

# Teaching Constructions in the Language Classroom: Hungarian Modal Existential wh-constructions

Mónika Dóla<sup>1</sup>, Zsolt Prohászka<sup>1</sup>, Anita Viszket<sup>1</sup>, Veronika Szabó<sup>1,2</sup>

University of Pécs<sup>1</sup>, University of Vienna<sup>2</sup>

(<dola.monika@pte.hu>, <prohaszka.zsolt@pte.hu>,  
<viszket.anita@pte.hu>, <szabo.veronika@pte.hu>)

## *Abstract*

This paper investigates the teaching of constructions (Goldberg 2006) through the example of the Hungarian modal existential wh-construction (MEC). After describing the schematic MEC for form and function, we present corpus data from Old to Modern Hungarian to show how the frequency and productivity of the construction increased over time, and how the pattern became associated with particular lexical expressions. Our analysis shows that while the abstract MEC-template associated with a general default meaning is highly productive, lexically, and grammatically fixed MEC-expressions are associated with distinct meanings, and they display a higher degree of invariability. Based on this, we propose to place various MECs along a scale of idiomaticity (Michaelis 2017) – arguing that in language teaching, all MEC-types must be paid attention to. Finally, we suggest ideas for the teaching of various MEC-constructions in the language classroom with the help of corpora and follow-up activities.

## *Keywords*

constructions; corpora; Hungarian; language teaching; modal-existential wh-construction

## 1. Introduction

Linguistics has offered various models for how meaning is created in language, and the different trends in language teaching throughout the years have adapted these models to facilitate the learners to create meaning in the target language. In the 21<sup>st</sup> century, usage-based linguistics have suggested



that since «abstract representations of linguistic structure are derived from language users' experience with concrete linguistic tokens, grammatical patterns are generally associated with particular lexical expressions», and also that «frequency strengthens the representation of linguistic elements in memory [...], which in turn can have long-lasting effects on the development and organization of the linguistic system» (Diessel 2017, 1). By making it possible to analyze large amounts of authentic language data quickly and systematically, corpus linguistics has offered ample evidence for this lexical-grammatical patterning in language, and it has also provided statistical tools to measure the frequency of occurrence related to linguistic patterns. Consequently, it has gained increasing popularity in language teaching to pay great attention to patterns (constructions, multiword units, formulaic language etc.) and to use corpus-based methods and tools.

Corpora and corpus-based research have a wide range of language-pedagogical applications from corpus-based dictionaries and grammars to corpus-based textbooks, data-driven learning, and learner corpora (Mukherjee 2016). Most studies related to corpus-supported language learning and teaching focus on more commonly taught languages (Keck 2014), especially on English; however, in recent years, corpus-research, corpus-building, and corpus-supported teaching have also gained increasing popularity in the teaching of less commonly/widely taught languages (Preradović, Posavec, Unić 2019) – such as that of Hungarian. We now have several freely available high-quality annotated corpora and various natural-language processing tools (<<http://corpus.nytud.hu/nkp/>>), and even the building of learner and pedagogical corpora has started (Baumann *et al.* 2020). Nonetheless, corpus-related studies are still rare in Hungarian language teaching, and there is a general shortage of corpus-informed Hungarian as a second/foreign language resources (e.g., dictionaries – Sass *et al.* 2011 – grammars, textbooks, classroom applications and activities).

The present paper taps into two language-pedagogically relevant applications of corpora – corpus-based research for materials to be used in Hungarian language teaching, and as a tool in the classroom itself – through one specific example: the learning and teaching of the Hungarian modal existential *wh*-construction (MEC). After introducing the Hungarian MEC, we address the issue of how corpus-based and corpus-informed research can

inform the description of this specific grammatical construction. First, we describe the schematic modal existential wh-construction as a conventional form-meaning pairing. Then, based on our corpus-research, we identify its most frequent patterns and instances, which we analyze for their fixedness, regularity, transparency, and situation/function/genre-boundedness, and we propose to place the various types of the MEC along a possible scale of idiomaticity from hpaxes (with the ‘default’-meaning) to fully fixed idiomatic MECs. A pilot study with native speakers shows that these two endpoints are relatively easy to identify. Finally, we put forth ideas and ways for the corpus-assisted learning of MECs in the classroom: we suggest various data-driven learning activities that we expect, on the one hand, to increase the learners’ language awareness by making them explore the patterning of the MEC in Hungarian, and, on the other hand, to enable them to use the targeted construction ‘flexibly’ but satisfying the necessary constraints of the given construction – considering that Hungarian is an agglutinative language.

We argue that – with sufficient training and assistance – both teachers and students can benefit from the use of corpus-based tools and methods in Hungarian language teaching. Teachers can develop materials to demonstrate a construction’s semantic prosody and its most natural, native-like, and common instances, and learners can discover and observe constructions together with their most frequent lexical-grammatical realizations. If grammatical constructions such as the one discussed in the present paper are approached with corpus-based tools and methods in Hungarian as a second/foreign language, not only do we provide our learners with authentic data that they can use to learn ‘real’ language, but we also facilitate their learning by making them attentive to language patterns and their degree of idiomaticity (see later).

