

DISCOURSE VARIATION ANALYSIS OF THE DISCOURSE MARKERS *NO* ~ *NA* IN PRESENT-DAY HUNGARIAN

Csilla Ilona Dér

*ELTE Research Centre for Linguistics & Károli Gáspár University of
the Reformed Church, HU*

csillader@gmail.com

Abstract: Historically derived from interjections, the Hungarian discourse markers *no* and *na* ‘now, so, well’ have long been treated as functionally equivalent variants, despite a reversal in their relative status: *no* was formerly dominant, whereas *na* was regarded as regional. This study examines the extent to which their present-day alternation is shaped by regional and register-based variation, as well as by functional specialization. Drawing on spontaneous conversations from the Hungarian Spontaneous Speech Corpus (BEA) and on spoken press data from the Hungarian Gigacorporus 2 (MNSz2), it analyzes their frequencies, positions, functions, and co-occurrence patterns within a discourse-variationist framework. The findings reveal only limited regional differentiation, apart from a higher frequency of *na* in the Subcarpathian subcorpus, and do not support a broad opposition between spoken *na* and written *no*; rather, *na* shows a clear concentration in spoken press. Although the two markers often occur in similar positions and have overlapping discourse functions, *na* is especially associated with transitions, whereas *no* is more often used in urging, disagreement, and argumentative continuation, suggesting partial functional differentiation.

Keywords: corpus linguistics, discourse marker, interjection, register variation

DOI: <https://doi.org/10.12697/jeful.2026.17.1.07>

1. Introduction

The aim of this study is to map and compare the present-day Hungarian – primarily spoken – usage of the discourse markers *no* and *na*, both originating from interjections meaning ‘now, so, well’ (New Etymological Dictionary of Hungarian, ÚESz., 2011–2025). I define discourse markers as interactives with metatextual functions (Heine 2023: 129), including text organization, the expression of speaker

attitudes, and/or the management of speaker–hearer interaction. Discourse particles are a subclass of discourse markers, consisting of short, word-level, and invariant units (cf. van Kemenade & Links 2020: 3).

The study is motivated by an apparent reversal in the historical status of the two forms. Until the twentieth century, *no* was the dominant variant, whereas *na* was described as a regional alternative. From the 1940s onwards, however, *na* began to overtake *no* in frequency (Dér 2026). This reversal is also reflected in the treatment of the two forms in comprehensive dictionaries: whereas the Dictionary of the Hungarian Language (CzF. 4/2, 1862: 848) classified *na* as a dialectal form, the Comprehensive Dictionary of the Hungarian Language (ÉrtSz., V., 1961: 92–94, 220) applied this label only to *no*, describing it, in one of its uses, as a regional word employed to urge animals, especially horses, while *na* received no dialectal marking. Despite this change in their relative frequency and lexicographical status, the two markers have generally been treated as functionally equivalent variants. The central question addressed in this study is whether the present-day coexistence and alternation of *no* and *na* merely reflect an incomplete process of replacement, or whether the two forms have come to differ in their regional, register-based, positional, or functional distribution.

The study investigates this question by comparing their frequency and regional distribution in the Hungarian Gigacorp2 (MNSz2) and by examining their discourse positions, functions, and co-occurrences with other discourse markers in spontaneous conversations and spoken press data.

The structure of the paper is as follows: Section 2 presents the relevant literature on *no* and *na* and outlines the theoretical framework applied. Section 3 details the databases and research methods used in the analysis of spontaneous conversations from the Hungarian Spontaneous Speech Database (BEA) and data from the Hungarian Gigaword Corp2 (MNSz2). The results are discussed in Section 4: Subsection 4.1 presents the analyses conducted on the BEA conversations, while Subsection 4.2 examines frequency patterns across MNSz2 registers and regional varieties. Section 4.3 provides a detailed analysis of the transcribed spoken press subcorpus: with Subsection 4.3.1 presenting the characteristics of *no* and Subsection 4.3.2 those of *na*, each based on a random sample of 500 occurrences. Finally, Section 5 summarizes the main findings and concludes the paper.

2. Literature review

2.1. The origin of *no*

The Hungarian forms *no* and *na*, both in their shape and function, appear to be related to the discourse markers (discourse particles) *NU* and *NA*, which are found in numerous Indo-European and some non-Indo-European languages (e.g. Hebrew) (Auer & Maschler 2016), and which also occur in various forms in Uralic languages related to Hungarian. An essential difference, however, is that while in those languages these elements are often considered loanwords or are traced back to adverbial forms with temporal meaning (e.g. *no(w)*, German *nu(n)*; see Auer & Maschler 2016: 6ff.), neither of these etymological possibilities is raised in the Hungarian literature, where the forms are generally regarded as (originating from) interjections (cf. ÚESz.). To complete the picture, it is worth noting that in the eleven languages discussed in the aforementioned volume (Russian, Polish, Yiddish, Romani, Hebrew, Finnish, Estonian, German, Dutch, Icelandic, and Swedish), the possibility of interjective origin is mentioned among the Slavic languages (e.g. Sawicki 2016: 83).

Interjections and discourse markers share certain characteristics, even though the latter constitute a functional rather than a lexical category, since their members originate from a variety of word classes and structural levels. In both cases, emotional expression is often discussed in terms of meaning, function, or stance marking.

Heine (2023: 7–9) characterizes both interjections and discourse markers as interactives, defined as invariable deictic forms – that is, forms anchored in the discourse situation – which are semantically, syntactically, and prosodically (to varying degrees) distinct from their surrounding text, cannot be negated, and cannot be used to form questions. Heine (2023: 7–9) mentions eleven such categories, among which both interjections and discourse markers can be found. He defines interjections as items that express a sudden change in the speaker's cognitive or emotional state and that can grammaticalize into discourse markers (Heine 2023: 185, 280). As the (direct) precursors of interjections, Heine (2023: 234) identifies vocatives, which serve to attract the listener's attention (to call or address someone). Hungarian etymological dictionaries offer a different type of account: they commonly describe interjections as onomatopoeic or expressive forms originating in vocal outbursts (cf. ÚESz.; this is also the etymological classification of *no*).

2.2. The development of *no* and *na* in Hungarian

Outside the lexicographical literature, *no* and *na* are mentioned briefly in the chapters on interjections and sentence structure in the Historical Grammar of Hungarian (Gallasy 1992: 842ff.; Balázs 1995: 71ff.), which record their existence in Old Hungarian and assume their early origin. Two corpus-linguistic studies have attempted to explore the discourse marker functions of *no* and *na* in more detail from a diachronic perspective (Erdei 2021; Dér 2026). Based on these, it can be stated that *no* was the first to develop of the two forms; according to etymological dictionaries, its earliest occurrence is found in the Schlägli Glosses (SchlGl. 77a, c. 1430) as *nomar* (*no* + *már* ‘then, well’), glossed there with the Latin form *eŷa* ‘ah ha, oh, well’ (Szamota 1894: 94). Both Latin *eya* ~ *eia* and its Hungarian equivalent allow for multiple interpretations (e.g. urging or expression of surprise), but in the absence of contextual evidence their precise value cannot be determined. In *The Battle of Szabács* (SzabV., c. 1476), which dates from a period very close to the glosses, it also appears together with another element, *azért*, that later developed into a discourse marker (*Nwazert: no azért* ‘well, for that reason’). The ÚESz. identifies it as an onomatopoeic unit, whose original function was “to introduce exhortation, encouragement, consent, insistence, concession, relief, question, or request. Later (depending on context), additional functions such as expressing surprise, anger, praise, or threat became associated with it.” (my translation)¹

The multiple functions of *no* are also reflected in explanatory dictionaries. ÉrtSz.² associates *no* primarily with interjectional meanings, grouped under volition and emotion. These include imperative, urging, request-like, and animal-directed uses, as well as a range of expressive meanings such as irritation, surprise, doubt, praise, and concession (ÉrtSz. V. 220–221).

1 No direct URL can be provided for the relevant entry of the ÚESz.; all search results are displayed under the same general page: <https://uesz.nytud.hu/index.html>

2 The most comprehensive descriptive dictionary of Hungarian to date is the seven-volume dictionary ÉrtSz. (Comprehensive Dictionary of the Hungarian Language). In *Nagyszótár* (Great Dictionary of the Hungarian Language, <https://nagyszotar.nytud.hu/index.html>), which is still in progress, the entries for *no* and *na* have not yet been completed.

Defining *no* as “an expressive or coloring word in emotional sentences” (my translation), the dictionary lists eight main categories (e.g. expression of volition; urging in questions; adding emotional coloring to sentences; a mood-setting word used as a transition in narration; introducing a farewell; adding an item in an enumeration; a response to being addressed; an emotion-expressing word used instead of elaboration). Among these, however, there are also clear discourse-organizing uses associated with *no*-based discourse markers, especially *no mármot* ‘well then’ which the ÉrtSz. describes as a formula introducing a new line of thought or a summary in expository discourse (ÉrtSz. V. 221). The dictionary also records related uses of *no* in introducing questions, bridging parts of narratives, marking leave-taking, and reinforcing the closure of a statement. For *na*, the ÉrtSz. lists largely the same meanings and co-occurring elements as for *no*, with the additional use of *na* before deictic elements.

