



On diachronic variation and change of aligha 'unlikely' and alighanem 'probably' : from low intensity to epistemic modality

Varga, Mónika

(Citation)

Papers from the Second International Workshop on the Syntax of Predication and
Modification:1-19

(Issue Date)

2026-02-25

(Resource Type)

conference proceedings

(Version)

Version of Record

(JaLCD0I)

<https://doi.org/10.24546/0100501353>

(URL)

<https://hdl.handle.net/20.500.14094/0100501353>



On diachronic variation and change of *aligha* ‘unlikely’ and *alighanem* ‘probably’: from low intensity to epistemic modality*

Mónika Varga
ELTE Hungarian Research Centre for Linguistics

Abstract

This paper investigates changes in the distribution of the Hungarian epistemic and inferential modifiers *aligha* ‘unlikely’ and *alighanem* ‘probably’. Before the late 19th century, *aligha* predominantly expressed high probability which was syntactically marked by a negative particle (as a case of double negation). In Modern Hungarian, however, it primarily marks the negative pole by default, being incompatible with affirmative responses – which is an atypical feature among Hungarian modifiers. The extensive analysis of historical data, covering sources from the 16th to the mid-20th century, highlights that the structural characteristics of the host sentences containing *aligha* have also changed. Before the late 19th century, when *aligha* appeared in affirmative sentences, the predicate displayed neutral word order, whereas in Standard Modern Hungarian *aligha* triggers inversion between the verb and the verbal prefix. The study compares and contrasts functional variants in order to find plausible motivations for the observed changes.

Keywords: epistemic and inferential modification; historical corpus analysis; functional variation; word order; Hungarian; *aligha*; *alighanem*

1. Introduction

In standard Modern Hungarian, the epistemic and inferential modifier *aligha*, meaning ‘barely’ and ‘unlikely’, depending on the context, primarily expresses that the occurrence of the event it introduces or the existence of the state is unlikely. In the Hungarian Giga-word Corpus which contains recent Hungarian data, *aligha* ‘probably not’ has more than 27000 instances (for the description of the database, see Oravecz, Váradi, and Sass 2014). Consider (1).¹

- (1) *Rend, az persze aligha lesz*
order that of-course unlikely be-FUT-3SG
‘Order, that unlikely will be.’ (MNSZ2, #215159, 2003)

* This paper was presented at The 2nd International Workshop on the Syntax of Predication and Modification (IWSPM 2025) held in Budapest, Hungary on September 1-2, 2025. The research was funded by National Research, Development and Innovation Office – Hungarian Scientific Research Fund (NKFIH – OTKA) project grant 135186.: Variation in Middle Hungarian: a register perspective.

¹ Early Modern and Modern Hungarian data are cited according to corpora, with reference to database, to the identifier of the source and the year of the publication given as metadata.

The modifier *aligha* is used less frequently to express high probability as well: in this case, however, the form is different. *Alighanem* ‘probably, likely’ has less than 15000 instances in the Hungarian Giga-word Corpus. This is illustrated in (2).

- (2) *Mit rugdos? – Alighanem labdát*
 what-ACC kick.3SG probably ball-ACC
 ‘What is he kicking? – Probably a ball’ (MNSZ2, #977971, 2003)

Thus, as Modern Hungarian data suggests, *aligha* ‘barely, unlikely’ marks the negative pole that is incompatible with an affirmative response by default while those that refer to the positive pole are marked lexically and they occur less frequently. Nevertheless, *aligha* is a rather unique modifier in this respect, since marking the negative pole is an atypical feature for modifiers in Hungarian (Kugler 2009). The Second Edition of *the World Lexicon of Grammaticalization* (Kuteva et al. 2019) does not mention any examples suggesting that it is uncommon in other languages as well. Nevertheless there was a modifier in Middle English, *unethes* which collocated both with *any* and *n*-series (e.g. *never*), unlike *hardly* (Tubau and Ingram 2015).

The present paper argues that the current distribution of forms and functions of *aligha* is the result of a change that can be observed during the 19th century, according to historical corpus data (Section 3). Between the 16th and the 18th centuries, *aligha* predominantly expressed high (positive) probability, typically in clauses containing a negative particle, as it is shown in (3).²

- (3) *Urunk aligha le nem megyen*
 lord-POSS-1SG unlikely VPFX not go-3SG
 ‘Our lord probably goes down.’ (TMK, Kár. 32., 1705)

During the Early Modern Hungarian Period, the negative particle fused with the modifier *aligha*, giving rise to *alighanem* ‘probably, likely’ (Simonyi 1881, Klemm 1928–1942, D. Máta 2003). *Aligha* followed directly by *nem* can be attested as early as the 17th century, though it only lexicalized in the 19th century. The details of this change and its structural consequences are insufficiently explored yet (on recent use, see Kugler 2009, 2013, 2014).

The phenomenon under scrutiny does involve some syntactic implications to be unraveled, although this paper does not aim to describe this change within a specific theoretical framework. The main research questions are the following:

1. If *aligha* occurs in the negative sentence that contains a verb with a verbal prefix (or particle), what word order pattern do they exhibit?
2. Similarly, what word order does the affirmative sentence exhibit containing *aligha* and a verb with a verbal prefix (or particle, depending of the approach)?
3. Can any change be observed regarding a) the spread of *alighanem* ‘probably’ and b) the spread of *aligha* ‘probability not’ and what are these changes?

² Historical data are cited in a normalized form: the spelling reflects to Modern Hungarian, while the original morphologic structures are preserved. For full citations, see Sources.

Until the 19th century, the VPFX–NEG–V was the default order in negative sentences, e.g., (3) *le* (VPFX) – *nem* (NEG)– *megyen* (V). This is claimed to be a heritage from *ha* ‘if’, a constituent of *aligha* (Kugler 2009, Gugán 2021).V

In Modern Hungarian neutral sentences, the word order of verbs with a verbal prefix is VPFX–V. Earlier research suggests that clauses with among others *alig* meaning ‘hardly, barely’ are non-neutral in that respect that the ‘at least’ interpretation is excluded and the negative polarity adverbial triggers inversion of the verbal prefix (or particle) and the verb due to its exclusive behavior (É. Kiss 2006). Kugler (2009, 2013) argues that in Modern Hungarian, *aligha* behaves similarly due to the restrictive use of one of its constituent *alig*, which *aligha* inherited. For example, *aligha oldotta* (V) *meg* (VPFX/PRT) ‘(s)he barely resolved it’ is grammatical, while the order *aligha megoldotta* (VPFX–V) is ungrammatical. The present study investigates this problem involving qualitative and quantitative analysis on historical linguistic material (for the details, see Section 4) where the changing ratios of negative and affirmative patterns and the related structural variations and developments are concerned. The question of the interaction between negation and probability is also examined as well as the scope of *aligha* and *alighanem*.

