

Gradual expansion in the use of the definite article Checking a theory against the Old Hungarian Corpus

Barbara Egedi & Eszter Simon
Research Institute for Linguistics
Hungarian Academy of Sciences

1 Introduction

The *Hungarian Generative Diachronic Syntax* Project (2009-2013)

- supported by the Hungarian Scientific Research Fund (OTKA No. 78074)
- located in the Research Institute for Linguistics of the Hungarian Academy of Science

The aim of the project is twofold:

- 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
- to reconstruct the syntax of different synchronic systems, and to examine and model the grammaticalization processes

The aim of the present talk:

- to explore and present the possibilities and the limits of checking a linguistic hypothesis against a larger amount of data

Data: five (at least) normalized codices from the Old Hungarian Corpus

The hypothesis to be checked: the expansion in the use of the article

Table 1. Language stages of Hungarian and types of sources

Proto-Hungarian		1000 BC – 896 AD	No written documents; reconstructed
Old Hungarian	Early Old Hungarian	The age of the Árpád dynasty	Sporadic records, glosses, a few short texts, e. g. <i>Funeral Sermon and Prayer</i> (ca.1195)
	Late Old Hungarian	From around 1370 (time of compilation of Jókai Codex) to 1526	Mainly codices: translations of Latin religious literature + original Hungarian compositions (documents, poems and letters)
Middle Hungarian		1526-1772	Book printing New secular genres appear Significant increase in quantity
Modern Hungarian		1772- present day	

2 The definite article in Old Hungarian

2.1 The data

Modern Hungarian makes extensive use of the definite article
Definite article together with possessives:

Modern Hungarian

(1) a szerzetes-nek **a** könyv-e
the monk-DAT *the* book-POSS
'the book of the monk'

(2) **a** te könyv-ed
the your book-POSS.2SG
'your book'

Old Hungarian

(3) a(z) szerzetes-nek **∅** könyv-e
the monk-DAT book-POSS
'the book of the monk'

(4) **∅** te könyv-ed
your book-POSS.2SG
'your book'

Definite article together with demonstratives:

Modern Hungarian

(5) ez/az **a** könyv
this/that *the* book
'this/that book'

Old Hungarian

(6) e(z)/a(z) **∅** könyv
this/that book
'this/that book'

About the syntactic structure of the Hungarian noun phrase, see Szabolcsi 1994, É. Kiss 2000

Grammaticalization: distal demonstrative modifier > definite article

⇒ they cannot be distinguished from each other merely on formal grounds: formal and positional equivalence, and functional overlap (e.g. anaphoric use)

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

First attestations of the definite article in traditional grammars: pre-articles; transitional word-class; dual nature: "pronoun-article". (Bakró-Nagy 1999: 7; I. Galassy 1992: 721-722)
For the same problem in Old English, see e. g. Sommerer (2011)

Semantic/pragmatic criteria to identify an article (Himmelman 2001: 832-834)

- the larger situation use (= first mention of entities that are considered to be unique, hence generally identifiable, in a given speech community)
- the associative-anaphoric use (=first mention of an entity that is not unique per se but with respect to a previously mentioned referent)

For a more detailed account of the issue: Egedi (*forthcoming*), with further references

2.2 The hypothesis

Search for the regular absence of definite articles in one of the early codices (first half of the Late Old Hungarian period)

→ the Gospel of Matthew in the *Munich Codex* (composed between 1416 and 1441. The survived copy is from 1466)

Aim: an investigation carried out by a manual search on a closed uniform text
– to classify of the articleless noun phrases with a definite interpretation
– to understand why the article is still absent

Theoretical background for the notion of definiteness:

inter alia Lyons 1999, Abbott 2004, Alexiadou 2007: 51-157.

Key concepts:

Identifiability: the speaker signals that the hearer is able to assign a referent for a certain DP.