In the New Hungarian Dialect Dictionary (ÚMTSz.), *no* is associated especially with the northeastern dialect region, while *na* is described as widespread in several other regions. This is significant because, while the CzF. (1867) stated that *na* “is used only regionally instead of the usual *no*” (IV. 680), the twentieth-century dictionaries make no such remark about *na*, whereas in the entry for *no*, this restriction still appears in reference to its use for urging animals (see above, ÉrtSz. V. 220).

From the Modern Hungarian period onward, *no* and *na* exhibit substantial alternation, although the latter is already attested from the first decades of the Middle Hungarian period: in his drama *Elektra* (1558), Péter Bornemisza used both forms, the example in (2) represents the first occurrence of *na* (see Dér 2026 for details).³

- (1) KED, Bornemisza Péter: *Elektra*, 1558

Clytem:	<i>Most</i>	<i>latom,</i>	<i>io</i>	<i>baratom</i>	<i>volt</i>	<i>vrad</i>
	now	see:1SG.DEF	good	my.friend	was	your.lord
	‘Now I see that your lord was a good friend,					
<i>nekem,</i>	<i>kit</i>	<i>en</i>	<i>neki</i>	<i>Soha</i>		
to.me	whom	I	him	never		
To me, one whom,						

3 Since the study focuses on pragmatic phenomena, I use simplified glossing.

meg *nem* *szolgálhatok,*
 VPFX not serve:POT:1SG
 I can never repay,
de ***no***, *ne* *halgas!*
 but DM PROH be.silent:IMP.2SG
 But **come on**, do not be silent!’

(2) KED, Bornemisza Péter: Elektra, 1558

Clytem: *Na, ez aran lanch es imez erszenbe valo*
 DM this gold chain and this purse:ILL belonging
 ‘Well, here is this gold chain, and these gold coins for your
 purse,
araniak es imez ruhak, legien oeröm
 gold and these clothes be:IMP.3SG joy
mondasodert.
 your.saying:CAUSFIN
 and these clothes, may there be joy for your message.’

Based on the data of the Hungarian Historical Corpus (Magyar Történeti Szövegtár, MTSz) from the Modern Hungarian period (starting in 1772), it appears that *no* remained the dominant form until the 1940s, after which its frequency declined, while *na* began to increase from the early 1900s. In the early 1940s, their proportions were still similar, but from then on *na* became dominant.

The analysis of random samples of 500 occurrences from the MTSz showed that from the mid-twentieth century onward, both *na* and *no* increasingly shifted toward (syntactically non-integrated) clause-internal position, while their formerly dominant clause-initial positions became less frequent. Clause-final uses remained rare and relatively stable over time. The two forms also differ in their combinations with other elements: such combinations are less frequent with *na* than with *no*, and the most typical co-occurring elements only partly overlap.

The “role reversal”, that is the replacement of the formerly dialectal *na* by *no*, was made possible by the fact that their functions were practically identical. For both forms, the causative or urging use declined during the twentieth century, while topic-management functions (such as topic initiation or shift) became dominant, as reflected by their increasing occurrence in (syntactically non-integrated) clause-internal position (Dér 2026).

Analyses based on online historical databases show that the earliest uses of *no* already reflect an advanced stage of grammaticalization: it typically appears in (syntactically non-integrated) clause-internal position, marking transitions or conclusions (Dér 2026). Since it is assumed to have originated as an onomatopoeic interjection, its first function was likely that of expressing volition or emotion – thus serving as an urging or attention-getting device intended to prompt others to act. This was probably followed by its sequential or rhetorical (topic-organizing) uses: the conclusive role likely preceded the transitional one. The responsive *no* could easily have developed in reaction to an initiative *no* (hence their turn-initial positions may be earlier than the turn-internal one), while marking discourse or topic transitions is interpretable primarily within turns or narrative segments. The emotional/emphatic component remained present in both *no* and *na*, and can appear in various ways in connection with any of their main functions.

The preceding discussion has addressed the origin and historical development of Hungarian *no* and *na*. Both forms are generally treated in Hungarian sources as originating from interjections, while they can also be related to broader cross-linguistic discussions of discourse particles, discourse markers, and interactives. The earliest attestations of *no* already show uses that go beyond interjectional meaning and point toward discourse-marker functions, whereas *na* appears later and was long regarded as a regional variant. Although *no* was historically the dominant form, corpus data indicate that *na* gradually gained ground from the twentieth century onward and eventually became the more frequent form. This development can be explained by the substantial functional overlap between the two forms, together with the decline of their earlier urging uses and the increasing prominence of their discourse-organizing functions.

What remains less clear, however, is how this historical development is reflected in present-day Hungarian. Previous descriptions suggest both a close functional relationship between *no* and *na* and a possible difference in their regional or stylistic distribution, but their current use has not yet been examined systematically from a discourse-variationist perspective. In particular, it is still unclear whether the alternation between the two forms reflects a standard vs. dialectal distinction, register-based preferences, functional specialization, or differences in their combinations with other discourse markers.

The present study therefore addresses the following research questions:

- 1) Does the present-day distribution of *no* and *na* support a standard vs. dialectal distinction, or is their variation better explained in terms of register and discourse function?
- 2) How are the two forms distributed across discourse positions and functions in spontaneous conversation and spoken press language?
- 3) To what extent do *no* and *na* differ in their combinations with other discourse markers, and what do these combinations reveal about their functional specialization?

3. Data and method

The corpus analyses were carried out on two datasets representing present-day Hungarian: the Hungarian Gigacorporus 2 (MNSz2; Oravecz, Váradi & Sass 2014), which serves as the largest synchronic database, and a subcorpus of the Hungarian Spontaneous Speech Database (BEA) compiled specifically for the present study. Multiplicated forms of *no* and *na* are not discussed in the present study; for details, see Dér (2026).

Thirty conversations with a total duration of 7 hours and 44 minutes were selected from the BEA (Gósy et al. 2012, Horváth et al. 2019), amounting to 37,129 words. Each conversation involved three participants engaged in discussions about everyday topics, all of whom were residents of Budapest. Audio recordings for the BEA corpus began in 2007.

The 15 male and 15 female main speakers (A) were selected to represent three age groups: the youngest (20–35 years), the middle-aged (36–50 years), and the oldest (51–71 years), with five male and five female speakers in each group. Of the 30 participants, one had completed primary education, 11 had a secondary school diploma, and 18 had a higher education degree (college or university).

When compiling the BEA, one of the main principles – besides maintaining the spontaneous character of speech – was to ensure that the experimental subject, that is, the main speaker (Speaker A), spoke as much as possible. The other two interlocutor roles were therefore not represented by a speaker sample comparable to that of Speaker A: Speaker T1 was the experimenter, a 29-year-old female linguist at the

then Research Institute for Linguistics of the Hungarian Academy of Sciences, while Speaker T2 was one of her linguist colleagues, one man and two women in their twenties, who participated as the third speaker in the conversations. Because including Speaker T1 and T2 would have treated non-comparable roles as equivalent and overrepresented a small, relatively stable set of interlocutors, the statistical analyses include only the data of Speaker A. The detailed analytical criteria are described in Section 4.1.

For the BEA examples, I followed the transcription conventions of the BEA (see Gósy et al. 2012: 18ff). Sentence boundaries are not marked, but transcription otherwise follows Hungarian orthography. Symbols used in the examples: ! = “noise” (e.g. laughter, coughing, etc.); □ = perceived pause; * = unclear or unintelligible segment; () = overlapping speech; repeated vowels = hesitation

MNSz2 (v.2.0.5) is a corpus of over one billion word tokens, comprising seven subcorpora (written press, fiction-literature, scientific, official, personal: forums and social media, and spoken press language), which I interpret as registers. With the exception of the personal and spoken subcorpora, the material in each corpus derives from five regional varieties of Hungarian: Hungary, Slovakia, Ukraine (Transcarpathia), Romania (Transylvania), and Serbia (Vojvodina). For the analyses, I extracted complete samples as well as random samples of 500 occurrences (see Section 4.2 for details). The queries were carried out on 6 July 2022.

In the MNSz2 gigacorporus, which aims to represent present-day Hungarian language use, I examined the distribution of *no* and *na* across regional varieties and registers (referred to by the corpus compilers as “stylistic levels”). The interpretation of the data is limited by certain characteristics of the subcorpora: in the social media subcorpus, there are numerous duplicate or triplicate hits, as portions of posts and comments are often repeated or “quoted” when users reply to one another. These exact repetitions were filtered out and removed from the dataset prior to analysis. As some of the works in the literary subcorpus are re-editions of earlier texts and were not excluded, they do not reflect late twentieth- or early twenty-first-century usage; this constitutes another source of imbalance in the corpus.

The overall approach of the analyses draws on Pichler’s (2013) framework of discourse variation analysis, insofar as *no* and *na* are

treated as discourse variants within a system of variation. For the description of discourse positions and functions, the analysis takes as its starting point the NU ~ NÅ categories distinguished by Auer and Maschler (2016: 11–35) (Table 1):

Table 1. Discourse positions and functions of NU~NÅ according to Auer & Maschler (2016).