The outline of the paper is as follows. First, the constituents of the compound *aligha* – *alig* ‘hardly, barely’ and *ha* ‘if’ – are described. Section 2 also raises the question that *aligha* may have been grammaticalized from a bi-clausal structure (Simonyi 1881, Klemm 1928-1942, Kugler 2009, 2013). Because of the insufficiency of the available actual historical linguistic data, many aspects of the question remain open. Then, the studied sources and the method of the analysis is provided (Section 3). This is followed by the corpus analysis, comparing and contrasting Middle Hungarian, Early Modern Hungarian, and Modern Hungarian data (Section 4). Finally, the structural aspects raised by the studied corpus data are discussed, especially the syntactically marked *aligha* [...] *nem* and the changes in the word order.

2. *Alig, ha, aligha (nem)*: Origin and open questions

To describe *aligha* in advance: it is a compound of two constituents, namely the adverbial *alig* which is a lexicalized inflected form of *al* meaning ‘lower part’ and the terminative suffix *-ig*, meaning ‘to a low extent ~ with difficulty’. Therefore, *alig* is a minimizer with the meaning ‘to a very low degree, barely’.³ Consider (4) and (5).

- (4) *mely számot embernek ő elméje*
 which number-ACC people-DAT (s)he mind-POSS-3SG
csak alig avagy nagy nehezen foghatja be
 only barely or great hardly grasp-MOD-3SG VPFX
 ‘which number the human mind can only grasp barely or with great difficulty’ (ÓMK, ÉrsK.)

³ *Alig* has a variety of meanings since Old Hungarian, such as collocations expressing intensity, *alig vár* ‘hardly can wait’, e.g.: *Alig várnám, hogy általesném rajta*. ‘I can hardly wait to get through it’ (TMK, Bark. 297.); near simultaneity or anteriority e.g.: *Haller Gábor Uram, alig szálltam meg, mindjárt érkezett* ‘Mr. Gábor Haller, I had barely settled in, when he arrived’ (TMK, Bark. 259.); low quantity: *alig van három lépésnyire tőlem* ‘it is barely three steps away from me’ (Székely László). However, most of these uses are over the scope of the present paper.

- (5) *járni való ereje is alig volt*
 walk-INF be-PRS strength too barely was
 ‘he barely had the strength to walk.’ (TMK, Bosz. 400., 1767)

Meanings ‘to a very low degree’ and ‘with difficulty’ can be regarded as sources for the development of the meaning ‘probably (not)’, such as in (6).

- (6) *isteni édességet kér, de csak alig,*
 divine sweetness-ACC ask-3SG but only barely
avagy soha meg nem nyerheti
 or never VPFX not win-MOD-3SG-DEF
 ‘asks for divine sweetness, but can only barely, or never get it’ (ÓMK, VitkK.)

The second part of the compound is *ha* meaning ‘if’, which is the conjunction of conditional clauses, originally with temporal and ‘whether’ meanings (see Klemm 1928–1942, Juhász 1991). The meaning ‘whether’ is still actively used in the Middle Hungarian period (before 1772) but the temporal ‘when’ has become obsolete.

Aligha is described as a compound where the constituents are not related syntactically (D. Máta 2003: 658, 2011: 268). Also, *alig* and *ha* have only one common use as a minimizer meaning ‘barely’, but specifically restricted to quantities (7).

- (7) *lovag, gyalog ötezer, ha volt.*
 knight foot-soldier five-thousand if was
 ‘five thousand of knights and foot soldiers, if there were that much’ (TMK, Tel. 35.).

A very similar meaning can be constructed with *alig* as well (see Footnote 2). The former example is related to probability in a broad sense since it does not address the question whether there were any knights or foot soldiers but how many of them were there.

Aligha as a compound can be attested as early as the 16th century. It is claimed to be grammaticalized from a bi-clausal structure (Simonyi 1881, Klemm 1928–1942, Kugler 2009). However, no actual linguistic data have yet been found to confirm the actual preceding structure (Gugán 2021). According to earlier approaches, the possible candidates for the main verb are *van* ‘is’ or *lehet* ‘may/can be’ (Simonyi 1881, Klemm 1928–1942), and *hisz* ‘believes’ is also mentioned (Simonyi 1881). The available historical data do not confirm the assumptions regarding the verb *van* or *lehet*. Main clauses involving *alig van* ‘it is barely’ or *alig lehet* ‘it can be barely’ are only exceptionally attested with a negative particle in the subordinate clause (see also Gugán 2021). Also, *alig lehetséges* ‘barely possible’ is attested only in the 20th century.

The conjunction *ha* ‘if’ does not appear in these clauses either. If there is a bi-clausal structure with *van* ‘is’ or *találkozik* in this context, meaning ‘can be found’, it connected with a relative pronoun *amely* or *mely* ‘which’. Additionally, probability is a rather atypical semantic component of the instances under examination. Instead, *alig van* typically expresses a low degree – there is hardly any time, strength, and means – or a state that is difficult to maintain.

- (8) *Magam a nagy katarusba csak alig vagyok*
 myself the great cough-INE only barely be-1SG
 ‘Because of the heavy cough, I myself am only barely [alive].’, Bark II. 512, 1720).

Simonyi (1881) argues that *alighogy* ‘almost’ also occurs with the meaning of *aligha* and is easier to explain. The problem with this explanation is, that, according to Middle Hungarian data, *alighogy* indicates averted relationship, suggesting that a given past event was imminent but ultimately did not occur (see. Gugán and Varga 2025), as in (9):

- (9) *alighogy pofán nem csapám*
 barely-that face.on not slap.PST-1SG
 ‘I almost slapped him on the face’ (Székely László, 1763–1771)

From a structural point of view, the two constructions share the presence of a negative modifier *nem* and the VPFX–NEG–V order. Additionally, *alighogy* is also a compound where the constituents *alig* ‘hardly, barely’ and *hogy* ‘that’ are syntactically unrelated. Historical studies on the Old Hungarian period claim that *hogy* ‘that’ tends to collocate with other conjunctions and other elements (e.g. *merthogy* ‘because’, *úgyhogy* ‘so, therefore’) due to its frequency, semantic bleaching, and functional diversity (Juhász 1992: 789, Haader 1995: 650–651).

However, in the case of *alighogy*, the original bi-clausal structure can also be attested. These tend to contain a main verb indicating that an undesirable event has been avoided. This also confirms that the component *hogy* is the conjunction of complement clauses, not a modal determiner ‘how’ since the occurrences are subordinated specifying the object: the event which was (almost not) avoided. Consider (10).

- (10) *hirtelen ránk jöve nagy szélvész:*
 suddenly on-us come-PST.3SG great windstorm
csak alig szaladhattunk egy kis szakadékban,
 only barely run-MOD-1PL a small ravine-INE
hogy mind oda nem veszünk
 that all away not sweep-PST-1PL
 ‘Suddenly, a great storm came upon us: we could **barely** run into a small ravine to avoid being swept away.’ (Thököly Imre, 1693)

Returning to the plausible bi-clausal origins of *aligha* as a compound, Simonyi (1881) also suggests that the main verb presumably was *hisz* ‘believe’. According to Middle Hungarian data, the subordinate clause of *alig hisz* ‘barely believes’ occurs both with *hogy* and without any conjunction, and, only very rarely with *ha* ‘if’, as in (11) and (12):

- (11) *kit alig hiszek, ha*
 which-ACC barely believe-1SG if
megint más útra nem kelletik uramnak fáradni
 again another path-onto not must man-DAT take-trouble-INF
 ‘I believe my man probably will have to take another path.’ (letter: MHöLev Nr. 138., 1607);

- (12) *én alig hiszem,*
 I barely believe-1SG
ha Gáspár uram nem morgotta el.
 if Gáspár Mr. not grumble-PST-3SG VPFX
 ‘I believe Mr. Gáspár probably grumbled about it.’ (correspondence of Mihály Teleki, Vol. 2. Nr. 89., 1661).