Definite article is the grammaticalization of the semantic and pragmatic concept of definiteness

Conclusion of the research: The early definite article appears only in the constructions where the referent of the noun phrase is not anchored in another way, thus absent:

- with proper names and with a group of lexemes that describe entities with a prototypically unique referent
- with demonstratives
- possessor expressions
- in case of a generic reading of the noun phrase
- if the noun phrase is an adjunct rather than an argument of the verb

2.3 Absence of article in definite contexts

2.3.1 Proper names

2.3.1 Proper names and nouns with a unique referent

Nouns with special lexical properties → inherently referential
Rigid designators (Kripke 1972)

A group of lexemes (common nouns in present day language use) seem to behave as proper names in Old Hungarian

→ they 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, gospel*, etc.

→ 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.

→ if modified by an adjective, they tend to have a definite article more frequently than canonical proper names such as personal and place names. (Imre Samu 1953: 357).

(8) és Ø *atyá-t* senki sem esmerte hanemcsak Ø *fiú-t*
and father-ACC nobody not knew but son-ACC
'neither knoweth any men the Father, save the Son' [Matt 11:27]

(9) Mendez kedig lött, hogy betelyesednék, mely mondatott Ø *Úrtól*
all-this PRT became that be-fulfilled which was-said lord-ABL
Ø *próféta* miatt, mondván: Íme Ø *szűz* vall fiat ő méhében, és szül
Prophet through saying behold virgin has son-ACC her womb-INE and brings-forth
'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]

- (10) Az napot kedig és az időt senki nem tudja,
 that day-ACC PRT and that time-ACC nobody not knows
 sem Ø *mennynek* angyali, sem Ø Fiú, hanemcsak Ø Atya [Matt 24:36]
 neither heaven-DAT angels-PL.POSS nor son but-only father
 ‘Nobody knows the day and the time, not even the angels of heaven nor the Son, but only the Father.

2.3.2 Modified by a demonstrative

Demonstratives encode directly accessible reference → necessarily definite

Lyons (1999: 20-21): a kind of ‘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

About the syntax and semantics of demonstratives: Alexiadou 2007: 93-130 and Lyons 1999

- (11) *Az napok-ban* jövö Jánus baptista prédikálván Júdeá-nak kietlen-é-ben
 that days-INE came John Baptiste preaching 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*)
- (12) mondj, hogy *e kövek* legyenek kenyerek-ké
 Tell that *this* stones become.they breads-FAC
 ‘Tell these stones to become bread’ [Matt 4:3]
- (13) És ím egy kánanea-i némbéri *az vidék-ből* kijövén ivölt vala
 and behold a Canaan-ADJ woman that region-ELA out-coming cries AUX
 ‘And, behold, a woman of Canaan came out of the same coasts, and cried’ [Matt 15:22]

New pattern: *determiner doubling*

→ the simultaneous spelling out of the demonstrative modifier and the article

→ appears only in the 16th century and is characteristic of the Middle Hungarian Period

- (5’) az **a** könyv
 that *the* book
 ‘that book’

2.3.3 Possessive structures

Possessive constructions → prototypically definite; the referent of the possessed noun is existentially presupposed and identified via its relation to the referent of the possessor (Koptjevskaja-Tamm 2001: 964)

No determiner appears on the head noun!

Pronominal possessors

- (14) És elhozaték egy tálnyér-on \emptyset ő fej-e, és
 and was.brought a platter-SUP his head-POSS.3SG and
 az lány-nak adaték: és vivé \emptyset ő any-já-nak
 the girl-DAT was.given and carried her 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]

Nominal possessors

- (15) az gyermek-nek lelk-é-t (16) az papok fedelm-i-hez
 the child-DAT soul-POSS-ACC the priests chief-POSS.PL-ALL
 ‘the soul of the child’ [Matt 2:20] ‘to the chiefs of priests’ [Matt 26:57]

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.