Positions		Functions
1. Stand-alone or turn-initial ⁴ but pre-clausal position	A) initiative sequential position	a) urging, prompting
		b) marking a ‘next step’ as a consequence of prior talk
	B) responsive sequential position	a) ‘go-ahead’
		b) (emphatic) affirmation
		c) appreciative affiliation
		d) in third sequential position: reconfirming
2. Turn-internal pre-clausal position		e) surprise
		f) not quite the preferred response
		a) marking a transition into a next activity
3. Turn-final position		b) introducing the upshot of a speaker’s turn in an argumentative context
		c) marking a transition from information-giving to assessment
4. Syntactically integrated position		emphatic reinforcement which involves epistemic and/or affective stance
		a) introducing an account that marks a change in the trajectory of events
		b) mitigating
		c) marking affective stance, often implying disagreement
		d) tying a turn back to the preceding exchange of arguments from which it is presented as a consequence
		e) functioning as part of an idiomatic question

4 A turn is understood as the basic conversational unit extending from one speaker change to the next. Sequences are stretches of talk serving a common interactional goal.

4. Results

4.1. The use of *no* and *na* in the spontaneous conversations of the BEA

The most striking result of the analysis⁵ is that not a single occurrence of *no* was found, in contrast to *na*, which appeared 166 times in the corpus.⁶ Of these, 43.97% (73 tokens) were produced by the main participant (A). The overall mean frequency for female speakers was 0.02315 occurrences per word, compared with 0.00206 for male speakers. Although more than half of the *na* occurrences (61.64%) came from female speakers, neither the gender difference (Mann–Whitney $U = 94,000$, $p > 0.05$) nor the differences across the three age groups (ANOVA: $F = 2.077$, $df = 2$, $p > 0.05$) were statistically significant. In what follows, I describe the *na* tokens produced by the A speakers in terms of their conversational positions and functions.

The data associated with the A speakers show that three-quarters of the *na* tokens represent turn-internal use (54 cases, 73.97%), nearly one-quarter (19 cases, 26.03%) are responsive. Within a turn, speakers typically use *na* at the beginning of a quoted segment – either quoting themselves or others (3) – and one main discourse functions can be identified: signaling the onset of the next discourse unit (4):

(3) bea015, 60-year-old woman

A: *alig* *vártam*
 hardly wait:PST:1SG
 ‘I could hardly wait
hogy a férjem hazaérjen
 that the my.husband arrive.home:SBJV:3SG
 for my husband to come home
mondjuk hét óra negyed (nyolckor ű)
 say seven o’clock quarter eight:TEMP HES
 around seven or a quarter to (eight uh)’

5 Since individual speakers participated in conversations of unequal duration, frequency ratios were normalized by dividing the number of occurrences by the total number of words (ratio per word number: nr of hits / nr of words).

6 Upon repeated listening, I identified only one case in which *na* sounded more *no*-like.

T1: (*ühhüm*)

uh-huh

A: *és mondtam*
and say:PST:1SG
'and I said

hogy na most a tied
that DM now the yours
well now they're yours'

(4) bea023, 71-year-old woman

A: *hát ööö! ööö vi hajnalban ööö kelek*
well HES HES FS dawn:INE HES get.up:1SG
'well um um I get up early in the morning um

ööö éss ööö! busszal jövők
HES and HES bus:INS come:1SG
and um I take the bus

na most a hajnali közlekedés még elég jó!
DM now the early.morning:ADJ traffic still quite good
well the early morning traffic is still quite good'

A typical combination of *na* is with *mindegy* ('never mind'), with which the speaker closes a topic (5) or the entire turn:

(5) bea039, 33-year-old man

A: *a fele ugyanis biztos*
the half.of.it namely surely
'half of it um, namely, is surely the people who
hogy akik arra utaznak
that who:PL that.way travel:3PL
travel that way
az arra járó vonaton történnek a
the that.way travel train:SUPERS happen:3PL the
bliccelések
fare-dodging:PL
the fare-dodging happens on the train that goes that way
hogy ezt hogy mutatják ki
that this how prove:3PL.DEF VPFX
how they prove this
azt nem tudom
that:ACC not know:1SG.DEF
I don't know

na mindegy

DM never.mind

well never mind

nem is ez a lényeg

not even this the point

that is not really the point

hanem az hogy ööö hogy ugye

but the that HES that as.you.know

that, um, that, as you know,

elzárják a kordonnal

close.off:3PL.DEF the cordon:INS

they close it off with the cordon'

Cases in which the speaker produced *na* at the beginning of their own turn, but in response to a prior contribution rather than as the first move in a new sequence, were included among responsive uses. In these cases, *na* was employed in relation to the preceding situation or topic, expressing a stance located somewhere along the agreement–disagreement continuum, typically as part of an evaluative comment. Example (6) illustrates an agreeing, reinforcing response:

- (6) bea020, 52-year-old woman

[Context: The three participants are discussing the difficulties of the teaching profession.]

T1: *simán megkapja*

easily get:3SG.DEF

'she'll easily get it

mer annak tuti hogy van ismerőse

because that.DAT surely that be:3SG her.acquaintance

because surely she has some acquaintance

vagy nem tudom mi

or not know:1SG her.what

or whatever'

T2: *hát csak nem biztos*

well only not certain

'well but it's not certain

hogy ott tud maradni

that there can:3SG stay:INF

that she can stay there'

- T1: *hát ja*
 well yes
 ‘yeah right’
- A: *na hát a mai tanárokat*
 DM well the today teacher:PL:ACC
abszolút nem irigylem
 absolutely not envy:1SG.DEF
 ‘well I absolutely don’t envy today’s teachers’

A manifestation of disagreement may involve, for example, mentioning an exception or qualifying the content of the previous speaker’s statement from one’s own perspective:

- (7) bea062, 35-year-old woman
- T2: *az nincs benne a rövidített részbe*
 that not.be.3SG in.it the shortened part:ILL
 ‘that’s not in the shortened version’
- T1: *nem tudom*
 not know:1SG.DEF
 ‘I don’t know
de valami olyanra
 but something such:SUBL
 but something
ami tényleg nem volt abba benne
 that really not was that:ILL in.it
 that really wasn’t included in that part
 (és azért)
 and therefore
 (and therefore)’
- A: (*na jó*) *hát mondjuk sok minden*
 DM okay well let.us.say much everything
nincs a rövidített részekbe benne
 not.be.3SG the shortened parts:ILL in.it
 ‘(well okay) but let’s say there are a lot of things that aren’t included in the shortened versions’

A motivational (‘go ahead’) or urging use of *na* could also be observed in responsive contexts, also in echo-questions:⁷

7 I would like to thank the editors for this comment.

(8) bea010, 22-year-old woman

T2: *én ismerem a harmadik típust is*
 I know:1SG.DEF the third type:ACC too

‘I know the third type too’

T1: *na mi a harmadik*
 DM what the three:ORD

‘well what’s the third one?’

A: *na mi*
 DM what
 ‘well what?’

None of the speakers used the discourse marker stand-alone, in syntactically integrated or turn-final position.

As can be seen, in many cases *na* does not fulfill the given function on its own but appears together with other discourse markers, which it always precedes in position. This is indeed the general tendency: in four-fifths of all cases, *na* co-occurs with at least one other discourse marker, and only 20.54% of the tokens occur without such markers. The most frequent combinations are shown in Table 2.

Table 2. The most frequent combinations of *na* with discourse markers in BEA.

DM-combinations	Nr of occurrences
<i>na most (akkor)/(hát)</i> ‘well now (then)/(so)’	12
<i>na hát (akkor)/(aztán)</i> ‘well then (so)/(then)’	8
<i>na jó (hát) (de)/(mondjuk)/(szóval)</i> ‘well okay (so) (but)/(let’s say)/(so)’	8
<i>na de (hát) (ugye)</i> ‘well but (so) (you know)’	7

Depending on their respective functions, these discourse markers either nuance or emphasize the role of *na*, or they add further pragmatic meanings. A good example of the former is the combination *na de* ‘well but’, since *na* alone (with appropriate intonation) can express disagreement, while *de* makes this meaning explicit, and the additional markers (such as *hát* ‘well, so’ or *ugye* ‘you know’) contribute further nuances such as obviousness or presupposition of the listener’s agreement:

(9) bea035, 45-year-old woman

A: *ugye ez annak köszönhető volt*
 you.know this that:DAT due.to was
 ‘you know this was because
hogy (segítettek felzárkózni)
 that help:PST:3SG catch.up:INF
 (they helped her catch up)’

T1: (*hát hogy sokkal többet*) *foglalkoztak vele*
 well that much more:ACC deal.with:PST:3SG with.her
 ‘(well, because they spent much more) time with her
igen
 yes
 yes’

A: *na de hát ugye ez egy értelmiségi házaspár*
 DM but after.all you.know this a intellectual
 couple
 ‘well but you see this is an intellectual couple’

To summarize, the BEA data show that *na* occurs mainly in turn-internal and responsive positions. In these conversations, *na* typically serves discourse-organizing and stance-marking functions, and it most often appears in combination with other discourse markers, which refine or reinforce its pragmatic contribution.

4.2. The use of *no* and *na* in the MNSz2 corpus

Na occurs nearly 1.5 times as often as *no* in the dataset. However, when the data are broken down by subcorpus, this difference in relative frequency largely disappears – no register shows a difference exceeding 10%, with the largest variation (6.48%) found in the spoken press subcorpus.

Table 3. Frequency of *no* and *na* in MNSz2 by register.