The subordinate clause can be both negative or affirmative, but in the latter case the verb appears in the subjunctive (imperative or conditional) mood, a feature that may indicate the event is non-veridical.

The construction *nem hisz* ‘does not believe’ is also attested with a negative subordinate clause containing an imperative or conditional verb, yielding double negation and thus expressing high probability, similar to *aligha*, though the conjunction is typically *hogy* or *zero*. The further investigation of the original bi-clausal structure is beyond the scope of this paper in the absence of data.

3. Data and method

Turning to the present study, data were collected from the available digital historical corpora of Hungarian. The earliest written sources come from the Old Hungarian Period (before the first decades of the 16th century). *Aligha* is neither attested in contemporary codices nor in smaller secular records (cf. Old Hungarian Corpus, Simon and Sass 2012).

Considering the Middle Hungarian period (between the years 1526 and 1772), two digital databases are available. The Old and Middle Hungarian Corpus of informal language use (here abbreviated as TMK, for the description, see Novák et al. 2018) contains private letters and records of witch trials, especially witness depositions (consisting of more than 1.1 million tokens). *Aligha* was found with 34 instances (that is about 3 per 100,000 tokens), in three cases it was followed by *nem* (thus *alighanem* ‘probably’). The Corpus of Middle Hungarian memoirs and dramas (here abbreviated as KED, consisting of more than 200,000 tokens) had one instance of *aligha*. Since some of the memoirs are sampled in the database, the full length texts were processed by using AntConc (Anthony 2023) resulting in three more occurrences. Contemporary letters, dramas, and religious prose out of corpora were also processed with AntConc or fully manually (amounting to more than 3 million words of texts). In total, 194 instances were found from the Middle Hungarian period.

Regarding the Early Modern and Modern Hungarian Period, the Hungarian Historical Corpus (here abbreviated as MTSZ, for the description, see Sass 2017) was consulted. Since the texts are not normalized, word forms and character sequences were queried as well (“*alig ha*” and “*alig, ha*”, in total, 55 instances). In the case of the forms *aligha* and *alighanem*, samples were taken: 600 instances of *aligha*, from the earliest occurrences, i.e. 1793 to 1959 and 244 data points of the fused form *alighanem*, because those fall before 1959. In the case of *aligha*, the sample covers about half of the instances available from MTSZ, and in the case of *alighanem*, more than a third of the occurrences.

The aforementioned Hungarian Gigaword Corpus was consulted in order to find out whether the form *aligha* [...] *nem* is still attested. About fifty occurrences were found, all from archaic literature and historical sources, not particularly representing present-day language use.

Data were tagged manually according to the following aspects and research questions:

- form of the modifier (*aligha*, *aligha* [...] *nem*, or *alighanem*);⁴
- function: likelihood or improbability;
- clause: affirmative or negative;
- the scope of *aligha*
- word order of the verb with a verbal prefix;
- distribution of the data across sub-periods

From the register point of view, the majority of the Middle Hungarian data comes from letters, while Modern Hungarian data mostly come from fictional prose.

4. Corpus study

First, the variation and change of the functions of *aligha* and *alighanem* are examined. In other words, the ratios of how frequently *aligha* express improbability and likelihood in Middle Hungarian, in Early Modern and Modern Hungarian are determined and how this feature affects some of the structural characteristics of the clauses in which *aligha* and *alighanem* appear. Then, the study focuses on the word order of the predicates with verbal prefix and whether it is affected by the functional change of *aligha*.

4.1. Function and variation in the Middle Hungarian period

The distribution of *aligha* according to function during the Middle Hungarian period is observed as follows. The modifier under scrutiny predominantly indicates that the occurrence or state of affairs is very probable (90%), mostly expressed by the form *aligha* (see Figure 1).

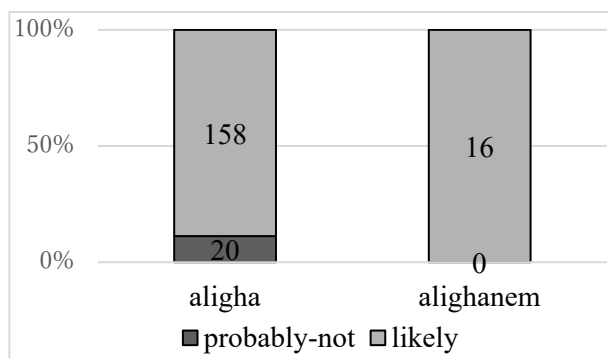


Figure 1: The distribution of *aligha* according to meaning between the 16th and 18th century (in total: 194 instances, sources: TMK, KED and manual data-collection)

Structure and function correlate with each other: improbability is always expressed in affirmative sentences (20 instances, 100%), while probability is expressed almost exclusively in negative sentences (173 instances, 99%). This feature can be regarded as a case of double negation, since *aligha* in itself belongs to the negative pole, and the addition of the negative particle triggers the highly probable (positive) interpretation (Gugán 2021). Note that the clause may contain not the negative particle *nem* ‘no, not’ but the negative form of the substantive verb, *nincs* ‘not exist’. Consider (13).

⁴ If *aligha* and *nem* occur directly adjacent, they are considered as *alighanem*, even if they are written separately in the historical sources since the published volumes followed different orthographic patterns.

- (13) *Besenyei István szegény Csomaközi is aligha oda*
 Besenyei István poor Csomaközi also unlikely toward
nincsen
 not.exist-3SG
 ‘István Besenyei and poor Csomaközi are also probably gone.’ (correspondence of Mihály Teleki, Vol. 6. Nr. 393., 1674)

This verbal form can be attested in the 20th century and even recently as well, as it is shown in (14).

- (14) *Ebben aligha nincs valami igazad*
 this-INE unlikely not.exist-3SG some truth-POSS-2SG
 ‘In this, you are probably right.’ (MTSZ, 1906–1907)

A very similar clause can be found in a radio interview: *aligha nincs is valami igaza* ‘he is also probably right’, MNSZ2, 2012).

As a result of the fusion of *aligha* and *nem* into *alighanem*, the meaning of the negative particle becomes schematized. The broader context of (15) is as follows: *Én egy gonosz, Istentől elrugaszkodott embert keresek, a korhelyeknek céhmesterét.* ‘I am looking for a wicked man, estranged from God, a master of drunkards.’

- (15) *Alighanem engemet, mert ez az én titulusom*
 probably me because this the my title
 ‘Probably me, because that’s my title.’ (Jezs.2.11_Botfalvai, second half of the 18th century)

However, the distinction is not always clear even in the 19th century. In (16), both the double negation and the actual negation of the proposition can be interpreted.