- (17) \emptyset Isten-nek igé-jé-t [Matt 13:20] (18) az ország-nak igé-jé-t [Matt 13:19]
 God-DAT word-POSS-ACC the kingdom-DAT word-POSS-ACC
 ‘the word of God’ ‘the word of the kingdom’
 (19) \emptyset ország-á-nak evangéliom-á-t [Matt 4:23]
 kingdom-POSS-DAT gospel-POSS-ACC
 ‘the gospel of his kingdom’
 (20) az ország-nak evangéliom-a [Matt 24:14]
 the kingdom-DAT gospel-POSS-NOM
 ‘the gospel of the kingdom’

2.3.4 *Generics*

Reference is made to a kind rather than an individual

Generic noun phrase: reference is made to the entire class of entities of which the denotatum of the noun is a member. (Alexiadou 2007: 175, for general discussion: Carlson and Pelletier 1995)

Generics ~ proper names

Generic reading: without article \leftrightarrow individual reading: with article

- (21) 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 \emptyset szelek
 are.amazed AUX saying what.kind this that winds
 és \emptyset 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]

(22) Látván kedig az gyölekezetek, félemének és dicsővejték Istent,
 seeing PRT the assembled-PL awed-3PL and praised-3PL God

ki adott ilyen hatalmat *Ø emberek-nek*. [Matt 9:8]
 who gave such power-ACC men-DAT

‘When the crowd saw this, they were awed and praised God, who had given such authority to men’

(23) Elég *Ø* tanejtvány-nak, hogy legyen, miként ő mestere,
 enough student-DAT that be-3SG like his master-POSS.3SG

és *Ø* szolgá-nak, hogy legyen, miként ő ura. [Matt 10:25]
 and servant-DAT that be-3SG like his lord-POSS.3SG

‘It is enough for the student to be like his teacher, and the servant like his lord’

2.3.5 Non-arguments

Argumenthood: syntactic reflex of the concept of referentiality

In Modern Hungarian: Non-referential nominal expressions → in the verb-modifier position or in structural focus position

In Old Hungarian: Non-arguments and focus constituents are optionally in post-verbal position; non-argument adjuncts can remain unspecified

3. The definite article in Old Hungarian – spreading

In the Middle Hungarian Period (from the 16th century onward) the definite article appears in new contexts:

- co-occurring with demonstratives
- preceding a possessed noun with dative-marked possessor

Manual checking of the NT *loci* in a later Old Hungarian ms. Codex Jordánszky (1516-1519)
 → expansion in article use with generic NPs and before possessive pronouns

Aims of Corpus checking:

- to demonstrate the proportional increase in the use of the article *within* the Old Hungarian Period
- to find out in which context(s) it took place earlier

Table 2. The codices used for checking

	Date	Content	Level of text-processing
Jókai Codex	a.1370/c.1448	life and deeds of Saint Francis of Assisi	morpho-syntactically annotated
Vienna Codex	a.1416/c.1450	books from the Old Testament	normalized
Guary Codex	before 1495	texts about righteous life, virtues, etc. written for nuns	morpho-syntactically annotated
Booklet on the Dignity of the Apostles	1521	A tractate about the apostles (with many dialogues)	morpho-syntactically annotated
Kazinczy Codex	1526-1541	sermons, legends, parables	normalized

Table 3. The proportion of definite articles in five Old Hungarian codices

	<i>a/az</i>	Tokens	%
Jókai Codex	550	22733	2,42
Vienna Codex	2225	54423	4,09
Guary Codex	1387	20288	6,84
Booklet	621	8745	7,10
Kazinczy Codex	2042	23591	8,66
<i>SzIT (Modern)</i>	68562	616131	11,12

NB. The hits always include the distal demonstratives, but it does not effect the proportional increase.

Spreading in various contexts

Table 4. Increase in the use of the article in specific contexts

Context	Method	Increase in article use
Nouns with a unique referent	One by one checking of lexical items in the Corpus	slightly; inconsistent results
Demonstratives	Automatic query	none
Possessives (pronominal)	Automatic query	√
Possessives (nominal)	Automatic query	none/minimal
Generics	Not possible automatically	???