Registers	Absolute frequency		Relative frequency		Size of subcorpus (million tokens)		Occurrence per token (normalized frequency)	
	<i>no</i>	<i>na</i>	<i>no</i>	<i>na</i>	<i>no</i>	<i>na</i>	<i>no</i>	<i>na</i>
Written press	7263	5358	12.96%	7.05%	364.8		1.99095×10^{-5}	1.46875×10^{-5}
(Transcribed) spoken press	644	5793	1.15%	7.63%	76.2		8.4514×10^{-6}	7.60236×10^{-5}
Personal (internet forums)	8297	13482	14.80%	17.76%	57.9		1.432988×10^{-4}	2.328497×10^{-4}
Personal (social media)	28907	40725	51.56%	53.65%	243.2		1.18861×10^{-4}	1.674548×10^{-4}
(Popular) science	5311	3228	9.47%	4.25%	117.9		4.50466×10^{-5}	2.73791×10^{-5}
Official (political)	1649	2950	2.94%	3.89%	99.0		1.66566×10^{-5}	2.9798×10^{-5}
Fiction/Literature	3990	4373	7.12%	5.77%	80.6		4.95037×10^{-5}	5.42556×10^{-5}
Total	56061	75909	100,00%		1039.7		5.39204×10^{-5}	7.30105×10^{-5}

In terms of normalized frequency, *na* clearly surpasses *no* in spoken press language and in the personal subcorpora, and it is also slightly more frequent in the official and literary registers (see Figure 1).

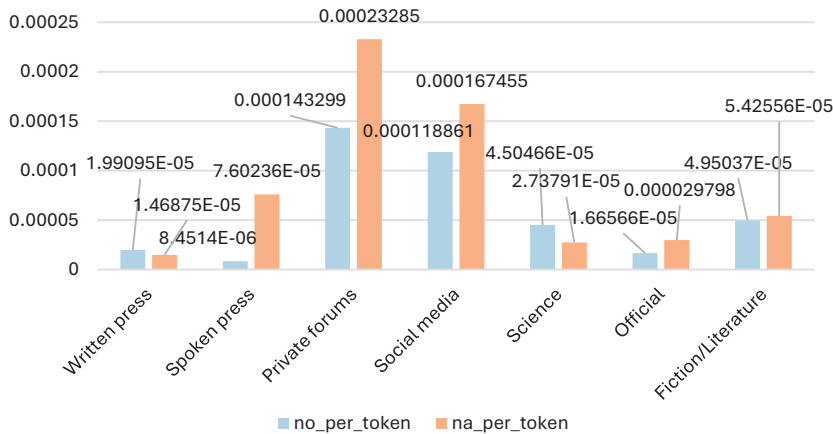


Figure 1. Normalized frequency of *no* and *na* by register.

Performing chi-square tests on the frequency data reveals clear patterns (see Table 3):⁸

- the difference is significant in favor of *na* in the spoken press, official, personal (forums, social media) and fiction/literature subcorpora;
- the difference is significant in favor of *no* in the written press and science subcorpora.

The subcorpora in which *na* is significantly overrepresented share a common feature: they represent either spoken language or spoken-like written language in computer-mediated communication (except the fiction subcorpus which also includes works rich in dialogue); the personal subcorpora naturally reflect more informal language use, while

8 In the study, goodness-of-fit chi-square tests were used as the default procedure (and are illustrated in the tables), but independence chi-square tests were first applied to both the register and regional data. For registers, the independence test showed that the frequency of *no* and *na* differs significantly across registers ($\chi^2 = 5683.60$, $df = 6$, $p < 0.001$), with a medium effect size (Cramér's $V = 0.208$), indicating a genuine association between register and form. Post-hoc analysis showed that *no* dominates primarily in written press, scientific, and literary registers, whereas *na* is predominant in spoken press, political, and personal (online) registers.

the others represent more formal usage. It is important to note, however, that the effect size is small (or very small) for almost all subcorpora, as well as for the corpus as a whole. Thus, although the difference – “*na* dominates in spoken(-like) registers” – is statistically significant, its actual magnitude is minor. The only exception is the (transcribed) spoken press register, where the association between register and the frequency of *na* is quantitatively strong. For the other subcorpora, the effect size remains small or negligible, even though the frequency difference between the two markers is not due to chance.

Table 4. Results of chi-square tests and Cramér’s *V* for *no* and *na* by register.⁹

Register	N	χ^2	<i>df</i>	<i>p</i>	Cramér <i>V</i>	effect size	more frequent form
(Transcribed) spoken press	6437	4118,72	1	<0.05	0,7999	large	<i>na</i>
(Popular) science	8539	508,1261	1	<0.05	0,24394	small	<i>no</i>
Official (political)	4599	368,0367	1	<0.05	0,282888	small	<i>na</i>
Written press	12621	287,538	1	<0.05	0,1509	small	<i>no</i>
Personal (internet forums)	21779	1234,4104	1	<0.05	0,23807	small	<i>na</i>
Personal (social media)	69632	2005,7606	1	<0.05	0,1697	small	<i>na</i>
Fiction/literature	8363	17,5402	1	<0.05	0,04579	very small	<i>na</i>

I also examined the regional distribution of *no* and *na*. The large size of the Hungarian subcorpus results in both markers appearing far more frequently there than in other regions. Not all registers are represented in every regional dataset: the spoken press register exists only in the Hungarian materials,¹⁰ while the personal subcorpora are absent from the Transcarpathian region, that is, precisely those registers in which *na* and *no* are generally the most frequent across the entire corpus.

⁹ The independence chi-square test also showed that the frequency of *no* and *na* differs significantly across regions ($\chi^2 = 91.63$, $df = 4$, $p < 0.001$), but the effect size is very small (Cramér’s *V* = 0.026), suggesting that although regional differences are statistically detectable, they are negligible in practical terms.

¹⁰ Characteristics of subcorpora by region: <https://clara.nytud.hu/mnsz2-dev/stat.html>

Table 5. Frequency of *no* and *na* in MNSz2 by region.

Region	Absolute frequency		Relative frequency (entire corpus)		Size of sub- corpus (million tokens)		Occurence per token (normalized frequency)	
	<i>no</i>	<i>na</i>	<i>no</i>	<i>na</i>	<i>no</i>	<i>na</i>	<i>no</i>	<i>na</i>
Hungary	53945	73131	96.25%	96.35%	1013.9		5.32054×10^{-5}	7.21284×10^{-5}
Slovakia	1005	1415	1.79%	1.86%	17.3		580925×10^{-5}	8.17919×10^{-5}
Subcarpathia	265	555	0.47%	0.73%	2.5		1.06×10^{-4}	2.22×10^{-4}
Transylvania	585	599	1.05%	0.79%	3.9		1.5×10^{-4}	1.535897×10^{-4}
Vojvodina	261	209	0.46%	0.27%	2.0		1.305×10^{-4}	1.045×10^{-4}
Total	56061	75909	100.00%		1039.7		$9,95596 \times 10^{-5}$	1.26802×10^{-4}

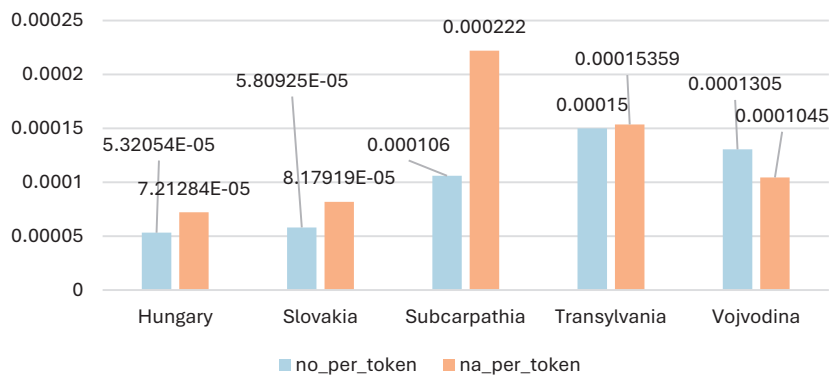


Figure 2. Normalized frequency of *no* and *na* by region.

Across the regions, *na* is more frequent in Hungary, Slovakia, and Subcarpathia, while Vojvodina shows the opposite pattern with *no* being the more common form. These differences are statistically significant in all four regions, but the effect sizes are small, indicating that despite the systematic preferences, the magnitude of the deviation from a balanced distribution is limited. Transylvania is the only region where no statistically significant difference was found between the two variants, suggesting a balanced use of *no* and *na*. The medium effect size observed in Subcarpathia indicates that the deviation from a balanced distribution is not merely a consequence of sample size, but reflects a genuinely stronger preference for the variant *na* in this region.

Table 6. Results of chi-square tests and Cramér's V for *no* and *na* by region.

Region	N	χ^2	df	p	Cramér V	effect size	more frequent form
Subcarpathia	820	102,561	1	<0.05	0,353659	medium	na
Vojvodina	470	5,753191	1	<0.05	0,110638	small	no
Transylvania	1184	0,165541	1	>0.05	0,011824	very small	na
Hungary	127076	2896,712	1	<0.05	0,150981	small	na
Slovakia	2420	69,462	1	<0.05	0,1694214	small	na

4.3. The spoken subcorpus of MNSz2

Since the BEA corpus did not contain any *no* tokens, it was not possible to conduct a functional analysis of *no* and *na*. Therefore, we extracted random samples of 500 occurrences each from the spoken subcorpus of MNSz2,¹¹ which contains transcriptions of spoken press materials: news, news interviews, reports, and features.¹² These genres differ in their degree of preparedness: some texts may be read aloud or pre-planned, whereas others, especially interviews, contain more spontaneous spoken interaction. The number of valid tokens was 457 for *no* and 499 for *na*. It was not possible to separate news items from features due to the presentation format of the data, but we took into account that variation between journalistic subgenres might influence the behavior of the discourse markers under analysis.