- (16) *Pedig az őz aligha nem nyomott többet,*
 although the deer hardly not weigh-PST-3SG more-ACC
mint lövője
 than hunter-POSS-3SG
 ‘Yet the deer hardly weighed no more than its hunter.’ (MTSZ, 1854-1855)

Both the variation of *aligha* [...] *nem* and *alighanem* and the fusion of the latter variant raise the question of the scope of the modifiers. The forms *aligha* ‘unlikely’ and *aligha* [...] *nem* ‘probably’ usually modify the whole event or proposition, whereas *alighanem* is more likely to function as a focalizer (Claridge 2020). The difference comes from the word order. In the former case, *aligha* itself is not always adjacent to the verb. Yet, in many cases, in 85 instances out of the 194 Middle Hungarian data points, they are still close to each other, only the verbal prefix is between them. Consider (17).

- (17) *aligha meg nem várom*
 unlikely VPFX not wait-1SG-DEF
 ‘I [will] probably wait for it.’ (TMK, Kár. 181., 1719)

Nevertheless, the negative particle is regularly located at a distance from *aligha* (47 instances), as it is illustrated in (18).

- (18) *aligha az erdélyi kereskedők is a Lipcsei*
 unlikely the Transylvanian merchant-PL too the Leipzig
húsvéti vásártól el nem rekesztetnek
 Easter Fair-ABL VPFX not exclude-PASS-3PL
 ‘Probably the Transylvanian merchants will be excluded from the Leipzig Easter Fair.’ (TMK, Peregr1. 13.)

Still, the negative particle modifies the verbal predicate and the two components *aligha* and *nem* enclose the whole proposition thereby emphasizing the probability of the entire event or state. In the case of *alighanem*, where *aligha* and *nem* are directly adjacent or even fused, the modifier is regularly followed by other constituents beyond the verb, and its scope may have been reduced as a result of the change in the position of the negative particle. See (19).

- (19) *alighanem most kezdődik itt az új komédia*
 probably now start-3SG here the new comedy
 ‘The new comedy is probably starting here now.’ (letter of Miklós Bercsényi, 1704–1712)

Considering the ratios of the different uses across subperiods, the meaning ‘likely to occur’ predominates until the end of the period (mostly expressed by *aligha* [...] *nem*). Although there are only a few data points from the final subperiod, the distribution follows the patterns of the preceding decades (see Figure 2).

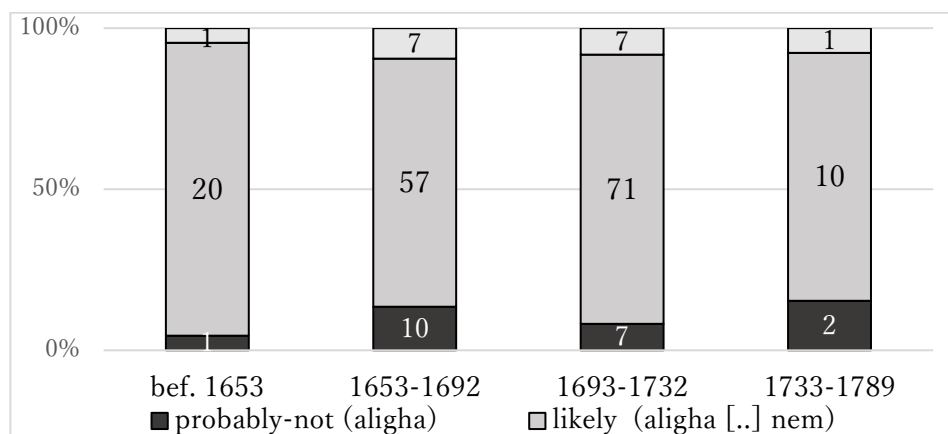


Figure 2: The ratios of the different uses of *aligha* between the 16th and the 18th century (in total: 194 instances, sources: TMK, KED and manual data-collection)

4.2. Early Modern and Modern Hungarian: a time for change

Turning to the Early Modern and Modern Hungarian periods, a change can be observed in the ratios of the different uses and forms according to data from the Hungarian Historical Corpus. *Aligha* still varies in expressing improbability and probability as well as the structure regarding the presence and the absence of the negative particle. Nevertheless, from the mid-1800s onward, *aligha* expressing improbability becomes

more frequent than uses with positive polarity covering the 60% or more of the occurrences in each subperiod (Figure 3).

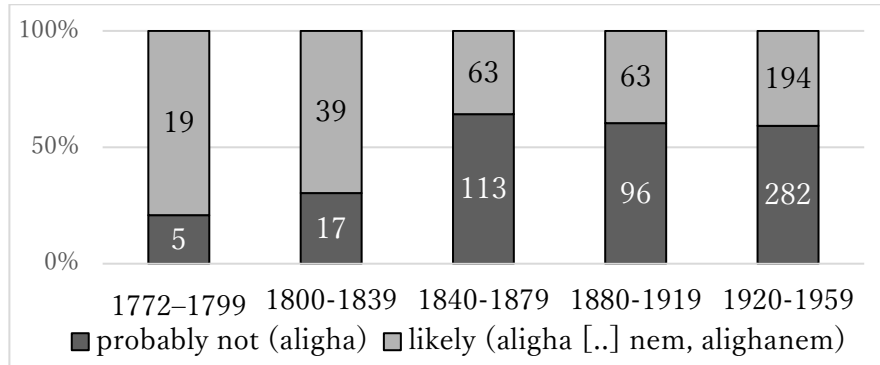


Figure 3: Change of the ratios of the modifiers *aligha*, *aligha*([...]*nem*) expressing high and low probability (total: 899 instances, corpus: MTSZ)

Over the same decades, *alighanem* begins to exceed the frequency of the unfused *aligha* [...] *nem* (Figure 4).

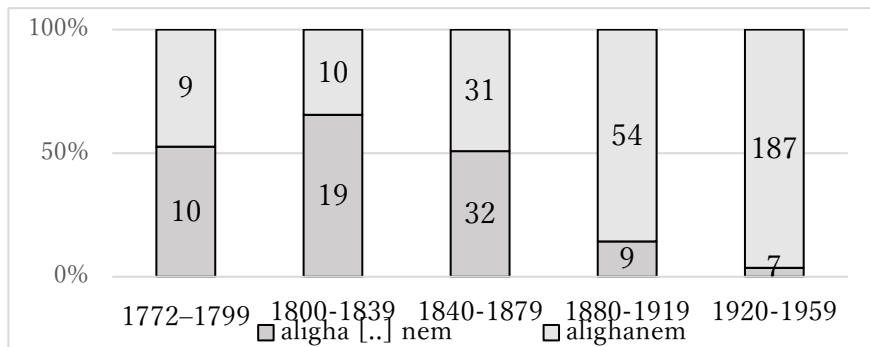


Figure 4: Change of the ratios of the functional variants *aligha* [...] *nem* and *alighanem* expressing high probability (total: 378 instances, corpus: MTSZ)

Dividing the raw frequencies into decades reveals that *aligha* meaning ‘probably’ in negative sentences remains relatively stable in frequency over time, while the affirmative uses indicating improbability have been spreading slowly but steadily since the 1830s and 1840s. The form *alighanem* also spreads, though much more slowly (see Figure 5).⁵

⁵ The raw frequencies could not have been normalized since the token numbers of the subperiods are currently not available and cannot be retrieved from the Hungarian Historical Corpus.