## 2. The Hungarian Modal Existential Wh-Construction

In Construction Grammar (Goldberg 1995, 2006, 2019), grammatical constructions are seen as learned and stored pairings of form and function. This means that particular abstract grammatical patterns evoke particular semantic representations on their own, i.e., irrespective of the words they include. In this sense, the Hungarian modal existential wh-construction

(MEC) can be defined as a pairing of a relatively fixed unique syntactic template (of a main existential predicate, a *wh*-pronoun, and a subordinated verb in the infinitive or in the subjunctive form), and a schematized representation of the situation that the possibility of a proposition is available or not (to someone) (for details, see below).

The term *modal existential wh-construction* highlights the three universal properties that characterize MECs in all the languages<sup>1</sup> that have them and which ones, at the same time, distinguish them from other related constructions (Šimík 2011): they express modality, they are embedded under existential predicates, and they obligatory contain fronted *wh*-words. A general property of the Hungarian MEC is the presence of three main elements strongly associated in a strict syntactic template. Most typically, these elements occur in a specific order: the main predicate is followed by a *wh*-pronoun and a subordinated verb.

The MEC as a grammatical construction has a specific function: it expresses that the possibility of a proposition is available or not (examples 1a-b). The proposition is expressed by the *wh*-pronoun and the subordinate verb, and its availability is expressed by the main predicate. Circumstantial modality is expressed by the construction as a whole: there is no overt modal element in the construction, the MEC itself as an abstract grammatical construction introduces a modal basis with a temporal perspective for the proposition. In (1a), the possibility to eat something at a given point of space and time is not available, in (1b), it is.

- (1) a. Nincs<sup>2</sup> mit enni.  
       'there-is-no what.Acc eat.Inf'  
       'There is nothing to eat. (We have no food to eat.)'  
       'There is nothing we can/could eat./We have no food we  
       can/could eat.'

<sup>1</sup> Šimík identified twenty-seven languages that have the MEC. (For a detailed cross-linguistic analysis, see Šimík 2011.)

<sup>2</sup> To negate the present tense substantive verb *van* 'there-is', the complex negative verb form *nincs* 'there-is-no' is used.

- b. Van mit enni.  
 'there-is what.Acc eat.Inf'  
 'There is something to eat. (We have food to eat.)'  
 'There is something we can/could eat./We have food we  
 can/could eat.'

Cross-linguistically, MECs occur with predicates whose semantic profile contains an existential component: the verbs' lexical meaning «supports existential quantification over their indefinite internal arguments» (Šimík 2011, 31). The verbs may assert the existence of something/someone or that of a proposition, or they may express availability or coming into being (ivi). MECs are related to other existential constructions through the presence of an existential verb. Languages may use copular verbs, possessive or locative expressions, or expletive/impersonal constructions to express existentiality (McNelly 2016). To this end, Hungarian primarily uses constructions with the verb *van* 'there-is' – both for the expression of existence, and for possession. Hungarian belongs to those languages where existential and possessive constructions are closely related due to the absence of a *have*-construction: in Hungarian, *van* is used in both expressions. In addition, *van* also features in referential and predicative sentences – as a copular verb (*to be; is*).

The main predicate of the Hungarian MEC is typically an existential verb, most often a form of the substantive verb *van* 'there is'. This verb takes different forms based on its mood and tense: *volt* 'there was', *legyen* 'there should be' (imperative), *lenne* 'there would be' (conditional), *lesz* 'there will be', *lehet* 'there can be', *lett volna* 'there would have been' (past conditional), and the negated forms of these can also appear in the MEC. Lipták (2003) claims that the existential verb *akad* 'occur' can also feature as the main predicate of the construction, however, such occurrences are rare.

The main existential predicate of the MEC can only appear in third person as it states the (non-)existence of a proposition (expressed by the other elements) (2a-e). However, person marking is possible by adding a dative-marked pronoun (2c), or by the conjugation of the subordinate verb (illustrated in 2d to be discussed in more detail later). In the first case (i.e., when a dative marked pronoun is present, and the subordinate infinitive is not inflected, as in 2c), the meaning of the construction is somewhat

ambiguous: it can either mean ‘You have something to say’ or ‘There is something to be said to you’ – the ambiguity can be resolved by using both the explicit dative marked pronoun and the inflected infinitive, as in (2e), meaning ‘I have something to say to you’.

- (2) a. Van/Nincs mit mondani.  
       ‘there-is/there-is-no what.Acc say.Inf’  
       ‘There is nothing to say.’
- b. Volt/Nem volt mit mondani.  
       ‘there-was/there-was-no what.Acc say.Inf’  
       ‘There was nothing to say.’
- c. Noked van mit mondani.  
       ‘you.Dat there.is what.Acc say.Inf’  
       ‘You have something to say./There is something to be said to you.’
- d. Van mit mondanom.  
       ‘there-is/there-is-no what.Acc say.Inf’  
       ‘I have something to say.’
- e. Noked van mit mondanom.  
       ‘there-is/there-is-no what.Acc say.Inf’  
       ‘I have something to say to you.’

Besides the existential *van* ‘there is’ (and its different forms), there are a few other verbs which can introduce a MEC: *tud* (in MECs meaning ‘can/be able to’) is also a productive MEC-predicate (3a). The verbs *bír* (which is used synonymously with *tud* in MECs, meaning ‘can/be able to’) and *mer* ‘dare’ can also occur in the construction, although they are rare (3b-c). MECs are most frequently used as negative constructions, that is, the main predicate usually occurs under negation.

- (3) a. Nem tudok mit mondani.  
       ‘no can.Sg1Indef what.Acc say.Inf’  
       ‘There is nothing I can say.’