In defining the functions, we again started from the categories summarized by Auer and Maschler (2016) (Table 1), which we modified when necessary in light of the data from the analyzed subcorpus. The necessary adjustments to Auer and Maschler's categories are specified in the analyses below. They mainly concern the treatment of position–function correspondences, the separation of initiative and responsive uses, and the classification of functionally ambiguous cases.

4.3.1. The characteristics of *no* in the spoken press subcorpus

Among the positional types, the turn-internal clause-preceding position clearly dominates for *no*: three-quarters of the valid tokens in the random sample occur in this position (343 out of 457, 75.05%), cf. (10).

11 The spoken subcorpus contains approximately 76 million words, consisting of selected materials from MR1 Kossuth Rádió between 2004 and 2012 (<https://clara.nytud.hu/mnsz2-dev/stat.html>). Kossuth Rádió is Hungary's public-service national news and current affairs radio channel, broadcasting mainly political, economic, cultural, and social programs. For a critical discussion of the subcorpus composition, see Lanstyák (2023: 67ff).

12 The sampling date was 18 September 2025. The examples are reproduced verbatim, including typos and punctuation or spelling inconsistencies, without [sic] notation. For easier readability, each speaker's contribution is presented on a new line. In some transcripts, only the interviewer/host is named, while the interviewee is not; in such cases, the latter's utterances also appear on separate lines.

(10) MNSz2, #885027717, spoken

Kovács László [politician of the Hungarian Socialist Party (MSZP)]:

(...) *Az, hogyha minden*
 that if everything
 ‘If everything
maradna így,
 stay:COND.3SG like.this
 stayed the same,
vagyis maradnának a párton belül
 that.is remain:3PL the party:SUPRESS inside
 if they remained within the party,
és továbbra is úgy viselkednének,
 and further.on too so behave:COND:3PL
 and continued to behave
mintha a párton belül egy másik
 as.if the party:SUPRESS inside a other
párt lennének.
 party be:COND:3PL
 as if they were a separate party within the party.
No, ez nem lesz megengedhető
 DM this not will permissible
a kongresszus után.
 the congress after
Well, this will not be allowed after the congress.’

Auer and Maschler (2016: 27–30) distinguish three functions for turn-internal position – transition to the next activity, final conclusion, and transition from providing information to drawing a conclusion – all of which can be observed in Hungarian as well. The signaling of transitions dominates, accounting for 47% (160 tokens) of the turn-internally positioned occurrences (10). Due to the characteristics of the register (narrative/expository sections, argumentation, evaluations, frequent use of quotations), additional functions typically associated with (turn-initial) initiative or responsive positions also appear with turn-internally placed *no*, such as directive or urging (11) and expressing disagreement (12).

(11) MNSz2, #893483771, spoken

(...) *egy találat, egyetlenegy találat döntheti*
 one hit single.one hit decide:POT:3SG.DEF
el azonnal Kárpáti javára az olimpiai
 VPFX immediately Kárpáti in.favor.of the Olympic
bajnokságot,
 championship:ACC
 ‘... one single hit, just one hit, could immediately decide the Olympic
 championship in Kárpáti’s favor,
vagy pedig úgy, hogy a lengyel jóvoltából
 or then so that the Polish thanks.to
 or, thanks to the Polish fencer,
egy bonyolult holtverseny alakul ki.
 a complicated tie develop.3SG VPFX
 a complicated tie could arise.
No Rudi, no Rudi, úgy, ahogy
 DM Rudi DM Rudi so as
megszoktuk tőled,
 get.used.to:PST:1PL.DEF from.you
Come on, Rudi, come on, Rudi, just as we’re used to from you,
nyugodtan, nyugodtan,
 calmly calmly
 calmly, calmly,
leengedi karját egészen mélyre,
 lower:3SG.DEF his.arm:ACC quite deep.to
 he lowers his arm all the way down,
majdnem a pástot éri,
 almost the piste:ACC reach:3SG.DEF
 almost touching the piste,
s onnan gyorsan fölemeli
 and from.there quickly VPFX.raise:3SG.DEF
 then quickly raises it
és aztán egy csel után támad.
 and then a feint after attack.3SG
 and attacks after a feint.’

- (12) MNSz2, #854340864, spoken

Valder Durante a világ egyik legjobb tollú
 Walter Duranty the world one.of best pen:ADJ
újságírója
 journalist:POSS.3SG
 ‘Walter Duranty, one of the world’s most talented journalists,
a New York Times moszkvai tudósítója volt.
 the New York Times Moscow:ADJ correspondent:POSS.3SG was
 was the Moscow correspondent for The New York Times.
Kiszáll Ukrajnába, Ukrajnában például
 go.out.3SG Ukraine:ILL Ukraine:INE for.example
 He goes to Ukraine – there are, for example,
jelenleg vannak egész hatalmas falvak,
 currently are quite huge village:PL
 entire huge villages
ahol ukrán nevet viselnek a lankák,
 where Ukrainian name:ACC bear:3PL the hill:PL
 where all bear Ukrainian names, the hills,
a folyók, az erdők
 the river:PL the forest:PL
 the rivers, the forests,
és az utcák is nincs ukrán
 and the street:PL too not.be.3SG Ukrainian
 and streets, too, but there isn’t a single Ukrainian
ebben a faluban, ezek orosz betelepültek.
 in.this the village:INE these Russian settler:PL
 in this village; they are all Russian settlers.
No de Valter Durante írja sorra
 DM but Walter Duranty write:3SG.DEF in.succession
a tudósításait
 the report:PL.POSS.3SG:ACC
Well but Walter Duranty keeps writing his reports:
ez egy rágalom nem igaz
 this a slander not true
 this is slander, it’s not true,
a Nagy Szovjetunióban nem éheznek az emberek.
 the Great Soviet.Union:INE not starve:3PL the people:PL
 people in the Great Soviet Union are not starving.’

The expressions of disagreement do not always reflect the speaker's (editor's, journalist's, or interviewee's) own professional or personal opinion; they may instead represent the views of previously mentioned (even fictional) figures, or the speaker may argue with themselves, presenting possible counterarguments and responding to them in the course of the discourse.

In turn-internal position, *no* is followed by another discourse marker in 76.09% of cases (261 out of 343), meaning that in three-quarters of the tokens it occurs in combination, most frequently with *no meg* 'well and', *no és* 'well and', and *no de* 'well but'.

The above functions can also co-occur, such as in cases where *no* simultaneously signals both exhortation and transition or conclusion:

(13) MNSz2, #932965423, spoken

<i>No</i>	<i>de</i>	<i>van-e</i>	<i>menekülés</i>	<i>ez</i>	<i>elől</i>
DM	but	be.3SG-QPTCL	escape	this	from
<i>a</i>	<i>fokozódóan</i>				
the	increasingly				
'But well, is there any escape from this increasingly					
<i>agresszív</i>	<i>napi</i>	<i>információ</i>	<i>özön</i>	<i>elől.</i>	
aggressive	daily	information	flood	from	
aggressive flood of daily information?					
<i>Nem</i>	<i>hiszik?</i>				
not	believe:3PL.DEF				
You don't think so?					
<i>No</i>	<i>akkor</i>	<i>hallgassák</i>	<i>Sárközi</i>	<i>Júliát.</i>	
DM	then	listen:IMP:3PL.DEF	Sárközi	Júlia:ACC	
Well then, listen to Júlia Sárközi.'					

The *no* used for appending additional information frequently co-occurs with *meg* 'and', *és* 'and', and *persze* 'of course':

(14) MNSz2, #874808745, spoken

[Context: Host opening the radio programme.]

<i>Jó</i>	<i>napot</i>	<i>kívánok,</i>	<i>10</i>	<i>óra</i>	<i>van</i>
good	day:ACC	wish:1SG	ten	o'clock	be.3SG
'Good morning, it's 10 o'clock,					
<i>Érczfalvay</i>	<i>András</i>	<i>vagyok.</i>			
Érczfalvay	András	be:1SG			
I'm András Érczfalvay.					

Nem kötelező védőoltások gyerekeknek,
 not obligatory vaccination:PL child:PL:DAT
 Non-mandatory vaccinations for children:
indokoltság, hatékonyság, ár, no meg persze
 justification effectiveness price DM and of.course
a veszélyek.
 the danger:PL
 justification, effectiveness, cost, **well, and of course**, the risks.'

No frequently appears in argumentative contexts where the speaker also expresses their opinion and the emotions connected to it or to the topic:

(15) MNSz2, #884975871, spoken

Természetesen az Edeledie port Edeledie
 of.course the Adelaide port Adelaide
futballklubok indulói is ott vannak,
 football.club:PL contestant:PL.POSS.3SG too there be:3PL
 'Of course, the contestants from the Adelaide port Adelaide football clubs are also there,
no *ezen nem vagyok meglepve,*
 DM this:SUPERS not be:1SG surprised
well, I'm not surprised,
elvégre ausztrálokról van szó,
 after.all Australian:PL:DEL be.3SG word
 after all, we're talking about Australians;
sportnemzet vagyunk.
 sport.nation be:1PL
 we're a sporting nation.'

