesség. Bizony az emberek
and became AUX big silence verily the men
csudálkodnak vala, mondván: Minemő ez, mert Ø szelek
were.amazed AUX saying what.kind this that winds
és Ø tenger engednek neki?
and sea obey.they to.him
'So he got up and commanded the winds and the sea, and it turned very calm. The men were amazed, saying: "What kind (of man) is this, that the winds and the sea obey him!"' [Matt 8:26–27]

(19) *Látván kedig az gyölekezetek, félemének és dicsővejték*
seeing CONJ the assembled.ones awed.they and praised.they
Isten-t, ki adott ilyen hatalm-at Ø emberek-nek.
God-ACC who gave such power-ACC men-DAT
'But when the crowd saw this, they were awed and praised God, who had given such authority to men' [Matt 9:8]

(20) *Elég Ø tanejtvány-nak, hogy legyen, miként ő mester-e,*
enough student-DAT that be like s/he master-POSS.3SG
és Ø szolgá-nak, hogy legyen, miként ő ur-a.
and servant-DAT that be like s/he lord-POSS.3SG
'It is enough for the student to be like his teacher, and the servant like his lord' [Matt 10:25]

Although the Gospel text abounds in examples for determinerless generics, this group will be the less resistant to the article in later manuscripts. Already in this early Gospel, I came across several places where I tended to interpret the given phrase with a generic reading, but the presence of article was confusing. Unfortunately, the spreading of article in the case of generics can hardly be tested automatically either, since generics are difficult to individualize without involving local semantic and pragmatic factors into the interpretational process.

4.4 Possessive structures

4.4.1 *Pronominal possessors*

In the codices under investigation normally we do not find the definite article in the presence of a pronominal possessor.

(21) *És elhozaték egy tálnyér-on Ø ő fej-e, és*
and was.brought a platter-SUP s/he head-POSS.3SG and
az lány-nak adaték: és vivé Ø ő any-já-nak
the girl-DAT was.given and carried s/he mother-POSS.3SG-DAT
'And his head was brought on a platter and given to the girl, and she carried it to her mother' [Matt 14:11]

There appear to be a few counter-examples: for instance, three occurrences in the Gospel of Matthew (interestingly, all of them coming up in the same chapter). Samu Imre, in his insightful study on the *Vienna Codex* (1953: 354–355), made a statistics for 25 pages and demonstrated a ratio of 294:2 between the noun phrases without any determiner and those in which a determiner and a pronominal possessor co-occur. (But see Section 5 for the spreading).

4.4.2 *Nominal possessors*

Traditional descriptive grammars report of a great oscillation of article use before the possessive construction, where the possessor is either marked by nominative or by dative case. However, on closer inspection, the head of the possessive construction as a whole, i.e. the possessed noun itself never takes an article. The determiners appearing in phrase initial position always belong to the possessor noun phrase only, and, accordingly, are subject to the same distributional rules as in non-possessed contexts. If they are common nouns with an individual reading, they must have an article as a rule (see (22–23)).

(22) *az gyermek-nek lelk-é-t* [Matt 2,20]
the child-DAT soul-POSS-ACC
'the soul of the child'

(23) *az papok fedelm-i-hez* [Matt 26:57]
the priests chiefs-POSS.PL-ALL
'to the chiefs of the priests'

The article is banned in well-defined cases: e.g. when the possessor is modified by a demonstrative (24), when it is a proper name or a noun with a unique referent (25), or the possessor expression itself is possessed, like in (26).

(24) *ez ország-nak kedig fia-i* [Matt 8:12]
this land-DAT CONJ son-POSS.PL
'but the sons of this kingdom'

(25) *mennyek-nek ország-a* [Matt 3:2; 4:17 and *passim*]
heavens-DAT land-POSS
'the land of heavens' (i.e. the kingdom of heaven)

(26) *te láb-aid-nak zsámoly-a* [Matt 22:44]
you foot-POSS.PL.2SG-DAT stool-POSS
'the stool of your feet'

The following Examples (27) with (28), as well as (29) with (30), constitute minimal pairs, where such a distribution can be clearly observed.