- b. Nem bírok mit mondani.  
'no bear.Sg1Indef what.Acc say.Inf'  
'There is nothing I can say.'
- c. Nem merek mit mondani.  
'no dare.Sg1Indef what.Acc say.Inf'  
'There is nothing I dare to say.'

The second element of the construction is the *wh*-word. Although these are formally interrogative, in the MEC they obtain an indefinite-pronoun like character (Prohászka *et al.* 2022). Prohászka discusses the different functions of the (formally) interrogative *wh*-pronouns in Hungarian. He claims that the *wh*-word of the MEC is a non-specific, non-referential element, which is similar in meaning to those indefinite pronouns which take the *vala*- 'some' morpheme (4a-b). It must be noted that these *vala*-pronouns can never occur in a MEC (4c).

- (4) a. Van kit meglátogatnom.  
'there-is who.Acc vp-visit.Inf.Sg1'  
'I have someone/people to visit.'
- b. Van valaki, akit meglátogathatok./Valakit meglátogathatok.  
'there-is someone who.Acc vp-visit.can.Sg1/someone.Acc  
vp-visit.can.Sg1'  
'There is someone who I can visit./I can visit someone.'
- c. \*Van valakit meglátogatnom.  
'there-is someone.Acc vp-visit.Inf.Sg1'

While most *wh*-words can occur in the MEC, there are three exceptions. The reason why *melyik* 'which', *hány* 'how many' and *mennyi* 'how much' cannot appear in the MEC is because they are referential elements (*melyik* is also specific).

The final element of the MEC is the subordinate verb, which is usually the infinitive form. While Šimík (2011) claims that the subordinate verb can also be in the subjunctive, a small-scale study of Prohászka (2022) suggests that the subjunctive is mostly used by speakers who use dialectal variants of Hungarian.

A subordinate verb and a *wh*-element (in a given case) can only be co-selected for a MEC if they can occur together in a verb-argument structure construction.

As previously mentioned, person can be marked in the MEC either by adding a dative-case element (e.g., a pronoun, see above) or by having an inflected infinitive. This is clearly possible in those constructions that use finite subjunctive verbs for subordination (illustrated in 5a) since those are obligatorily conjugated. But person-suffixation is not only present in subjunctives: Hungarian infinitives can also be inflected, even in the MEC (see 5b). We must note, however, that the results of the previously mentioned small-scale study of Prohászka (2022) show that most frequent MECs host an uninflected infinitive.

- (5) a. Van hova menjek.  
       'there-is where-to go.subj.Sg1'  
       'I have a place to go.'
- b. Van hova mennem.  
       'there-is who.Acc vp-visit.Inf.Sg1'  
       'I have a place to go.'

In the current paper, our focus is on MECs that host the subordinate verb in the infinitive form.

Although most instances of the MEC follow the same word order of the three main elements (that is: main predicate–*wh*-word–subordinate verb), some tokens display word orders different from the relatively strict pattern. Left dislocation of the infinitive verb is possible, as shown in (6a). Rarely, the MEC structure can also be interrupted by additional elements – such as in (6b) where the dative-marked subject occurs between the main predicate and the *wh*-word. Sluicing is also possible, as presented in (6c) (Šimík 2011).

- (6) a. Menekülni vizsont nincs hova.  
       'escape.Inf but there-is-no where-to'  
       'However, there is nowhere to escape.'



- b.      Nem volt Péternek mit csinálnia.  
          ‘no there-was Peter.Dat what.Acc do.Inf.Sg3’  
          ‘Peter had nothing to do.’
  
- c.      A férfi nem is dolgozik, mert nincs hol.  
          ‘the man no also work.Sg3 because there-is-no where’  
          ‘The man doesn’t work, either, because he has nowhere to  
          do so.’

In sum, the MEC is an abstract grammatical construction (Figure 1) which involves a relatively fixed unique syntactic template of an existential construction with a main predicate and a subordinated verb–argument structure construction of a wh-pronoun and a subordinate verb which appears in the infinitive or in the subjunctive form. Person marking is possible by conjugating/inflecting the subordinate subjunctive/infinitival verb and/or by adding an optional dative-marked element. The construction introduces a circumstantial modal base for the content expressed by its component parts. The MEC is associated with the default semantic content of the possibility of a proposition being available or not (to someone). Since several aspects of its form and function are «not strictly predictable from its component parts or from other constructions recognized to exist» (Goldberg 2006, 5), the MEC must be regarded as a learned pairing of form and function.

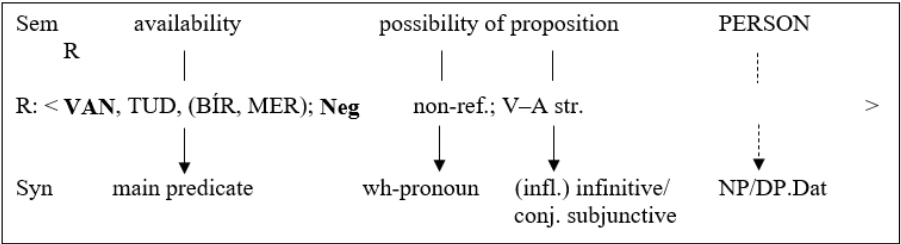


Figure 1<sup>3</sup> – The abstract grammatical MEC-construction<sup>4</sup>.

<sup>3</sup> This general template of the abstract grammatical MEC-construction was prepared by the authors following Goldberg’s (1995) representations of various constructions.