Because of register-specific features and the way the MNSz2 data are presented (see above), I did not distinguish between (turn-initial) initiative and responsive uses (112 out of 457, 24.50%), contrary to my original intention. This is because the interviews do not always have a clear question–answer structure. The interviewer, when taking back the floor, often responds to the interviewee's answers, arguments, or even questions, while also introducing a new point. There were therefore few cases that were purely (turn-initial) initiative or purely responsive. In example (16), the use of *no* cannot be assigned unambiguously to a single category. It marks the interviewer's uptake of his interview

partner Endre Szemerédi's (ESz) preceding answer, with a nuance of curiosity or mild surprise, while at the same time prefaces a follow-up question concerning the relevance of this answer for mathematical thinking:

(16) MNSz2, #888126214, spoken

Mv.: *Tudom, hogy izgul,*
 know:1SG.DEF that be.nervous.3SG

Host: 'I know you're nervous,
nem nagyon szereti a mikrofont,
 not very like:3SG.DEF the microphone:ACC
 you don't really like the microphone;
jobban szereti a papírt és a ceruzát
 more like:3SG.DEF the paper:ACC and the pencil:ACC
 you prefer paper and pencil,
vagy a tollat, ugye?
 or the pen:ACC right
 or a pen, right?'

ESz: *Vagy a labdát.*
 or the ball:ACC
 'Or the ball.'

Mv.: *No, annak mi köze van*
 DM that:DAT what connection:POSS.3SG be.3SG
 'Well, what does that have to do
a matematikai gondolkodáshoz?
 the mathematical thinking:ALL
 with mathematical thinking?'

ESz: *Hát nem sok, elősegíti,*
 well not much facilitate:3SG.DEF
 'Well, not much, it helps though;
nagyon szeretek focizni, meg teniszezni,
 very like:1SG football.play:INF and tennis.play:INF
 I really love playing football, tennis,
pingpongozni, labdajátékokat.
 ping-pong.play:INF ball.game:PL:ACC
 ping-pong, ball games.'

In the (turn-initial) initiative and responsive category, several functions are dominant: expressing disagreement (typically combined with *de* 'but') (17); signaling a conclusion (characteristically co-occurring

with *hát* ‘well’) (18); and indicating exhortation or urging. However, the most frequent function is introducing a transition in discourse, which is often accompanied by exhortation (19); together, these last two account for 38% of the tokens.

(17) MNSz2, #938005822, spoken

[Context: Interview with Attila Fodor (FA), the communications director of the CBA retail chain.]

FA: (...) *Szeretném még egyszer hangsúlyozni,*
want:COND:1SG.DEF still once emphasize:INF

‘(...) I would like to emphasize once again

hogy ez nem azonos a Magyarországon

that this not same the Hungary:SUPESS

that this is not the same as the CBA stores found in Hungary,

megtalálható CBA-val, hanem ez egy franchise

findable CBA-INS but this a franchise

partner ország.

partner country

this is a franchise partner country.’

Mv: *No de a logójuk ugyanúgy néz ki,*
DM but the logo:POSS.3PL same.way look.3SG out

‘Well, but their logo looks the same,

a betűtípus ugyanúgy néz ki,

the font same.way look.3SG out

the font looks the same,

a szín minden.

the color everything

the color, everything.’

(18) MNSz2, #906193609, spoken

R.: *Van-e olyan, aki már járt*
be.3SG-QPTCL such who already go:PST.3SG
iskolába?

school:ILL

‘Is there anyone who’s already been to school?’

?: *Most harmadikos.*

now third.grade:ADJ

‘Now in third grade.’

R.: *És hány éves?*

and how.many year:ADJ

‘And how old?’

?: *Harmincegy.*

thirty.one

‘Thirty-one.’

R.: *No, hát egy kicsit későn kezdte, nem?*

DM well a little late start:PST.3SG.DEF not

‘Well, a bit of a late start, isn’t it?’

(19) MNSz2, #939623338, spoken

[Context: Interview with editor-reporter Judit Godány (GJ).]

GJ: (...) *az volt az érzésünk,*

that be:PST.3SG the feeling:POSS.1PL

‘(...) we had the feeling

hogy akármelyik fotóalbum,

that any photo.album

that any photo album

akár a mi fotóalbumunk is lehetne.

even the our photo.album:POSS.1PL too be:COND.3SG

could have been our own.’

Rep.: *No, akkor lapozzunk bele.*

DM then leaf.through:IMP.1PL into.it

‘Well then, let’s turn a few pages.’

The only turn-final occurrence of *no* was presented as a separate unit in the transcription:

(20) MNSz2, #852769694, spoken

[Context: Interview with Péter Boross (BP), politician, member of the MDF (Hungarian Democratic Forum) party.]

Rep.: *És afelé halad most a párt?*

and that.direction.to move.3SG now the party

‘And is the party moving in that direction now?’

BP: *Már ellentétes azzal is,*

already contrary that:INS too

‘It’s already contrary to

amit Dávid Ibolya többször megerősített,

what:ACC Dávid Ibolya several.times confirm:PST.3SG

what Ibolya Dávid has repeatedly confirmed,

hogy az MDF balra zárt. No.

that the MDF left:SUBL close:PST.3SG DM

that the MDF has closed ranks to the left. Well.’

Three other instances are similar in that *no* occurs in units segmented as separate in the transcription (21), twice in combination with another discourse marker (*de* ‘but’, *tehát* ‘therefore’). Their functions vary (summarizing, signaling topic shift), so these were treated as a separate category.¹³

- (21) MNSz2, #883220128, spoken

De az emberek hol lépdelenek? A történelemben.
 but the people:PL where walk:3PL the history:INE
 ‘But where do people walk? In history.
S az mennyi időre méretezett? Mondjam-e?
 and that how.much time:SUBL designed say:IMP:1SG.DEF-QPTCL
 And how long is that designed to last? Shall I say it?
No. Ehhez kívánok most neked, néktek és
 DM to.this wish:1SG now you.DAT you.PL.DAT and
magunknak
 ourselves:DAT
Well. I wish you, all of you, and ourselves
is könnyű teherrel, nótaszóval, nem erőltetett
 too light burden:INS song:word:INS not forced
 a cheerful march with light burdens, with song, not forced
menetben, de célba érő hittel, vidám
 march:INE but goal:ILL reaching faith:INS cheerful
menetelést.
 marching:ACC
 but with faith that reaches the goal.’

4.3.2. The characteristics of *na* in the spoken press subcorpus

In the random sample of 500 *na* tokens, the proportion of turn-internal-position occurrences is practically the same as for *no*: 74.74% (373 out of 499 valid tokens). The distribution of functions, however, shows a slightly different pattern, as the signaling of discourse transitions appears to be a more prominent function: three-quarters of the

13 A reviewer raised the question of whether such discourse-marker combinations should be regarded as sentences. This is an important and theoretically interesting issue, but it lies beyond the scope of the present paper. Here I use “separate unit” in a transcription-based sense, without making any claim about the syntactic sentence status of these items.

turn-internally placed tokens (278, or 74.53%) fulfill this role, primarily in forms combined with *most* ‘now’ (235 tokens) (22). Other discourse markers that frequently co-occur include *akkor* ‘so’, *aztán* ‘then’, *hát* ‘well’, *persze* ‘of course’, and *ugye* ‘right?’

- (22) MNSz2, #935704998, spoken

Úgy mondom hogy a magyarság
 so say:1SG:DEF that the Hungarian.people
 ‘I should mention that Comrade Serov already
 elleni megtorlás Szerov elvtárs már ott
 against:ADJ reprisal Serov comrade already there
 practiced his reprisals against Hungarians there
 gyakorolta Kárpátalján.
 practice:PST.3SG:DEF Subcarpathia.SUPES
 in Subcarpathia.
Na most november 13-án jelenik meg egy ilyen
 DM now November 13-SUPES appear:3SG VPFX a such
Well, now, on November 13,
hadparancs.
 military.order
 such a military order appeared.’

In addition to conclusion (23), disagreement (24), and exhortation or urging (25) – the latter typically in quoted speech – can also be observed as functions of *na*, though at a much lower rate than in the case of *no*:

- (23) MNSz2, #864935009, spoken

De szóval azért amikor csak piacot vásárol
 but let.us.say PTCL when only market:ACC buy.3SG
 ‘But, well, when a company buys up a market
 a privatizáció okán egy cég, és aztán
 the privatization because.of a company and then
 through privatization and then
 odébbáll amikor máshol olcsóbb neki,
 move.away.3SG when elsewhere cheaper to.it
 moves elsewhere when it’s cheaper,
 a piaca viszont megmarad.
 the market:POSS.3SG however remain.3SG
 yet its market share remains.

Na szóval itt a bibi.
 DM let.us.say here the problem
Well, that's the problem right there.'