(27) *Ø Isten-nek igé-jé-t* [Matt 13:20]
God-DAT word-POSS-ACC
'the word of God'

(28) *az ország-nak igé-jé-t* [Matt 13:19]
the kingdom-DAT word-POSS-ACC
'the word of the kingdom'

(29) *Ø ország-á-nak evangéliom-á-t* [Matt 4:23]
kingdom-POSS.3SG-DAT gospel-POSS-ACC
'the gospel of *his* kingdom'

(30) *az ország-nak evangéliom-a* [Matt 24:14]
the kingdom-DAT gospel-POSS
'the gospel of the kingdom'

As for the possessed part of the expression, no determiner ever appears on the head noun in this manuscript. Possessive constructions, though not necessarily, but prototypically are definite, because the referent of the possessed noun is existentially presupposed and identified via its relation to the referent of the possessor (Koptjevskaja-Tamm 2001: 964). Accordingly, Old Hungarian does not need an article in these constructions, since definiteness becomes evident by the presence of a possessor expression. It is remarkable to note that, as far as I observed during my inquiries, possessives and demonstratives mutually exclude each other in this corpus, suggesting that these expressions occupy the same structural position (probably the specifier of the DP).

4.5 Non-arguments

A further factor influencing determination will be argumenthood that is assumed to be the syntactic reflex of the concept of referentiality. It is the D position that turns

a nominal expression into an argument; consequently, DPs can be arguments, NPs cannot (Longobardi 1994: 620 and 628; also pointed out by Stowell 1989; first proposed in Szabolcsi 1983).[13]

4.5.1 *Predicative constituents*

In Modern Hungarian, arguments can occupy two canonical positions: they either remain in a post-verbal complement position, inside the VP, or undergo topicalization (Alberti 1997). A non-referential nominal expression, however, must leave the VP-domain and move to the predicative part of the clause. It normally lands in a directly preverbal position, in the so-called *verb-modifier* position.

In a smaller collection of data, we seem to have non-referential noun phrases in verb-modifier position (as in (31) and (32)), but in others bare nouns follow the verbs (see (33) and (34)). This latter construction would not be grammatical in Modern Hungarian, where bare noun phrases uniformly appear pre-verbally as they are assumed to occupy the specifier position of the so-called Predicate Phrase (É. Kiss 2006).

At the time of completing this manuscript, our project team is eventually examining the nature of predication and focus structures in Old Hungarian. The preliminary results show that grammaticalization of the preverbal slot had not yet taken place in the period of these early codices. Hence post-verbal bare noun phrases can hardly be tested for argumenthood. The contrastive diachronic investigation of some later translations of the gospels suggests that these are actually non-referential nominals and they can remain undetermined for this reason.

(31) *ki az elhagyott-at vendi, törvény-t tör* [Matt 5:32]
who the left.one-ACC will.take law-ACC breaks
'anyone who marries a left (woman) runs counter to the law'

(32) *Holval levén kedig tanálcs-ot tartanak mend az papi*
morning being CONJ counsil-ACC hold.they all the pristly
fejedelmek és az népek-nek véni Jézsus ellen,
chiefs and the people-DAT elders-POSS.PL Jesus against
hogy őtet halálnak ad-nák
that him death-DAT give- COND.3PL
'In the morning, all the chief priests and the elders of the people held a consultation against Jesus in order to give him to death (i.e. how to have Jesus executed)' [Matt 27:1]

13. Szabolcsi (1994) also speaks about NP subordination. One of the two requirements that make the definite article a compulsory morpheme, proposed by Crisma (2011: 177–178), follows the same line of reasoning: the category D is obligatorily spelled out when the noun phrase is an argument, while non-arguments can be simple NPs.

(33) *ő étk-e kedig vala sáska és vad méz*
s/he food-POSS.3SG CONJ was locust and wild honey
'and his food was locusts and wild honey' [Matt 3:4]

(34) *Bódogok kik éhez-nek és*
happy-PL who-PL hunger.PRES.3PL and
szomjúhoz-nak igazság-ot
thirst.3PL righteousness-ACC[14]
'blessed are those who hunger and thirst for righteousness' [Matt 5:6]

The same phenomenon can be observed with respect to focus constituents. Modern Hungarian has a structural position for identificational focus in the preverbal field, which appears to already exist in Old Hungarian, but is only optionally filled by certain types of constituents. For instance, explicitly contrasted elements tend to remain to the right of the verb, as exemplified in (35) and (36).