<sup>4</sup> VAN ‘there-is’; TUD ‘can’; BÍR ‘bear’; MER ‘dare’. The English translations of the relevant items in the figures are given in the footnotes.

### 3. MECs in Hungarian corpora

In this section we present our MEC-related corpus findings. If Hungarian native speakers – including language teachers – were asked to list typical and frequent infinitival constructions in Hungarian, they would probably mention various modal auxiliary + infinitive constructions but the list would possibly not include the MEC. However, if we consult a corpus of modern Hungarian, we can see that MECs are within the first 20 most typical infinitival constructions in Hungarian (there are more than 1500 verbs with infinitival arguments, see Szabó 2020). Native speakers' intuition about the frequency of a construction does not always overlap with the statistical data (see Reppen 2010), and it follows that neither course content developers, nor language teachers can rely solely on their language competence when selecting grammar points for their curricula.

By consulting modern corpora, however, we can identify the most frequent constructions of contemporary language use, and it can also be interesting to study historical corpora to understand the formal and functional features of given constructions in depth. Our corpus analysis in this paper is based on a predefined construction, the MEC, which contains a matrix predicate, a *wh*-word and an infinitival or subjunctive element.

#### 3.1. A brief history of MECs

We worked with two historical corpora to study the MEC and its possible constructional changes over the centuries: the *Ómagyar korpusz* (Old Hungarian Corpus) (<<http://omagyarkorpusz.nytud.hu/>>, cf. Simon, Sass 2012) and *A Történeti magánéleti korpusz* (TMK) (The Old and Middle Hungarian corpus of informal language use) (Novak *et al.* 2017-, Dömötör *et al.* 2017, Novak *et al.* 2018).

It must be noted that although, in general, corpora provide a large amount of data to examine real language use, we had to exercise great caution when interpreting the findings drawn from the *Old and Middle Hungarian corpora*. Since they are not balanced and only cover a narrow range of text categories and genres, they cannot be considered representative of the language use of the given period (Simon 2019). A further problem presented itself regarding

the discussion of the relative formal flexibility of MEC constructions, which was due to search-related challenges: it is hard to define a good query expression on the different query surfaces. For this reason, our database cannot be totally exhaustive. Nevertheless, it can be used to establish tendencies in the use of MECs across the different stages of Hungarian.

The results of the queries show that the MEC was already present in Old Hungarian (15th century); based on this, it can be considered a typical Hungarian infinitival construction. The construction with some universal formal characteristics appears in many, often typologically and genetically different languages, for example, in Slavic languages (Šimík 2011), which were in language contact with Hungarian. Consequently, it cannot be stated with certainty if MECs were present in Proto-Hungarian or if they were later borrowed from Slavic languages. From the point of view of language teaching, however, where students' native languages can/should be considered, it is interesting to note that students with a Slavic language as their mother tongue can be expected to be familiar with the construction, while for learners with a different linguistic background, the MEC may be new.

In our data from Old Hungarian (896-1526), the main predicate is mostly the substantive verb, but the modal verb *tud* 'can' also appears 7 times. The occurrence of this modal auxiliary is exceptional in MECs across languages; it seems to be specific to the Hungarian MEC. The subjunctive-infinitive ratio is different from today's language use: we find more subjunctive-type MECs than infinitival ones. It is also curious that at the time, the infinitive could also be inflected next to the already conjugated modal verb, which is not allowed in modern Hungarian. However, this was typical of other infinitival constructions of the time, as well [e.g., infinitives with a modal verb as predicate (Dékány 2014)]. The three compulsory elements (the matrix verb, the *wh*-word, and the infinitive) appear fixed in the data, and the core meaning of the MEC is the same as in Modern Hungarian: it expresses that the availability of a proposition is possible or not. They appeared mainly in denials.

Since Old Hungarian texts are Bible translations, the most typical examples express poverty, or the lack of basic human needs. In addition to such instances, we found several hapaxes, which demonstrates the productivity of the MEC in the period (e.g., see example 7 from the Munich Codex – 1466).

- (7) emberfiának kedig nincs hova fejét hajtani #276860<sup>5</sup>  
 'son\_of\_man.Dat and there-is-no where-to head.Poss.Sg3.Acc lay.Inf'  
 'And the son of man does not have a place to rest his head.'

We can see that the first recorded function of the MEC was the central canonical meaning.

As for Middle Hungarian, we had private correspondences to study for the use of the MEC, so it came as no surprise that the MECs we found here were embedded in new contexts. By this time, the MEC had become a frequent way to describe one's inability to do something. We found, for instance, a great number of expressions which were used to express the inability to further elaborate or say anything new on a topic (8).

- (8) A doktor felől most sem tudok kegyelmednek mit írnom #1064046  
 'the doctor about now also-no can.Sg1 excellency.PossSg2.Dat  
 what.Acc write.Inf.Sg1'  
 'There is still nothing that I could write about the doctor to your excellency.'

As for form, the ratio between the *van/nincs* 'there-is(-no)' and the *nem tud* 'cannot' type predicates was like the one in the Old Hungarian sample, but the number of the infinitival type examples increased at the expense of the subjunctive-type MECs. In addition, we found idiomatic MECs, as well, in the Middle Hungarian sample. The personal letters of the time frequently made use of the MEC not only to describe the inability to do something in a practical sense, but also to express the incomprehensibility of a situation, and the inability to change the situation. Expressions such as *nincs mit tennem* ('there-is-no what.Acc do.Inf.Sg1') meaning 'there is nothing I can do about it' can be found 21 times in the sample.