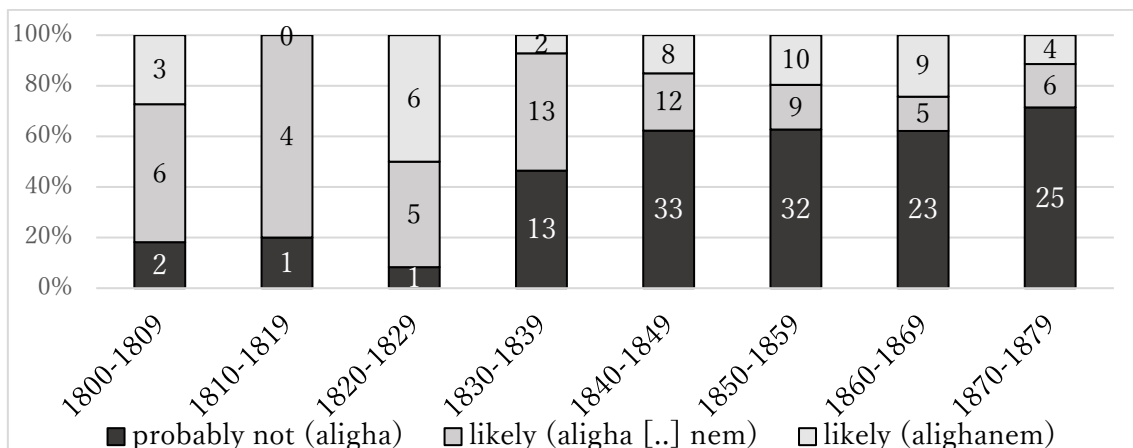


Figure 5: The changing ratios of the variants of *aligha* during decades (between 1800 and 1879, corpus: MTSZ)

Note, that during this period the scope of the modifiers became more varied as well. *Alighanem* and *aligha* meaning ‘unlikely’ can be observed as responsive marker as well – the former has instances by the end of the 18th century, while the latter is attested from the mid-19th century onward, as it is illustrated in (20) and (21), respectively.

- (20) [BERÉNYI]: *Nem ismérsz tehát szavamról?*
 [BERÉNYI]: No know-2SG then word-1SG-DEL
 [JOLÁNTA]: *Alighanem. – De mégsem.*
 [JOLÁNTA]: Probably. But still-not.
 ‘[BERÉNYI]: So you don’t know me from my word? [JOLÁNTA]: Probably (I do). – But still not.’ (MTSZ, #7087019, 1795)

- (21) *Sándor olvasta? – Azt nem tudom. De aligha.*
 [Sándor] read-PST-3SG that-ACC no know-1SG But unlikely
 ‘Has Sándor read it? – I don’t know. But unlikely.’ (MTSZ, #9163012, 1917)

The responsive uses highlight that the constituents of the modifier are lexicalized (Simonyi 1881). As an interim summary, corpus data indicate that during the mid-19th century there was a shift in the preference from the predominant use of *aligha* [...] *nem* expressing high (positive) probability toward *aligha* expressing improbability. This resulted in the increase of a unique type of modification associated with the negative pole. The lexicalization and the growing frequency of the variant *alighanem* had an influence on the scope of the modifier as well, giving space to the focalizer and responsive uses as well. The syntactic aspect of the word order and how is it related to the change described above is discussed in the following section.

5. Word order patterns: variation and change

It has been argued (Kugler 2009, 2013, Gugán 2021) that the word order of *aligha*, which used to be VPFX-NEG-V, reflects the typical pattern of conditional clauses inherited from the second component of the compound *ha* ‘if’. Considering the Middle Hungarian data, when *aligha* indicates high probability with a negated predicate and a verbal prefix is present in the structure (92 instances), the word order is predominantly VPFX-NEG-V (85 instances, 92%, see Examples 3, 17, and 18).

In a few cases (3 instances), the observed word order displays an inversion of the neutral VPFX–V order, namely: NEG–V–VPFX. The inversion is presumably related to the scope of the modifier: in these cases, *aligha* does not modify the entire proposition or clause referring to the probability of the whole event or state of affairs but rather a specific circumstance. In other words, the issue is not whether the event actually occurs, but whether one of its constituents or circumstances is affected by *aligha* functioning as a focalizer. Example (22) does not raise the question of whether the mentioned people leave their territory but the actual place is the subject of (un)certainty.

- (22) *aligha Verestoronyán nem járnak ki*
 unlikely Verestorony-ADE not go-3PL VPFX
 ‘They probably go out by Verestorony.’ [= a gorge in the South Carpathians]
 (correspondence of Mihály Teleki Vol. 8. Nr. 546., 1679)

Also, four instances display neutral VPFX–V order, due to temporal (e.g. *fog* ‘will’) or modal verb predicates (*kell* ‘must’) where the non-finite form of the verb has a verbal prefix, consider (23).

- (23) *aligha magamnak nem kell bemennem*
 unlikely myself.DAT not must VPFX.go-INF-1SG
 ‘Probably I myself must go in.’ (correspondence of Mihály Teleki, Vol. 4., Nr. 86., 1667)

By contrast, when *aligha* does not indicate probability and its host sentence is affirmative, the order is exclusively VPFX–Verb (10 instances), as illustrated in (24) and (25).

- (24) *Az ország gyűlésének én az elejére*
 the country assembly-DAT I the beginning-SUB
aligha bemehetek
 unlikely VPFX.go-MOD-1SG
 ‘I can unlikely go to the beginning of the country’s assembly.’ (correspondence of Mihály Teleki, Vol. 5. Nr. 16, 1670)

- (25) *amint arányzom, aligha a hadat elkerülhetjük*
 as suppose-1SG unlikely the army-ACC VPFX.avoid-MOD-1PL
 ‘As I suppose we can hardly avoid the army/fight.’ (letter: Bálint Lépes, Litpol., Nr. 138, 1614)

The observed pattern in (24) and (25) is ungrammatical in Modern Hungarian. The question arises when the change took place in this respect and whether it is affected by the spread of the frequency of the uses of *aligha* expressing improbability. According to Early Modern and Modern Hungarian data, *aligha* [...] *nem* consistently preserved the VPFX–NEG–V order even in the 20th century when this pattern had become infrequent. The broader context for (26) is the following: *Hitemre, nem tudom, mint végződött volna a dolog* ‘Upon my faith, I don’t know how it would be ended’.

- (26) *de valamelyikünk aligha le nem nyúzta volna*
 but one-of-us unlikely VPFX not skin-PST-3SG be-COND-3SG

a másik bőrét
 the other skin-POSS-3SG
 ‘but one of us probably would pull the other’s skin off.’ (MTSZ, #9275031, 1950)

In the case of *aligha* meaning ‘unlikely’, however, a change can be observed along with the spread of its frequency. Until the mid-19th century, the VPFX–V order was neutral as it is illustrated by (27) and (28).