(35) *Én valóbizony keresztellek tütök-et víz-ben penitenciá-ra,*
I verily baptize you-ACC water-INE repentance-SUB
Ki kedig jövendő énutánam (…) az keresztel tütök-et
who CONJ coming after.me that baptizes you-ACC
szent lélek-ben és tűz-ben
holy spirit-INE and fire-INE
'I baptize you with water for repentance, But who comes after me (…)
he will baptize you with the Holy Spirit and fire.' [Matt 3:11]

(36) *Mendezek-et beszéllé Jézsus az gyülekezetek-nek*
all-these-ACC spoke Jesus the assembled.ones-DAT
példabeszédek-ben, és példabeszéd nélkül nem beszél
parables-INE and parable without not speak
vala nekik
AUX to.them
'Jesus spoke all these things to the crowd in parables, and he was not
speaking to them without (using) a parable' [Matt 13:34]

4.5.2 *Non argument adjuncts*

There are some cases of determinerless noun phrases in which the lack of an article may be due to the fact that these expressions are adjuncts rather than arguments, thus they may remain unspecified with respect to definiteness.

14. Note that the conjugation of the verb is indefinite, so the object (*righteousness*) cannot be interpreted as a definite generic noun phrase.

(37) *Nem vetekedik, sem üvölt, sem hallja valaki*
not quarrels and.not cries and.not hears someone
ő szavá-t *Ø utcák-ban*
s/he word-POSS.3SG-ACC streets-INE
'He will not quarrel or cry out, no one will hear his voice in the streets'
[Matt 12:19]

5. Spreading of article use and expansion in the left periphery

When trying to apply the criterion of obligatoriness to the supposed Old English definite article, Lotte Sommerer (2011: 193–204) carried out a query that resembles the one presented here to a great extent. Through a quantitative and qualitative analysis of two manuscripts, she collected the unmodified (and thus undetermined) common nouns which occur in clearly definite contexts, and found that definiteness marking is already quite consistent in these texts. It is remarkable that the cases that still resist the use of an article in Old English records can be explained more or less with the same contextual conditions as the ones in my research of Old Hungarian (i.e. with nouns of unique reference, generics, in presence of possessives).

Expansion in the use of the article did not happen simultaneously in the above listed contexts. As a preliminary and tentative research, I made a contrastive analysis checking the articleless noun phrases collected in the Munich Codex against the corresponding loci in a parallel gospel text of a later date. My choice has fallen on Jordánszky Codex, a manuscript that still belongs to the Old Hungarian Period but was compiled about a hundred years after the Hussite Bible translations, between 1516 and 1519. According to my observations, the definite article spread into the generic function first, and, at the same time, also started to appear before possessive pronouns. The latter pattern, however, still shows a massive variation as to having a definite article or not. The modern structures illustrated in (1) and (3), in which the definite article co-occurs with a demonstrative or with a possessive expression, is still not attested. These constructions are more characteristic of the Middle Hungarian Period.[15]

15. These observations were recently confirmed by the automatic queries I carried out using five normalized codices in the Old Hungarian Corpus. First of all, the proportional increase of the definite article is significant. No article can be attested with demonstratives and nominal possessors (this latter had less than 10 hits in the five codices), but steady increase may be observed with pronominal possessives and generics. The paper about these results is in preparation.

These Middle Hungarian phenomena, the definite article appearing together with demonstratives on the one hand, and preceding a possessed noun with dative marked possessor on the other, seem to be interrelated. They might reflect a substantial structural change in the left periphery of the noun phrase. Generative analyses for the Modern Hungarian noun phrase structure (*inter alia* Szabolcsi 1994, É. Kiss 2000) agree on generating the demonstratives (*ez/az* 'this/that') in the specifier position of DP. As far as the dative-marked possessor is concerned, Katalin É. Kiss (2000) proposes that the possessor expression undergoes a noun phrase internal topicalization and lands in a phrase-initial specifier position (Spec, TopP). These assumptions are proved by various word order constraints within the noun phrase. For instance, if both a demonstrative and a nominal possessor are present, the possessor must precede the demonstrative.