- (24) MNSz2, #918388794, spoken

Erre lehet azt mondani,
 this:SUBL can that:ACC say:INF
 'On this one can say,
amit a pénzügyminiszter mond,
 what:ACC the finance.minister say.3SG
 as the Minister of Finance does,
hogy hát ha majd bevallják.
 that well if then declare:3PL.DEF
 that it depends on whether it's declared or not.
Na, de bocsánat, ez az adósság ez létezik,
 DM but excuse.me this the debt this exist:3SG
Well, but excuse me, this debt exists,
független attól, hogy bevalljuk vagy
 independent that:ABL that declare:1PL.DEF or
 regardless of whether we declare it
nem valljuk be,
 not declare:1PL.DEF VPFX
 or not,
ezzel kezdeni kell valamit.
 this:INS start:INF must something:ACC
 we have to deal with it.'

- (25) MNSz2, #865590858, spoken

Rep.: *De a Dunaszerdahelyen történt magyar*
 but the Dunaszerdahely:SUPESS happen:PTCP Hungarian
 'The Hungarian community
labdarúgó szurkolókat érintő szlovák rendőri
 football fan:PL:ACC affecting Slovak police:ADJ
 in Slovakia also has its opinion on the Slovak police
erőszakról is megvan a véleménye a
 violence:DEL too VPFX:be.3SG the opinion:POSS.3SG the
 violence against Hungarian football
szlovákiai magyaroknak.
 Slovakia:ADJ Hungarian:PL:DAT
 fans in Dunaszerdahely.'

?: *Na, figyeljen ide! Én nagyobb magyar vagyok,*
 DM listen:IMP:3SG here I greater Hungarian be:1SG
 ‘Well, listen here! I’m more Hungarian
mint a nagyobb magyar, érti...
 than the greater Hungarian understand:3SG.DEF
 than the more Hungarian, you see...’

For the reasons mentioned earlier, the (turn-initial) initiative and responsive categories were merged for *na* as well, yielding a total of 123 tokens (24.65%). In these cases, co-occurrence with other discourse markers was frequent (102 instances), with *na most* ‘well now’ dominating as a transition-marking (topic-introducing or topic-shifting) formation (48 tokens):

(26) MNSz2, #885201986, spoken

[Context: Interview with Éva Várhegyi, economist and scientific adviser at Pénzügykutató Ltd.]

Mv.: *Nálunk viszont a hangsúly*
 at.us however the emphasis

Host: ‘In our case, the focus is
elsősorban a lakásokon vannak.
 primarily the apartment:PL:SUPSS be:3PL
 primarily on housing.’

VÉ.: *Itt most így van.*
 here now so be.3SG
 ‘That’s how it is now.’

Mv.: *Na most, a Nemzeti Eszközkezelő Társaság*
 DM now the National Asset.Management Company
 ‘Well, now, the National Asset Management Company
ugye azt tervezné, hogy
 you.know that:ACC plan:COND.3SG.DEF that
piacról vesz fel hiteleket (...).
 market:DEL take.3SG up loan:PL:ACC
 is planning to take out loans from the market (...)

Although the most prominent function, both with and without co-occurring discourse markers, is indeed signaling transition (64 tokens, 52.03%), expressions of disagreement (27) are also noteworthy (36 tokens, 29.26%). These are almost always clarifying or corrective objections on the part of the interviewer:

(27) MNSz2, #885563092, spoken

[Context: The participants are discussing whether the price of gold can be expected to fall. Pálfi Gergely (PG) suggests checking the recent chart data, which the host extends to data from the past year.]

PG: *Illetve az elmúlt évi chartot, igen.*
 or.rather the past year:ADJ chart:ACC yes
 ‘Yes, or rather last year’s chart.’

Mv.: *Na de hát akkor is előfordulhat,*
 DM but after.all then too occur:POT.3SG

Host: ‘**Well**, but still, it could happen
hogy egyszer csak zuhanni kezd.
 that once only plummet:INF begin.3SG
 that it suddenly starts to plummet.’

It must be emphasized, however, that distinguishing between transition and conclusion, as well as determining the presence of emotional expression, is often not clear-cut (since emotional overtones typically overlay existing functions). In such cases, the data were categorized according to the function judged to be most dominant.

A single example of a turn-final *na* appeared in the random sample, just as with *no*, in an emphatic confirmative role (‘that’s right / that’s the truth’):

(28) MNSz2, #864581272, spoken

[Context: Mihály Kupa, economist and former Minister of Finance of Hungary, discussing the proposed bank tax.]

(...) *higgye el nekem ahhoz,*
 believe:IMP:3SG.DEF VPFX I.DAT that:ALL

‘(...) believe me,

hogy 100 milliárdot csökkentsenek
 that 100 billion:ACC reduce:SBJV:3PL

to cut 100 billion,

ahhoz erre a a 30 milliárdra szükség van,
 to.that this:SUBL the the 30 billion:SUBL need be.3SG

this 30 billion is needed,

mert nincs erre fedezet na.
 because not.be.3SG this:SUBL cover DM

because there’s no cover for it, **well**.’

Two tokens of *na* were segmented as separate units in the transcription, though both can also be interpreted as turn-internal occurrences. They appear in combination with other discourse markers (*na jó, na most*). In (29), *na* together with *jó* introduces a summary, but the reporter also conveys a sense of reluctant acceptance of the situation:

(29) MNSz2, #853241326, spoken

[Context: The reporter is speaking with a caller about the negative financial consequences of a prize draw.]

- Rep.: *Én pedig nagyon szépen köszönöm,*
 I and very nicely thank:1SG.DEF
 ‘And I thank you very much
hogy elmondta, mert ez sok
 that VPFX:say:PST.3SG:DEF because this much
 for sharing that, because it’s worth more than,
minden riportnál,
 everything report:ADE
 many reports
meg kérdésnél, meg beszélgetésnél többet
 and question:ADE and conversation:ADE more:ACC
ért.
 be.worth:PST.3SG
 questions, or conversations.’
- Tel.: *Na jó. A dolog tehát így megy*
 DM good the thing therefore like.this go.3SG
- Call: **‘Well, all right.** So, this is how things go
Magyarországon,
 Hungary:SUPRESS
 in Hungary,
sok helyen, sok társaságnál, nyilván nem
 many place:SUPRESS many company:ADE obviously not
mindegyiknél.
 each.one:ADE
 at many places, in many companies, though of course not every-
 where.’

A further point concerns emotional expressions. Although both *no* and *na* are discourse markers of interjective origin, emotional expressions are not dominant in the spoken press subcorpus. This is probably related to the register itself, which tends toward more formal

and publicly oriented language use. Emotional or evaluative overtones nevertheless remain observable in certain combinations, such as *na persze* ‘well of course’ and *na mindegy* ‘well never mind’, and in evaluative commentaries.

5. Conclusions

The present study examined the variation between Hungarian *no* and *na* in two corpora representing present-day Hungarian: spontaneous conversations from the BEA database and the MNSz2 gigacorpus. Since the regional subcorpora of the MNSz2 differ in size and composition, the regional patterns observed in the data must be interpreted with caution. The data suggest that the present-day distribution of *no* and *na* cannot be reduced to a simple regional contrast. Most regional differences are small, with *na* being more frequent in the Subcarpathian subcorpus. At the same time, the absence of clearly identifiable *no* tokens from the BEA conversations and the high frequency of *na* in the spoken press subcorpus of MNSz2 suggest that *na* has a strong presence in the spoken data examined here, especially in the spoken press material.

With regard to register, the clearest finding concerns *na* in the spoken press subcorpus, where the association is supported by a meaningful effect size. Other register differences are weaker, and *no* does not show a similarly clear register-related pattern. The register findings therefore do not support a broad opposition between spoken or speech-like *na* and written *no*. Rather, they point to a specific concentration of *na* in the spoken press data, alongside smaller differences elsewhere.

The functional analysis of the spoken data¹⁴ showed substantial positional overlap but more visible functional differentiation. In the BEA conversations *na* occurred predominantly in turn-internal position, with responsive uses also present. In the spoken press subcorpus of MNSz2, both *no* and *na* occurred mainly in turn-internal position, while

14 I thank the reviewers of the paper for their suggestion that, based on the examples, two main function groups seem to stand out: (1) evaluation in a broad sense (stance expression, emotional reaction), and (2) conversation management (introduction of a new development or closing of the conversation). I agree with this observation, although I believe that in some examples these functions cannot be clearly separated from one another, as the former may permeate the latter.

turn-final uses were rare. However, transition marking was considerably more prominent with *na*, whereas *no* showed a more varied functional distribution, including uses related to urging and disagreement.

The two forms also differ in their combinations with other discourse markers. Both *no* and *na* frequently occur in such combinations, but they tend to combine with different elements. In the spoken press data, *na* is strongly associated with *most* ‘now’, typically in transition-marking uses, whereas *no* more often occurs in recurrent combinations such as *no meg* ‘well and’, typically in additive continuations; *no és* ‘well and’, often in argumentative extensions; and *no de* ‘well but’, in contrastive contexts. These patterns indicate that the two forms are not fully interchangeable, even though their functional overlap remains substantial.

Finally, the findings suggest that the coexistence of *no* and *na* in present-day Hungarian is linked to both overlap and differentiation. Their shared discourse-organizing roles make alternation possible, while their differences in discourse-pragmatic use and preferred combinations point to partial specialization. The register results add a more cautious picture: they show a clear concentration of *na* in the spoken press subcorpus, but not an equally clear register-related pattern for *no*. Thus, the greater prominence of *na* in the spoken data examined here does not amount to the complete replacement of *no*, but rather to a reorganization of the relationship between the two discourse markers.