### 3.2. MECs in Modern Hungarian corpora

We used two corpora to study the MEC in modern Hungarian. One of them was the *Hungarian Historical Corpus* of 27 million tokens coming from

---

<sup>5</sup> The #numbers are the corpus identification numbers.

different text types and genres (Sass 2017). This corpus is a collection of texts written between 1772 and 2010 (New Hungarian), where 40% of the texts are dated from the second half of the 20th century. The other corpus we used here is the *Hungarian Gigaword Corpus* of 1.5 billion tokens (Oravecz, Váradi, Sass 2014), which is a collection of texts written in the 20th century. Most texts belong to the press genre, but the corpus also contains other subcorpora such as (transcribed) spoken language samples.

In our corpus search on modern Hungarian, we found MECs from Facebook comments to newspaper headlines. In the *Hungarian Gigaword Corpus*, we found over 50,000 MECs. Our queries revealed that in modern Hungarian, the MEC is a very productive pattern and is not limited to one or two infinitive types.

Our data show that similarly to other periods, the vast majority of MECs stands in the negative form. Compared to the previous periods, new verbs appear as the main predicate of the construction (besides the substantive verb and the modal verb *tud* 'can'): e.g., *talál* 'find', *lehet* 'possible', *bír* 'bear' (i.e., cannot stand), *győz* 'win' (i.e., cannot stop) and *mer* 'dare'. Some further examples include modal verbs such as *kell* 'must', *akar* 'want to' and *szeretne* 'would like to'. Instead of regarding these rare examples as performance errors, we consider them as evidence for analogy which is often claimed to account for «the productive use of language (...) as well as certain types of language change » (Diessel 2017, 14). We assume that although MECs were and are mainly existential in nature, idiomatic wh-word + infinitive combinations have begun to live an independent life and can now also appear in modal verb + infinitive constructions. This process is likely to be facilitated by the formal and functional similarities between the infinitival modal existential wh-construction containing a substantive verb and the highly frequent *tud* 'can' modal verb + infinitive construction (see Goldberg 2019). The idea is supported by the fact that Old Hungarian MECs already contained the modal verb *tud* 'can', next to which the infinitive could also appear in an inflected form, which we do not see in modern and new Hungarian. When teaching the MEC today to learners of Hungarian, the two basic, prototypical main predicates – i.e., the substantive verb and the modal verb *tud* 'can' – must be presented, while the other above-listed predicates, due to their very low frequency, can be left out of the curriculum.

Modern Hungarian MECs have a more flexible syntactical structure than they had in earlier periods: in 80% of the cases, the structure is very strict, however – also as Lipták (2003) notes – certain elements such as the dative-marked subject can occur before the *wh*-item. In around 10% of the examples, we found discourse particles, adverbs and pronouns interrupting the [main predicate + *wh*-word + subordinate verb] sequence.

#### 4. MECs and idiomaticity

Although it may not be evident to all participants in second/foreign language teaching, idiomaticity is very important to consider in the teaching and learning of grammatical constructions. Learners should not only be familiarized with an abstract construction's grammatical buildup and its basic, central function/meaning, but they should also become aware of those very frequent instances of the construction where the given structural form is very strongly associated with particular lexis. If lexical elements are frequently co-selected in a given grammatical pattern, they gradually clump and forge together, and often take on a distinct meaning. The parts may (partly) lose their independent meanings, and the meaning of the whole may not be inferred compositionally. The holistic function/meaning of the expression will put constraints on the pattern, which will result in less variability or even regularity than what is characteristic of the schematic grammatical pattern. This idiomaticization can happen to various degrees – from semi-preconstructed phrases to fixed expressions, phraseological units, and idiosyncratic phrases (Sinclair 1991, 2008). As a result, learners must also be made aware of the frequency, productivity, and idiomaticity of various instances of a construction.

A salient property of the MEC in modern Hungarian is that several highly frequent instances in the database seem more idiomatic in nature than the canonical MEC – some to a lesser, some to a greater degree. This conventionalization of certain MEC-expressions is the main constructional change that can be observed over time regarding the MEC (Szabó, Prohászka 2021). But how idiomatic exactly can a MEC be?

Usage based linguistics argues that idiomaticity is not a clear-cut issue but rather a scalar phenomenon. As Michaelis (2017) puts it, «every pattern of language, from the fixed formulas to the fully productive phrase-structure rules» can be placed on a scale with vague boundaries between the types along the idiomaticity-continuum (Figure 2).

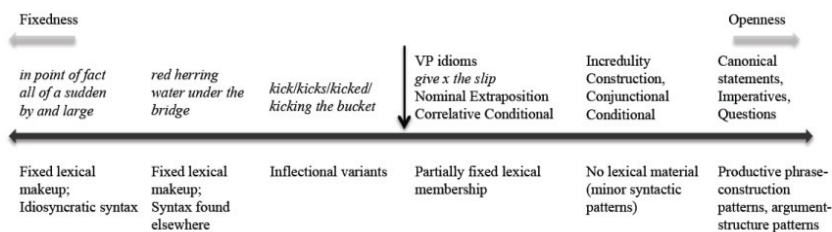


Figure 2 – Scale of idiomaticity (Michaelis 2017).