The diachronic data, however, show that the transition from the old system to the present one could not happen in one step only, namely by the emergence of a topic position and the movement of dative-marked possessors thereto. In Middle Hungarian, the combination of demonstratives and possessives, though already well attested, shows a peculiar distribution, proving significant differences in the structure of the left periphery.

(38) *Az-ok az Angliá-nak nemes-ei*
that-PL the England-DAT nobleman-POSS.PL
'Those noblemen of England'

(39) *Ar-rul is az bibliá-nak rész-é-rűl*
that-DEL also the Bible-DAT part-POSS-DEL
'About that part of the Bible'

(38) and (39) would not be grammatical in Modern Hungarian, where the position of the demonstratives is already fixed in Spec,DP and only the possessor can leave the DP domain. In Middle Hungarian, however, the demonstrative does not form a prosodic unit with the article, can be separated by various elements, and observably can precede the dative-marked expression as well. The possibility of a more extended noun phrase obviously emerged in this period, but at first two expressions of different nature competed for the same outermost position. In a possible diachronic syntactic analysis, a simple but instable adjunction operation grammaticalized as a Topic phrase within the noun phrase – in accordance with the sentence structure of present-day Hungarian.[16]

16. Cf. Elly van Gelderen's (2008: 250) third universal economy principle, the so-called 'Specifier Incorporation' that claims for elements coming from outside to tend to be a specifier rather than an adjunct.

6. Conclusions

As a closing remark, let me add some thoughts on gradualness in the spirit of Traugott and Trousdale (2010). In most cases change involves small and discrete micro-steps. A given construction may undergo changes at different points in time, and, consequently, the change regarding the construction as a whole may appear to be gradual. The grammatical encoding of referentiality in Hungarian had several stations between the first records to the present day and major restructuring of the nominal left periphery only took place a few centuries after the emergence of the definite article.

Considering the use of the early definite article in the period chosen here (the first part of the Late Old Hungarian Period), the marking of definiteness might be assumed to be incomplete, or rather it is to be analyzed in terms of synchronic 'gradience'. The overt marking of *pragmatic* definiteness is already obligatory, thus the definite article appears if the referent of the noun phrase either cannot be identified by its independent semantic properties (e.g. inherently unique reference), or its definiteness is not encoded by an alternative syntactic device (e.g. a demonstrative modifier). After the reanalysis of the article as a default definiteness marker, however, a further extension can be observed in its use, which finally results in the syntactic reformulation of the noun phrase as a whole. One point of change must have been the more and more frequent spelling out of the D head, which generalized to nominals whose reference was identifiable for semantic reasons (i.e. generics and nouns with unique reference). To put it differently, semantic definiteness also started to be encoded in syntax by the grammaticalized definiteness marker. The gradualness of this change can also be detected in the early corpus of the Munich Codex by the unstable status of generic noun phrases, which already show a slight oscillation with respect to their resistance to the article. This micro-step was followed by a further generalization: from the Middle Hungarian period, the definite article also had to be spelled out in contexts where the reference was already anchored by other syntactic strategies, such as the presence of a demonstrative modifier or a possessor expression. With the increased frequency of structures headed by the article, analogy may also have had a role in the process during which article use expanded to almost every pattern embodying a definite description. The sequence of these micro-changes had further consequences in the form of a substantial structural change in the left periphery of the noun phrase (the details of which have been sketched in Section 5). The precondition of the successive reanalyses that can be observed through the interplay between the Middle Hungarian possessives and demonstratives, undoubtedly, was the gradual expansion of the syntactic role of the article.

Focus on the dynamic interface between synchrony and diachrony

The *Grammatical encoding of referentiality in the history of Hungarian* is a paper about the Old Hungarian synchronic system of definiteness marking, and aims to demonstrate that the grammaticalization of referential identification in the form of the definite article may take place in a gradual manner, by subsequent micro-steps as well.