(27) *ha kérdi tőlem, mi a szerelem,*
 if ask-3SG from-me what the love
aligha azt megmondhatom
 hardly that-ACC VPFX.say-MOD-1SG
 ‘If (s)he asks me what love is, I can hardly say.’ (MTSZ, #7165014, 1794)

(28) *kedvezőtlenebb kép aligha feltárul*
 unfavorable-CMP picture unlikely VPFX.reveal-3SG
a magyar történetben
 the Hungarian history-INE
 ‘a more unfavorable picture in Hungarian history can hardly be revealed.’
 (MTSZ, #8505001, 1870)

Moreover, other modifiers also preceded the verb during the Early Modern Hungarian period which occur in post-verbal position in standard Modern Hungarian. Consider *készen lehetek* (29) and *bérbe adta* (30).

(29) *hogy aligha készen lehetek a kiszabott időre,*
 that unlikely ready be-MOD-1SG the determine-PRS time
 ‘that I can unlikely be ready by the deadline.’ (MTSZ, # 8173059 1854)

(30) *Más fejdelem aligha bérbe adta volna*
 other prince unlikely rent-ILL give-PST-3SG be-COND-3SG
 ‘Any other prince would unlikely have rented out [part of his palace to the merchants].’ (MTSZ, #8235013, 1858)

As Figure 7 suggests, the inverse order that is typical for non-neutral sentences became common only from the last decades of the 19th century as it is illustrated by (31).

(31) *hogy aligha cserélhetnők fel egymást valaki jobbal*
 that unlikely replace-MOD-1PL VPFX each-other someone better
 ‘that we could unlikely replace each other with someone better.’ (MTSZ, #9290042 1913)

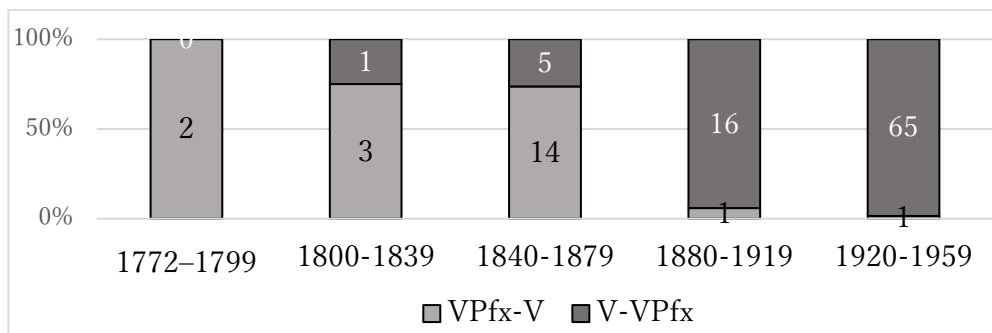


Figure 7: Changing ratios of the neutral and non-neutral word order in clauses involving *aligha* (raw numbers, corpus: MTSZ)

However, an example from the paper of the linguist Zsigmond Simonyi (1913: 92) suggests ongoing variation in this respect as well, at least in the scientific register.

- (32) *nevezetes jelenség, mely a magyaron kívül*
 specific phenomenon which the Hungarian-on other-than
más nyelvekben aligha előfordul
 other language-PL-INE barely VPFX.appear-3SG
 ‘it is a specific phenomenon which unlikely appears in other languages than in Hungarian.’ (Simonyi, 1913)

Both the aforementioned utterance and the low frequencies available from MTSZ suggest that the question of the actual change requires further investigation and additional data collection from the 19th and 20th century. This could be a succeeding step of this research.

Nevertheless, the studied corpus data indicates the shift in the preference of *aligha* expressing improbability apparently involved not only declining use of double negation and the spread of *alighanem* but change in word order the spread of inversion of the verbal prefix as well.

6. Discussion

The results of the corpus study highlight some issues to be discussed. One of them is the syntactically marked structure of *aligha* [...] *nem* and why that variant is more generally used in Middle Hungarian. In terms of function, there was basically no functional variant that directly expressed high probability during that time. Other elements indicated more general epistemic uncertainty, such as *talán* ‘perhaps’, *netalán* ‘maybe’. *Valószínű* ‘probable’ and *valószínűleg* ‘probably’ can only be attested from the 19th century onward, similarly to *feltehetőleg*, *gyaníthatóan*, and *elképzelhető* which also indicate probability, marking the positive pole by default (see MTSZ and ÚESz.).

In terms of the structural features, several verbal phrases triggered logically not interpreted, in other words expletive negation as well as double negation in subordinate clauses, especially, before the 19th century (such as *tilt* ‘forbid’, *tagad* ‘deny’, *fél* ‘be afraid’, *kétkelkedik* ‘doubt’, Gugán 2022). The use of *aligha* [...] *nem* might also be related to this tendency – as is was already mentioned, the averted use of *alighogy* ‘almost’ involved expletive negation. This assumption requires further investigation.

Another issue to discuss is the explanation for the change in word order in the case of *aligha*, expressing unlikelihood. According to Kugler (2009: 146–147), the VPFX–NEG–V word order shows the word order of the conditional clauses and it is inherited

from the component *ha*. She argues that once the functional composition had become fixed, meaning that *alig* and *ha* did appear not only at the boundary of a clause but also at the beginning or within a clause, its order was replaced by the word order pattern characteristic of the restrictive modifier *alig*, that is inversion.

As it was mentioned in Section 2, there are no original bi-clausal utterances attested while *aligha* occurs at least as a collocation if not as a compound as early as the 16th century. It is also unusual that an element wedges itself between the constituents. According to data from TMK, KED, and manual sources, in total based on more than 900 data points of *alig*, there is only one exceptional example regarding how likely is that the modified event will actually happen. Consider (33).

- (33) *Alig Édesem ha januáriust is itt nem töltöm*
 hardly sweet-POSS-2SG if January-ACC too here not spend-1SG-DEF
 ‘I probably spend January here too, my dear.’ (TMK, Kár. 23, 1704)

Also, there is an example where not the event itself but the time of the actuation is questioned. Note, that in (34) the word order is neutral (VPFX–V):

- (34) *A szegek csak alig hétfőn estére*
 the nail-PL just hardly Monday-SUP evening-SUB
ha elkészülhetnek
 if VPFX.be-made-MOD-3PL
 ‘The nails can only hardly be made by Monday evening, if (not later).’
 (Correspondence of Mihály Teleki, Vol.1. Nr. 155., 1658)

From at least the 16th century onward, *aligha* tends to occur clause-internally as well, indicating that the compound is functionally fixed. By all means, it exhibits VPFX–NEG–V order for example in (35) and (36).

- (35) *Álljatok meg, mert aligha meg nem*
 stop-SUBJ.PL2 VPFX because unlikely VPFX no
kelletik változtatnunk akaratumkat
 must-PASS-3SG change-INF-1PL will-POSS-1PL
 ‘Stop, because we probably must change our mind.’ (Aesop’s Fables translated by Gábor Pesti, 1535)

- (36) *s a gyűlésbe is aligha el nem megyek*
 and the meeting.ILL too unlikely VPFX no go1SG
 ‘and I will probably attend the assembly.’ (letter of Miklós Bethlen, 1667)

As it has been shown in Section 5, these clauses displayed neutral VPFX–V order between the 16th and the 18th century even in cases when *aligha* occurs within the sentence, such as (37).