In fact, the canonical MEC itself can be regarded as idiomatic in the sense that its formal aspects such as the arrangement of and the restrictions on the elements and the function of the construction, i.e., the fact that it expresses the possibility of a proposition being available or not (to someone) cannot be predicted from its component parts or from other constructions (see the Goldbergian definition of constructions). As we have seen in Sections 2 and 3, however, the canonical MEC is a productive pattern which allows for a relatively high degree of formal variability, so this could stand at the more open end of the spectrum (cc. productive phrase-construction patterns, see Figure 2). In the Hungarian Gigaword Corpus, there are over 50,000 thousand MEC-examples. It must be noted though that due to the high variability of the construction, corpus-query on the MEC in general is relatively difficult.

Nonetheless, the meanings of certain very frequent instances of the MEC cannot be predicted from the canonical MEC construction, and many highly frequent MECs allow less formal variability than the canonical MEC. Let us consider here six such constructions along various points of the idiomaticity continuum, proceeding from the most open to the most fixed MEC.

#### 4.1. Types of idiomatic MECs

The first expression is *Ezzel nem tudok mit kezdeni* (9). This is a *TUD* ‘can’ type of MEC which hosts the [‘indefinite/wh-pronoun.Acc *kezd* ‘begin’ DP.Instr’] verb–argument structure construction.

- (9) *Ezzel nem tudok mit kezdeni.*  
 ‘this.Instr no can.Sg1Indef what.Acc begin.Inf’’  
 ‘I can’t do anything about/with this./I can’t handle this.’

The [*NEM TUD.Infl MIT KEZDENI DP.Instr*] is a lexically fixed construction with one open slot for the instrumental-case argument (cc. partially fixed lexical membership, see Figure 2). The construction allows for great formal variability. Although the main predicate can only stand in the negative form, it can be freely conjugated, and the instrumental-case definite noun phrase (or an equivalent pronoun) can stand after the infinitive, before the infinitive, or before the negative main predicate. Due to the open slot, it is difficult to find all the relevant examples in the corpus. However, we can affirm that this is the first most frequent MEC in the Hungarian Gigaword Corpus with nearly 8500 results (e.g., 10).

- (10) *A készen kapott szabadsággal a város lakói nem tudtak mit kezdeni.* #2698471  
 The city dwellers could not do anything with the ready-made freedom.’

At the same time, the verb–argument construction hosted in this MEC is idiomatic both in the sense that its verb *kezd* ‘begin’ requires an accusative and an instrumental case argument (in addition to the subject) where the accusative argument can only be a pronoun, and also in the sense that it has a non-compositional meaning: *valaki valamit kezd valamivel* ‘somebody begins something with something’ expresses the distinct meaning of ‘somebody does something about/with something’.

The second example is the ambiguous expression *Van/Nincs miből* (11), where one form (the MEC) may map onto two unrelated meanings. One meaning can be the canonical MEC meaning: in this case, the expression is a MEC where the subordinate verb (any verb that can take a noun in the



relative case as an argument) is deleted. The other meaning can be ‘can/cannot afford it’: in this case, we have a lexically fixed expression where a semantically specific subordinate verb is deleted – one that belongs to the semantic field of paying (*kifizetni* ‘to pay’, *finanszírozni* ‘to finance’, *megélni* ‘to live on’ etc.), and which now by itself means the concept of (not) having the resources to afford something.

- (11) Van/Nincs miből.  
 ‘there-is/-no what.Elat’  
 ‘Meaning 1: there is something/nothing to .... from’  
 ‘Meaning 2: can/cannot afford it’

The above-mentioned ambiguity is illustrated in (12).

- (12) Állnak a kocsmában, isznak, ha...van miből. #293768181  
 ‘They would be standing in the pub drinking if...there-is what.Elat’

If we wanted to complete the MEC in (12), we could choose two ways. Based on the previous context, we could say *ha van miből inni* (‘if there is anything to drink from’), or, based on our encyclopedic knowledge related to the pub-situation, we could say *ha van miből kifizetni* (‘if there is money to pay for it’). In fact, the expression *van/nincs miből* is so highly conventionalized in the ‘can(not) afford it’ sense, that even the former full MEC (*van miből inni*) can be interpreted as ‘if they can afford drinking’. This also supports the idea that the basic meaning of this construction is that of (not) having the resources to afford something.

With over 1,400 corpus hits, the substantive verb + *miből* (what.Elat) construction counts as a relatively frequent MEC. Although the most frequent version in the corpus data is the present tense negative *nincs* ‘there-is-no’ type, the main predicate shows a high degree of variability: it can be in the present (735 hits), in the past (364 hits), and in the future form (322 hits), and it can stand in the negative and in the positive form. Also, although MECs in general randomly feature the dative-marked subject, its appearance in this particular construction is more typical.

In sum, [VAN/NINCS.Infl MIBŐL] can be regarded as a lexically fixed expression with little inflectional variability (regarding the tense and mood of the main predicate), which despite does not contain a subordinate verb

and is thus associated with the distinct meaning of resourcing and payment (cc. inflectional variants, see Figure 2).

The third expression is *Nem tudja hova tenni*, (13). It is a polysemous construction where one form has two meanings that are related under the category of the MEC. One meaning is the central or basic canonical MEC-meaning, while the other one is a metaphorical extension of this prototypical sense. Viewing idiomaticity as a cline, we can say that the extended meaning is more idiomatic than the central meaning.

- (13) *Nem tudja hova tenni.*  
 no can.Sg3Def where-to put. Inf  
 Basic meaning: She/He has nowhere to put it/him/her.  
 Extended meaning: She/He cannot place him/her./She/He cannot make sense of it.