Table 5. Examples for nouns with a unique referent:

Codex	Jókai		Vienna		Guary		Booklet		Kazinczy	
	–	+	–	+	–	+	–	+	–	+
hold ‘moon’	1	0	7	0	0	9	0	0	5	0
föld ‘earth’	24	11	143	47	14	6	0	2	3	11
pokol ‘hell’	8	0	8	1	17	10	10	0	3	0
isten ‘god’	176	2	264	2	141	11	83	0	125	26

Barbara Egedi: egedib@yahoo.com
 Eszter Simon: simon.eszter@nytud.mta.hu
 Corpus query tool: <http://ohc.nytud.hu>
 Downloadable publications: <http://www.nytud.hu/oszt/elmnyelv/mgtsz.html>

References

- Abbott, Barbara 2004. Definiteness and Indefiniteness. In: Laurence R. Horn – Gregory Ward (eds.): *The Handbook of Pragmatics*. Blackwell Publishing, 122-149.
- Alexiadou, Artemis et al. 2007. *Noun Phrase in the Generative Perspective*. Studies in Generative Grammar 71. Berlin – New York: 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: Büky László – Forgács Tamás (szerk.): *A nyelvtörténeti kutatások újabb eredményei I. Magyar és finnugor mondattörténet*. Szeged, JATE Magyar Nyelvészeti Tanszék, 5-13.
- Carlson, Greg N. and F. Jeffrey Pelletier (eds.) *The Generic Book*. Chicago, The University of Chicago Press
- Egedi Barbara (forthcoming) Grammatical encoding of referentiality in the history of Hungarian In: Anna Giacalone Ramat – Caterina Mauri – Piera Molinelli (eds.): *Synchrony and Diachrony: a Dynamic Interface. Studies in Language Companion*. Amsterdam: John Benjamins
- É. Kiss Katalin 2000. The Hungarian Noun Phrase is like the English Noun Phrase. In: Alberti Gábor – Kenesei István (eds.) *Approaches to Hungarian 7*. Szeged, 121-149.
- I. Gallasy Magdolna 1992. A névelők. [= The articles] In: Benkő L. (szerk.), *A magyar nyelv történeti nyelvtana II/1. A kései ómagyar kor. Morfematika*. Budapest: Akadémiai Kiadó. 716-770.
- Himmelmann, Nikolaus P. 2001. Articles. In: Martin Haspelmath et al. (eds.): *Language Typology and Language Universals*. Handbücher der Sprach und Kommunikationswissenschaft. Volume 20.1. Berlin – New York: Walter de Gruyter, 831-841.
- Imre Samu 1953. A határozott névelő használata a Bécsi Kódexben. [The use of the definite article in Codex Vienna] *Magyar Nyelv* 49. 348-359.
- Koptjevskaja-Tamm, Maria 2001. Adnominal possession. In: Haspelmath, Martin et al. (szerk.) *Language Typology and Language Universals*. Handbücher zur Sprach- und Kommunikationswissenschaft. Band 20.2. Walter de Gruyter, Berlin – New York. 960-970.
- Kripke, Saul A. 1972. Naming and necessity. In: Donald Davidson and Gilbert Harman (eds.): *Semantics of natural language*. Dordrecht: Reidel. 253-355.
- Lyons, Christopher 1999. *Definiteness*. Cambridge University Press, Cambridge.
- Sommerer, Lotte 2011. Old English *se*: from demonstrative to article. A usage-based study of nominal determination and category emergence. PhD dissertation, Universität Wien
- Szabolcsi, Anna 1994. The noun phrase. In Kiefer, F. – Katalin É. Kiss (eds.) *The syntactic structure of Hungarian*. Syntax and semantics 27. San Diego: Academic Press. 179-274.