- (37) *kinek különböztető színét aligha*
 who-DAT distinctive color-ACC unlikely
valamely virág felülhaladja
 some flower surpass-3SG

‘whose distinctive color is unlikely surpassed by any other flower.’ (Lippay, 1641)

Historical corpus data suggests that the presumably inherited order changes lately, despite the fact that *aligha* occurred clause internally from the 16th century onward, indicating that it is functionally fixed. Inversion in affirmative clauses became widespread only in the last decades of the 19th century (Section 5).

Klemm (1928–1942) argues that the inversion might have been influenced by adverbials with similar meanings, such as *nehezen* ‘difficultly’ and *bajosan* ‘problematically’. Those however, triggered inversion in Middle Hungarian period as well, and they infrequently expressed low probability even later on. Also, structures involving *nehezen* and *bajosan* exhibited neither expletive, nor double negation.

Note that the first constituent, *alig* itself, mostly triggered inversion from the earliest instances onward, even when expressing near-simultaneity. The exceptional environments include the aforementioned avertive construction, in which the word order is triggered by the expletive negation (Gugán 2022). Another specific environment is *alig hisz* ‘hardly believes’ followed by a subordinate clause, which is structurally different from the occurrences of *aligha*. Nevertheless, when the scope of *alig hisz* is sentential and the subordinate clause is affirmative, the word order is neutral, as in (38), similar to the variants involving *aligha* (see also Examples (24) and (25)).

- (38) *alig hiszem, én az gyűlésre bemehessek*
hardly believe-1SG I the assembly-SUB VPFX.go-MOD-SUBJ-1SG
‘I hardly believe (that) I can go to the assembly’ (Correspondence of Mihály Teleki, Vol. 5. Nr. 35, 1670)

Presumably, the change of the word order involving *aligha* is related to a more general development affecting negation. There was a spread of inversion in the negative sentences during the 19th century. In the preceding period, the VPFX–NEG–V word order was common and pragmatically neutral; therefore, the change can be regarded as a specific instance of the Jespersen cycle (see Gugán 2017, 2021, 2025). The consequences of this change require further investigation of expletive negation and double negation, including the environments of *aligha* and other epistemic and inferential modifiers.

7. Summary

The present study investigated the history of the Hungarian modifier *aligha* ‘unlikely’ and its variant *alighanem* ‘probably’. *Aligha* is a specific modifier, as it indicates negative polarity – the negation of the modified proposition – by default. Considering the plausible bi-clausal origin of the constituents *alig* ‘hardly, barely’ and *ha* ‘if’, the provided historical data analysis suggests that the predicate of the main clause may originally have been a verb other than *van* ‘is’, *lesz* ‘will be’, or *lehet* ‘may be, can be’. One of the most plausible candidates is *hisz* ‘believe’. Although the construction *alig hisz* co-occurring with the conjunction *ha* ‘if’ is attested, it is extremely rare. The latter constituent, the conditional conjunction *ha* may have been motivated by at least two factors. First, *alig* also occurred with the default conjunction of the complement clauses *hogy* ‘that’, however, in that case, the structure *alig [...] hogy ~ alighogy* served a different function, namely the avertive. Second, *ha* ‘if’ had a potential to indicate that the clause did not

describe a fact but an event assumed with strong conviction. Nonetheless, this question remains open due to the lack of sufficient historical data.

The extensive historical analysis confirmed a change in the usage during the 19th century, resulting in the stronger preference of *aligha* expressing improbability, whereas in earlier periods, *aligha* in collocation with the negative particle *nem* ‘no, not’ was used to indicate high probability. Moreover, the shift in the preference presumably affected the syntactic features of the modifier as well. On the one hand, these features include the fusion and schematization of the negative particle when expressing high probability, yielding *alighanem*. On the other hand, *aligha* previously exhibited neutral word order (VPFX–V) when expressing improbability, which was gradually replaced by non-neutral (V–VPFX) order. The V–VPFX pattern was typical of the first constituent, *alig* ‘barely’ as early as the Old Hungarian period. However, *aligha* displayed a (VPFX–V) order similar to the collocation of *alig hisz* ‘barely believes’ until the late 19th century.

At the current state of the research, it can be assumed that the motivation of the observed change may be related to a more general development affecting negation and word order. Nonetheless, the details require a further investigation of the language use in the late 19th century, extending the analysis to main clauses with mental verbs, such as *hisz* ‘believe’, *gondol* ‘think’, as well as to other earlier patterns of double and expletive negation.

Sources

- Aesop’s Fables = *Ezópusz fabulái* [Aesop’s Fables]. Translated by Gábor Pesti. 1535.
- Bark. II. = Kovács, Ágnes and Fogarassy Zoltán (2017). *Barkóczy Krisztina levelei férjéhez, Károlyi Sándorhoz. Második kötet (1712–1724)*. [Correspondence of Krisztina Barkóczy. Vol. 2.] Debrecen.
- Bethlen Miklós = Jankovics, József (ed.) (1987). *Bethlen Miklós levelei* [Correspondence of Miklós Bethlen]. Budapest: Akadémiai, Vol. 1. 1657–1698; Vol. 2. 1699–1716. <https://mek.oszk.hu/21600/21614/pdf/>
- Correspondence of Mihály Teleki = Sámuel Gergely (ed.) (1905–1926). *Teleki Mihály levelezése* [Correspondence of Mihály Teleki.] 1656–1679. Vol. 1–8. Budapest: Athenaeum-Királyi Magyar Egyetemi Nyomda.
- KED = Katalin Gugán, Enikő Bácsi, Adrienne Dömötör, László Horváth, Zsuzsanna Mohay, Mónika Varga, and Bálint Sass (2023). *Középmagyar emlékirat- és drámakorpusz*. [Middle Hungarian Corpus of Dramas and Memoirs] <https://ked.nytud.hu/#concordance>
- Lippay = Lippay, János (1664). *Posonyi Kert* [The garden of Pozsony].
- Litpol = Jankovics, József (1981). *Literátor-politikusok levelei Jenei Ferenc gyűjtéséből (1566–1623)* [Letters of literate politicians]. Budapest–Szeged: MTA–JATE.
- MHöLev = Magyar Hölgyek = Deák, Farkas (1879). *Magyar Hölgyek levelei 1515–1711*. [Letters of Hungarian Ladies.] Budapest: MTA <https://digitalia.lib.pte.hu/hu/pub/deak-farkas-magyar-holgyek-levelei-bp-1879-4655>
- MNSZ2 = Magyar Nemzeti Szövegtár [Hungarian Gigaword Corpus] <https://mnsz2-ud.nytud.hu/#open>
- MTSZ = *Magyar Történeti Szövegtár* [Hungarian Historical Corpus] <http://clara.nytud.hu/mtsz/>
- ÓMK = *Ómagyar Korpusz* [Old Hungarian Corpus] http://clara.nytud.hu/omagyar/run.cgi/first_form

Székely László = Fehér, Andrea (ed.) (2019). *Gróf Székely László önéletírása* [Memoir of László Székely] 1763–1772. Budapest–Kolozsvár: MTA BTK–EME.