The two examples below (14-15) show the two meanings of the expression.

- (14) *Bútort nemigen tudnék hova tenni az Astrában (nem kombi).*  
 #113017672  
 I don't really have any space to put furniture in the Astra (it's not a station wagon).
- (15) *Amúgy az agresszivitást nem tudom hova tenni.* #1138796279  
 By the way, I can't understand aggression.

The subordinate verb can only be changed in the basic meaning (e.g., *to put*, *to place*, *to lay*, *to sit*, etc.), the expression associated with the extended meaning is lexically fixed.

If we want to do a corpus search for this construction, the adequate query expression is the string *hov.\* tenni* since there are two variants of the wh-word where-to in Hungarian: *hova* and *hová*. Over twice as many of the hits contain the *hova* wh-element (854 hits) than the *hová* variant (379 examples). Out of the 1,233 hits, 95% feature in the following pattern: NEM TUD.Infl HOVA/Á TENNI. This indicates that the string *hova/á tenni* is very strongly associated with this type of MEC (and not with the substantive verb type). Although this MEC is lexically fixed, it also shows conjugational variability.

ty: the main predicate *tud* 'can' can stand in any mood, tense, and person. In the data, the first-person singular form is the most frequent one (over 48%), followed by the third-person singular form (20%). The construction also allows its subordinate verb to appear in the subjunctive, although we only found 114 examples for the string *tud+hov.\*+tegy.\**. However, the main predicate is grammatically fixed in the sense that it can only stand in the negative: *nem tud* 'cannot'.

We have seen that [NEM TUD.Infl HOVA/Á TENNI] construction has a high degree of lexical and grammatical fixedness. In addition, an overwhelmingly large number of the query results express the extended, more idiomatic 'cannot make sense of it, do not understand' meaning. (N.B. We cannot do statistical corpus search for idiomaticity.) Based on this, we can propose that this particular construction is highly idiomatic (cc. fixed lexical makeup, syntax found elsewhere – with conjugational variants; see Figure 2), and that in the teaching of Hungarian as a foreign language, it should primarily be taught as a fixed expression meaning 'cannot understand something' / 'cannot recognize someone'.

The fourth construction is the expression *Nincs mit*, (16).

(16) 16. *Nincs mit*.

there-is-not what.Acc

Meaning: Not at all./You're welcome. (As a polite response to thanking)

This MEC is totally fixed both lexically and grammatically, but it contains syntactic structure and could thus be assimilated to the "syntax found elsewhere"-stage in the linear scale represented in Figure 2. It does not allow for any variability regarding either the main existential predicate *nincs* or the wh-pronoun *mit*, or the word order of these two elements. No subordinate verb can occur in the construction, yet it is implicitly understood that the unpronounced infinitive is (*meg*)*köszönni* 'to thank': *nincs mit* (\**megköszönni*); 'there is nothing to thank for'. The word string has fixed prosody, as well: the existential verb carries the stress, the wh-pronoun cannot be stressed. The construction with these idiosyncratic lexical and grammatical constraints has a unique function: it is used as a conventional response bound to the situation of thanking. In this sense, the expression

*Nincs mit* can be regarded as idiomatic, and it could be placed at the more fixed end of the continuum of idiomaticity.

The fifth example is the expression *Mit volt mit tenni*, (17).

- (17) *Mit volt mit tenni.*  
 what.Acc there-was what.Acc do.Inf  
 Meaning: There was nothing for it (but to...)

This MEC has irregular idiosyncratic syntax (see Figure 2), and it is open to very little variability. The lexical elements are fixed, and so is their order. The main predicate is always in the past form. Only the infinitive *tenni* ‘to do’ can be inflected. However, out of the 82 exemplars that we found in the Hungarian Gigaword Corpus, 75,6% contained the non-inflected infinitive, so this form can be assumed to be the prototypical one (13,4% of the results contained the subordinate verb in the first-person singular form, and only 8 examples were found to stand in the third-person singular form). The expression is often used in narration, especially in folktales, to express that someone had no other choice but to do what they finally did in the story. In the teaching of Hungarian as a second/foreign language, the non-inflected variant should be presented to the (more advanced) learners, possibly using longer stories, for instance Hungarian folk tales, as input texts. Students can identify the string in the text, they can discuss its idiosyncrasy, discover its meaning, and they can note that it is used in storytelling. This should lead them to the realization that it is an idiomatic construction.

The final expression to be considered in this section is *Van mit a tejbe aprítani*, (18).

- (18) 18. *Van mit a tejbe aprítani.*  
 there-is what.Acc the milk.Illat chop.Inf  
 Meaning: She/He can put meat on the table.

This lexically fixed non-compositional MEC expresses the meaning of ‘having plenty of money’, and it is often used with a bit of irony. The construction is unique in the sense that it contains an idiom which can only occur in the MEC-template, which otherwise has regular syntax and is open to structural variability. In contrast to other MECs, this construction most

typically occurs with the positive present tense form of the substantive verb. Only 6% of the examples are in the negative form, and around 10% in the past or in the future form. Half of the examples contain the uninflected form of the infinitive. Word order shows variation, as well. The illative case marked element *a tejbe* 'into the milk' can stand before or after the infinitive *aprítani* 'to chop'. However, over 70% of the examples have it before the infinitive. Despite its formal variability, the expression *Van mit a tejbe aprítani* can be regarded as highly idiomatic due to the idiom it contains (this type is not featured in the scale of idiomaticity in Figure 2).