Compared to the present day distribution, the article in the first part of the Late Old Hungarian Period appears in a smaller amount and in fewer contexts. A modern reader may even have the impression that the variation in its occurrences is somewhat arbitrary or unsystematic. This paper argues for the true article status of this early determiner rather than its being a transitional category. The Old Hungarian article is a fully grammaticalized element encoding definiteness on a syntactic level, but has a more restricted use than it does in the subsequent language stages. Systematically examining the regular absence of articles in definite contexts, in a selected closed corpus, reveals that the definite article appears only in the constructions where the referent of the noun phrase either cannot be identified by independent semantic properties, or its definiteness is not encoded by an alternative syntactic device (e.g. a demonstrative modifier or a possessive construction). What follows from these observations is that the emergence of a new grammatical category does not necessarily entails the re-articulation of the system as a whole in one step only. In fact, the gradual spreading of the article can be observed first to encode semantic definiteness, and only much later expands to every construction whose interpretation is definite. The sequence of these micro-changes will have a major structural effect in Middle Hungarian only, when new patterns arise for demonstrative modification and its combination with possessives.

Abbreviations

1	first person
3	third person
ABL	ablative
ACC	accusative
ALL	allative
AUX	auxiliary
COND	conditional
CONJ	conjunction
DAT	dative
DEL	delative

FAC	factive
ILL	illative
INE	inessive
PL	plural
POSS	possessive
SG	singular
SUB	subessive
SUP	superessive

Old Hungarian text editions consulted

P. Balázs János 1981. *Jókai-kódex XIV–XV. század*. Budapest: Akadémiai Kiadó, 9–20.

A Jordánszky-kódex bibliafordítása. Sajtó alá rendezte és kinyomtatta: Toldy Ferenc, az eredetivel összevetette, és előszóval ellátta: Volf György. Régi magyar nyelvemlékek 5. Buda, 1888.

Müncheni kódex [1466] A négy evangélium szövege és szótára. Décsy Gyula olvasata alapján a szöveget sajtó alá rendezte és a szótári részt készítette Szabó T. Ádám. Budapest: Európa Könyvkiadó. 1985.

Nyíri Antal 1971. *A Müncheni Kódex 1466-ból. Krirtikai szövegkiadás a latin megfelelővel együtt* [Codices Hungarici 7]. Budapest: Akadémiai Kiadó.

References

Abbott, Barbara. 2004. Definiteness and Indefiniteness. In *The Handbook of Pragmatics,* Laurence R. Horn & Gregory Ward (eds), 122–149. Oxford: Blackwell.

Alberti, Gábor. 1997. Restrictions on the degree of referentiality of arguments in Hungarian sentences. *Acta Linguistica Hungarica* 44: 341–362.

Alexiadou, Artemis, Haegeman, Liliane & Stavrou, Melita. 2007. *Noun Phrase in the Generative Perspective* [Studies in Generative Grammar 71]. Berlin: Mouton de Gruyter.

Bakró-Nagy, Marianne. 1999. A magyar határozott névelő kialakulásának szintaktikai vonatkozásai (Syntactic aspects of the development of the Hungarian definite article). In *A nyelvtörténeti kutatások újabb eredményei I. Magyar és finnugor mondattörténet,* Büky László & Forgács Tamás (eds), 5–13. Szeged: JATE Magyar Nyelvészeti Tanszék.

Carlson, Greg N. & Jeffry Pelletier, Francis (eds). 1995. *The Generic Book*. Chicago IL: The University of Chicago Press.

Crisma, Paola. 2011. The emergence of the definite article in English. In *The Noun Phrase in Romance and Germanic. Structure, Variation, and Change* [Linguistik Aktuell/Linguistics Today 171], Petra Sleeman & Harry Perridon (eds), 175–192. Amsterdam: John Benjamins.

É. Kiss, Katalin. 2000. The Hungarian noun phrase is like the English noun phrase. In *Approaches to Hungarian 7,* Alberti Gábor & Kenesei István (eds), 121–149. Szeged: JATE Press.