Thököly Imre = Nagy, Iván (ed.) (1863). *Késmárki Tököly Imre naplója 1693–1694 évekből* [Diary of Imre Thököly.] Budapest: Akadémiai.
<https://mek.oszk.hu/05200/05239/05239.pdf>

TMK = Történeti magánéleti korpusz [The Old and Middle Hungarian Corpus of informal language use] <http://tmk.nytud.hu>

References

- AntConc = Anthony, Laurence (2023). AntConc (Version 4.2.4) [Computer Software]. Tokyo, Japan: Waseda University. <https://www.laurenceanthony.net/software>.
- Claridge, Claudia (2020). Epistemic adverbs in the Old Bailey corpus. In Ewa Johnson and Tove Larsson (eds.) *Voices Past and Present – Studies of Involved, Speech-related and Spoken Texts. In honor of Merja Kytö*. 133–151. Amsterdam: Benjamins.
- É. Kiss, Katalin (2006). Jól megoldottuk, rosszul oldottuk meg? [Explaining the word order of inclusive and exclusive expressions.] *Magyar Nyelv* 442–458.
- Gugán, Katalin (2017). A magyar tagadó mondatok szórendje és a konstansráta-hipotézis. In Katalin É. Kiss, Attila Hegedűs, and Lilla Pintér (eds.) *Nyelvelmélet és diakronia* [Theoretical Linguistics and Diachrony] 3, 91–110. Budapest–Piliscsaba: PPKE BTK.
- Gugán, Katalin (2021). A Jespersen-ciklus előszobájában: a mondattagadás változatai a középmagyarban [In the antechamber of the Jespersen Cycle: varieties of sentence negation in Middle Hungarian]. In Adrienne Dömötör, Katalin Gugán, and Mónika Varga (eds.) *Versengő szerkezetek a középmagyar kor nyelvén* [Competing Variants in Middle Hungarian]. 129–153. Budapest: Akadémiai.
- Gugán, Katalin (2022). „Félő, az Isten is el ne hagyjon” – Pleonasztikus tagadás a középmagyarban [Expletive negation in Middle Hungarian]. In Tamás Forgács, Miklós Németh, and Balázs Sinkovics (eds.) *A nyelvtörténeti kutatások újabb eredményei* [Recent Results of Historical Linguistic Research] XI, 95–112. Szeged: SZTE BTK.
- Gugán, Katalin (2025). A tagadás változatai a regiszterek tükrében [Variants of negation in the mirror of registers]. (to appear: *Regiszterfüggő változatok a középmagyarban* [Variation in Middle Hungarian: A Register Perspective])
- Gugán, Katalin and Varga Mónika (2025). Avertives in Middle Hungarian: Inherent variation and expletive negation. (Conference presentation: *Avertives in European Languages*)
- Haader, Lea (1995). Az alárendelő mondatok [Subordinate clauses]. In Loránd Benkő, (ed.) *A magyar nyelv történeti nyelvtana II/2* [Historical Grammar of Hungarian] 506–664. Budapest: Akadémiai.
- Juhász, Dezső (1991). A kötőszók. [Conjunctions] In Loránd Benkő (ed.) *A magyar nyelv történeti nyelvtana II/1* [Historical Grammar of Hungarian] 772–814. Budapest: Akadémiai.
- Klemm, Antal (1928–1942). *Magyar történeti mondattan* [Historical Syntax of Hungarian]. Budapest: MTA.
- Kugler, Nóra (2009). A módosítószók funkcionális megközelítése [Functional approach of modifiers] In Borbála Keszler and Szilárd Tátrai (eds.) *Diskurzus a*

- grammatikában – grammatika a diskurzusban* [Discourse in Grammar – Grammar in Discourse], 139–148. Budapest: Tinta.
- Kugler, Nóra (2013). The negativity of Hungarian modal adverbs (*aligha, bajosan, nehezen*). In Ilona Koutny and Piotr Nowak (eds.) *Language. Communication. Information* 8, 9–19. Poznań.
- Kugler, Nóra (2014). Az episztemikus (főmondati) predikátumok és a nyelvi polaritás [Epistemic predicates in the main clause and polarity]. In Ferenc Havas, Katalin Horváth, Nóra Kugler, and Zsuzsa Vladár (eds.) *Nyelvben a világ. Tanulmányok Ladányi Mária tiszteletére* [The World in Language. Papers in honor of Mária Ladányi]. 55–65. Budapest: Tinta.
- Kuteva, Tania, Bernd Heine, Bo Hong, Haiping Long, Heiko Narrog and Seongha Rhee. (2019). *World Lexicon of Grammaticalization*. 2nd edn. Cambridge: Cambridge University Press.
- D. Máta, Mária (2003). Szófajttörténet [History of word classes]. In Jenő Kiss and Ferenc Pusztai (eds.) *Magyar nyelvtörténet* [Diachronic Linguistics of Hungarian] 630–662. Budapest: Osiris.
- D. Máta, Mária (2011). *Szófajttörténet* [History of Word Classes]. Budapest: Argumentum.
- Novák, Attila, Katalin Gugán, Mónika Varga, and Adrienne Dömötör (2018). Creation of an annotated corpus of Old and Middle Hungarian court records and private correspondence. *Language Resources and Evaluation* 52, 1–28.
- Oravecz, Csaba, Tamás Váradi, and Bálint Sass (2014). The Hungarian Gigaword corpus. In Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asunción Moreno, Jan Odijk and Stelios Piperidis (eds.), LREC 2014 – *Ninth International Conference on Language Resources and Evaluation*, 1719–1753. Lisbon: European Language Resources Association.
- Sass, Bálint (2017). Keresés korpuszban: a kibővített Magyar történeti szövegtár új keresőfelülete. In Tamás Forgács, Miklós Németh, and Balázs Sinkovics (eds.) *A nyelvtörténeti kutatások újabb eredményei IX.* [Recent Results on Historical Linguistic Research.] 267–277. Szeged: SZTE.
- Simon, Eszter, and Bálint Sass (2012). Nyelvtechnológia és kulturális örökség, avagy korpuszépítés ómagyar kódexekből [Language technology and cultural heritage or corpus building from codices]. In Gábor Prószéky and Tamás Váradi (eds.) *Nyelvtechnológiai kutatások* [Language Technology Researches], 243–64. Budapest: Akadémiai.
- Simonyi, Zsigmond (1881–1883). *A magyar kötőszók, egyuttal az összetett mondat elmélete* [Hungarian Conjunctions and the Theory of Complex Sentences]. I–III. Budapest: MTA.
- Simonyi, Zsigmond (1913). *A jelzők mondattana* [The Syntax of Attributives]. Budapest: MTA.
- Tubau, Susagna and Ingham, Richard (2015). Some historical notes on English negation: *unethes, almost, and hardly*. In Fabienne Toupin and Brian Lowrey (eds.) *Studies in Linguistic Variation and Change. From Old to Middle English*. 215–226. Cambridge: Cambridge Scholars Publishing.
- ÚESz. = Gerstner, Károly et al. (2011–2022). *Új magyar etimológiai szótár* [New Etymological Dictionary of Hungarian.] Budapest: Nyelvtudományi Kutatóközpont. <https://uesz.nyttud.hu/index.html>