Abbreviations

1 – first person, 2 – second person, 3 – third person, ACC – accusative, ADE – adessive, ADJ – adjectivizer, ALL – allative, CAUSFIN – causal-final, COND – conditional, DAT – dative, DEF – definite conjugation, DEL – delative, FS – false start; HES – hesitation, ILL – illative, IMP – imperative, INE – inessive, INF – infinitive, INS – instrumental, ORD – ordinal, PL – plural, POSS – possessive, POT – potential, PROH – prohibitive, PTCL – particle; PTCP – participle, PST – past, QPTCL – question particle, SBJV – subjunctive, SG – singular, SUBL – sublative, SUPESS – superessive, TEMP – temporal, VPFX – verbal prefix.

Sources

- BEA = *Beszélt Nyelvi Adatbázis* [Hungarian Spontaneous Speech Database]. <https://fonetika.nytud.hu/bea-beszelt-nyelvi-adatbazis/?lang=en>
- CzF. = Czuczor, Gergely & Fogarasi, János (ed.). 1867. *A magyar nyelv szótára IV. kötet* [Dictionary of the Hungarian Language, Vol. IV]. Pest: Emich Gusztáv Magyar Akadémiai nyomdásznál. <https://www.arcanum.com/hu/online-kiadvanyok/Lexikonok-a-magyar-nyelv-szotara-czuczorfogarasi-55BEC/>
- ÉrtSz. = MTA Nyelvtudományi Intézete (ed.). 1961. *A magyar nyelv értelmező szótára V. Mo– S.* [Comprehensive Dictionary of the Hungarian Language, Volume V (Mo–S.)]. Budapest: Akadémiai. <https://www.arcanum.com/hu/online-kiadvanyok/Lexikonok-a-magyar-nyelv-ertelmezo-szotara-1BE8B/>
- KED = *Középmagyar Emlékirat- és Drámakorpusz* [Middle Hungarian Memoir and Drama Corpus] <https://ked.nytud.hu/#open>
- MNSz2 = *Magyar Nemzeti Szövegtár* [Hungarian Gigaword Corpus 2.] <http://clara.nytud.hu/mnsz2-dev/>
- MTSz = *Magyar Történeti Szövegtár* [Hungarian Historical Database]. <http://clara.nytud.hu/mtsz/>
- SchlGl. = Szamota, István. 1894. *A Schlägli magyar szójegyzék* [The Schlägli Hungarian Glossary]. Budapest: MTA.
- SzabV. = Imre Samu. 1958. *A Szabács viadala* [The Battle of Szabács]. Budapest: Akadémiai Kiadó.
- ÚESz. = Gerstner, Károly (ed.). 2011–2025. *Új Magyar Etimológiai Szótár* [New Etymological Dictionary of Hungarian]. MTA Nyelvtudományi Intézet/HUN-REN Nyelvtudományi Kutatóközpont. <https://uesz.nytud.hu/index.html>
- ÚMTSz. = B. Lőrinczy Éva (ed.). 2002. *Új magyar tájszótár IV. N–S* [New Hungarian Dialect Dictionary Vol. IV. N–S]. Budapest: Akadémiai Kiadó.

References

- Auer, Peter & Maschler, Yael. 2016. The family of NU and NÁ across the languages of Europe and beyond: Structure, function, and history. In Peter Auer & Yael Maschler (eds), *NU/NÁ. A family of discourse markers across the languages of Europe and beyond*, 1–47. Berlin/Boston: De Gruyter.
- Balázs, Judit. 1995. A mondat szerkezete [The structure of the sentence]. In Loránd Benkő (ed.), *A magyar nyelv történeti nyelvtana II/2. A kései ómagyar kor. Mondattan, szöveggrammatika* [Historical Grammar of the Hungarian Language II/2. The Late Old Hungarian Period. Syntax and Text Grammar], 43–120. Budapest: Akadémiai Kiadó.

- Dér, Csilla Ilona. 2026. *A no~na diskurzusjelölők története az ómagyar kortól napjainkig* [The diachronic development of the discourse markers *no ~ na* from Old Hungarian to the Present]. Manuscript, to appear In Katalin Gugán & Mónika Varga (eds), *Regiszterfüggő nyelvi változók a középmagyarban* [Register-dependent linguistic variables in Middle Hungarian], 137–165. Budapest: Akadémiai Kiadó.
- Erdei, Tamás. 2021. *A na* diskurzusjelölő korpuszalapú vizsgálata. In Éva Basch, Tamás Erdei, & Valéria Juhász (eds), *Határtalan nyelvészet alkalmazásban, határtalan alkalmazás a nyelvészetben 2* [Boundless linguistics in application, boundless application in linguistics 2], 108–125. Szeged: SZTE JGYPK Magyar és Alkalmazott Nyelvészeti Tanszék.
- Gallasy, Magdolna. 1992. Az indulatszók [Interjections]. In Loránd Benkő (ed.), *A magyar nyelv történeti nyelvtana II/1. A kései ómagyar kor. Morfematika* [Historical Grammar of the Hungarian Language II/1. The Late Old Hungarian Period. Morphematics], 839–852. Budapest: Akadémiai Kiadó.
- Gósy, Mária, Dorottya Gyarmathy, Viktória Horváth, Tekla Etelka Grácz, András Beke, Tilda Neuberger & Péter Nikléczy. 2012. BEA: Beszélt nyelvi adatbázis [BEA: Hungarian Spontaneous Speech Database]. In Mária Gósy (ed.), *Beszéd, adatbázis, kutatások* [Speech, database, research], 9–24. Budapest: Akadémiai Kiadó.
- Heine, Bernd. 2023. *The grammar of interactives*. Oxford: OUP. <https://doi.org/10.1093/oso/9780192871497.001.0001>.
- Horváth, Viktória, Krepisz, Valéria, Gyarmathy, Dorottya, Hámori, Ágnes, Bóna, Judit, Dér, Csilla & Weidl, Zsófia. 2019. Háromfős társalgások annotálása a BEA-adatbázisban: elvek és kihívások [Annotating three-person conversations in the BEA database: principles and challenges]. *Nyelvtudományi Közlemények* 115. 255–274.
- Lanstyák, István. 2023. *Magyar nyelvi váltakozások vizsgálata a Magyar nemzeti szövegtár segítségével* [Investigating Hungarian language alternations using the Hungarian Gigaword Corpus]. Pozsony: Szenci Molnár Albert Egyesület. https://fphil.uniba.sk/fileadmin/fif/katedry_pracoviska/kmj/Korpvalt.pdf
- Oravecz, Csaba, Váradi, Tamás & Sass, Bálint. 2014. The Hungarian Gigaword Corpus. In *LREC 2014 Proceedings*. Reykjavik: ELRA, 1719–1723. ISBN 978-2-9517408-8-4. <https://doi.org/10.63317/3vncys6frpn>.
- Pichler, Heike. 2013. *The structure of discourse-pragmatic variation*. Amsterdam/Philadelphia: John Benjamins. <https://doi.org/10.1075/silv.13>.
- Sawicki, Lea. 2016. The Polish multifunctional particle *no*. In Peter Auer & Yael Maschler (eds), *NU/NÁ. A family of discourse markers across the languages of Europe and beyond*, 81–103. Berlin/Boston: De Gruyter. <https://doi.org/10.1515/9783110348989-003>.
- van Kemenade, Ans & Links, Meta. 2020. Discourse particles in early English: Clause structure, pragmatics and discourse management. *Glossa: a journal of general linguistics* 5(1). 3. 1–23. <https://doi.org/10.5334/gjgl.1020>

Kokkuvõte. Csilla Ilona Dér: **Tänapäeva ungari keele diskursusemarkerite *no* ~ *na* diskursusevariatsiooni analüüs.** Ajalooliselt interjektsioonidest tuletatud ungari diskursusemarkereid *no* ja *na* 'no, noh' on pikka aega käsitletud funktsionaalselt samaväärsete variantidena, hoolimata nende suhtelise staatuse pöördumisest: *no* oli varem domineeriv, samas kui *na*-d peeti regionaalseks. Käesolev uurimus vaatleb, mil määral kujundab nende tänapäevast vaheldumist regionaalne ja registripõhine varieerumine, samuti funktsionaalne spetsialiseerumine. Tuginedes BEA korpuse spontaansetele vestlustele ja Ungari Gigacorpus 2 (MNSz2) suulise ajakirjanduse andmetele, analüüsitakse nende esinemissagedusi, positsioone, funktsioone ja koosinemise mustreid diskursusevariatsiooni raamistikus. Tulemused näitavad vaid piiratud regionaalset eristumist, välja arvatud *na* suurem sagedus Taga-Karpaatia alamkorpuses, ega toeta laiaulatuslikku vastandust suulises keeles levinud *na* ja kirjaliku *no* vahel; pigem näitab *na* selget kontsentratsiooni suulises ajakirjanduses. Kuigi need kaks markerit esinevad sageli sarnastes positsioonides ja neil on kattuvad diskursuse funktsioonid, on *na* tavaliselt seostatud üleminekutega, samas kui *no* on kasutatud sagedamini ärgitamise, lahkarvamuse ja argumenteeriva jätkamise puhul, mis viitab osalisele funktsionaalsele eristumisele.

Märksõnad: korpuslingvistika, diskursuse marker, interjektsioon, registri- vahetus