É. Kiss, Katalin. 2006. Focussing as predication. In *The Architecture of Focus*, Valéria Molnár & Susanne Winkler (eds), 169–196. Berlin: Mouton de Gruyter.

I. Gallasy, Magdolna. 1992. A névelők (The articles). In *A magyar nyelv történeti nyelvtana,* II/1: *A kései ómagyar kor. Morfematika,* Loránd Benkő (ed.), 716–770. Budapest: Akadémiai Kiadó.

van Gelderen, Elly. 2008. Linguistic cycles and economy principles. The role of Universal Grammar in language change. In *Grammatical Change and Linguistic Theory. The Rosendal Papers* [Linguistik Aktuell/Linguistics Today 113], Thórhalur Eythórsson (ed.), 245–264. Amsterdam: John Benjamins.

Giusti, Giuliana. 2001. The birth of a functional category: From Latin ILLE to the Romance article and personal pronoun. In *Current Studies in Italian Syntax. Essays Offered to Lorenzo Renzi* [North Holland New Linguistic Series 59] Guglielmo Cinque & Giampaolo Salvi (eds), 157–171. Oxford: Elsevier.

Hawkins, John A. 1991. On (in)definite articles: Implicatures and ungrammaticality prediction. *Journal of Linguistics* 27: 405–422.

Himmelmann, Nikolaus P. 2001. Articles. In *Language Typology and Language Universals* [Handbücher der Sprach und Kommunikationswissenschaft, Band 20.1], Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds), 831–841. Berlin: Walter de Gruyter.

Imre, Samu. 1953. A határozott névelő használata a Bécsi Kódexben (The use of the definite article in Vienna Codex). *Magyar Nyelv* 49: 348–359.

Koptjevskaja-Tamm, Maria. 2001. Adnominal possession. In *Language Typology and Language Universals* [Handbücher zur Sprach- und Kommunikationswissenschaft, Band 20.2], Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds), 960–970. Berlin: Walter de Gruyter.

Kripke, Saul A. 1972. Naming and necessity. In *Semantics of Natural Language,* Donald Davidson & Gilbert Harman (eds), 253–355. Dordrecht: Reidel.

Longobardi, Giuseppe. 1994. Reference and proper names: A theory of N-movement in syntax and logical form. *Linguistic Inquiry* 25: 609–665.

Longobardi, Giuseppe. 2001. The structure of DPs: Some principles, parametres, and problems. In *The Handbook of Contemporary Syntactic Theory,* Mark Baltin & Chris Collins (ed.), 562–603. Oxford: Blackwell.

Lyons, Christopher. 1999. *Definiteness* [Cambridge Textbooks in Linguistics] Cambridge: CUP.

Roberts, Craige. 2003. Uniqueness in definite noun phrases. *Linguistics and Philosophy* 26: 287–350.

Sommerer, Lotte. 2011. Old English *se:* From Demonstrative to Article. A Usage-based Study of Nominal Determination and Category Emergence. Ph.D. dissertation, Universität Wien.

Stowell, Timothy. 1989. Subjects, specifiers, and X-bar theory. In *Alternative Conceptions of Phrase Structure,* Mark Baltin & Anthony Kroch (eds), 232–262. Chicago IL: University of Chicago Press.

Szabolcsi, Anna. 1983. The possessor that ran away from home. *The Linguistic Review* 3: 89–102.

Szabolcsi, A. 1994. The noun phrase. In *The Syntactic Structure of Hungarian* [Syntax and semantics 27] Ferenc Kiefer & Katalin É. Kiss (eds), 179–274. San Diego CA: Academic Press.

Traugott, Elisabeth Closs & Trousdale, Graeme. 2010. Gradience, gradualness and grammaticalization: How do they intersect? In *Gradience, Gradualness and Grammaticalization* [Typological Studies in Language 90], Elisabeth Closs Traugott & Graeme Trousdale (eds), 19–44. Amsterdam: John Benjamins.

Westerståhl, Dag. 1985. Determiners and context sets. In *Generalized Quantifiers in Natural Language*, Johan van Benthem & Alice G.B. ter Meulen (eds), 45–72. Dordrecht: Foris.