#### 4.2. Pilot study on the idiomaticity of MECs

We conducted a small-scale study to find out where native speakers would place various MECs along the scale of idiomaticity. Twenty native speakers of Hungarian participated to the study. Since we wanted to make sure that the participants are familiar with the concept of idiomaticity, but they would not be biased by theoretical assumptions, we recruited first-year students of linguistics at the University of Pécs.

We designed a ten-minute questionnaire in Google Forms with nine example-constructions that appeared in minimal context. Participants were asked to answer the following questions after each item:

How adequate is the use of the construction in the given context? (Mark your answer on the five-point Likert-scale.)

Can you modify the construction? If yes, how? (Give a short answer.)

How idiomatic is the construction? (Mark your answer on the five-point Likert-scale.)

Can you replace the construction with another one? If yes, with what? (Give a short answer.)

The findings of the pilot study confirmed our analyses: the participants' answers yielded a very similar ordering of the expressions for their idiomaticity as our scale outlined above (Figure 3).

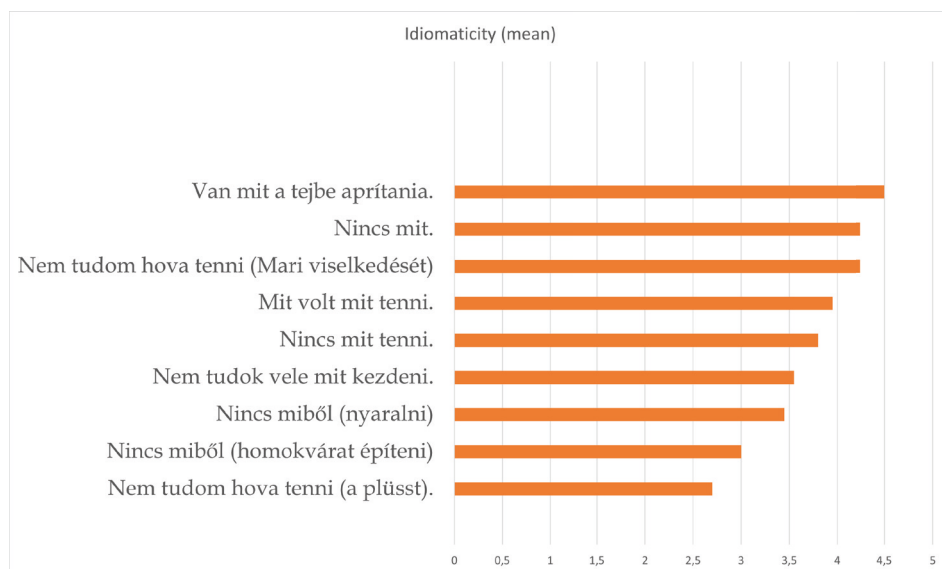


Figure 3 – Native speakers' idiomaticity judgements on a five-point Likert-scale (average)<sup>6</sup>.

The answers showed a high degree of agreement at the two ends of the scale (Figure 4). It seems that the participants of the pilot study had no difficulty identifying the most and the least idiomatic MECs. They found the expression *Van mit a tejbe aprítania* (there-is what.Acc the milk.Illat chop. Inf.Sg3) 'He/She has plenty of money'/'He/she can put meat on the table' to be the most idiomatic of the nine MECs. This construction contains an

<sup>6</sup> *Van mit a tejbe aprítania*; there-is what.Acc the milk.into chop.Inf.Sg3; 'She/He can put meat on the table'; *Nincs mit*; there-is-no what.Acc; 'Not at all'; *Nem tudom hova tenni (Mari viselkedését)*; no can.Sg1Def where-to put.Inf (Mari.NOM behaviour.Poss3.Acc); 'I can't make any sense of it/(Mary's actions)'; *Mit volt mit tenni*; what.Acc there-was what.Acc do.Inf; 'There was nothing for it'; *Nincs mit tenni*; there-is-no what.Acc do.Inf; 'There is no way around it'; *Nem tudok vele mit kezdeni*; no can.Sg1Indef with-it what.Acc begin.Inf; 'I can't do anything about it'; *Nincs miből (nyaralni)*; there-is-no what.Elat (be-on-holiday.Inf); '(We) can't afford it/(going on a holiday)'; *Nincs miből (homokvárat építeni)*; there-is-no what.Elat (sand-castle. Acc build.Inf); 'There's nothing (to build a sand-castle from)'; *Nem tudom hova tenni (a plüsst)*; no can.Sg1Def where-to put.Inf (the soft-toy.Acc); 'I have nowhere to put it/(the soft toy)'.

idiom which can only appear in the modal existential wh-construction. They also found the expression *Nincs mit* (there-is-no what.Acc) ‘Not at all/You’re welcome’ idiomatic. This construction is fully fixed and conventionalized as a polite response to thank you. Those constructions were found less idiomatic by the respondents that could be predicted or inferred based on the form and function of the abstract MEC with a canonical meaning: *Nincs hova tenni a játékot* (there-is-no where-to put.Inf the toy.Acc) ‘There is no space to put the toy anywhere’, and *Nincs miből várat építeni* (there-is-no what.Elat castle.Acc build. Inf) ‘There is nothing to build a castle from’. However, the participants were uncertain about those expressions that are somewhere halfway between the two extremes.

In conclusion, the participants of the study tended to find a construction idiomatic if it is formally and lexically fixed, and if it has a distinct function or meaning that cannot be predicted based on the canonical MEC.

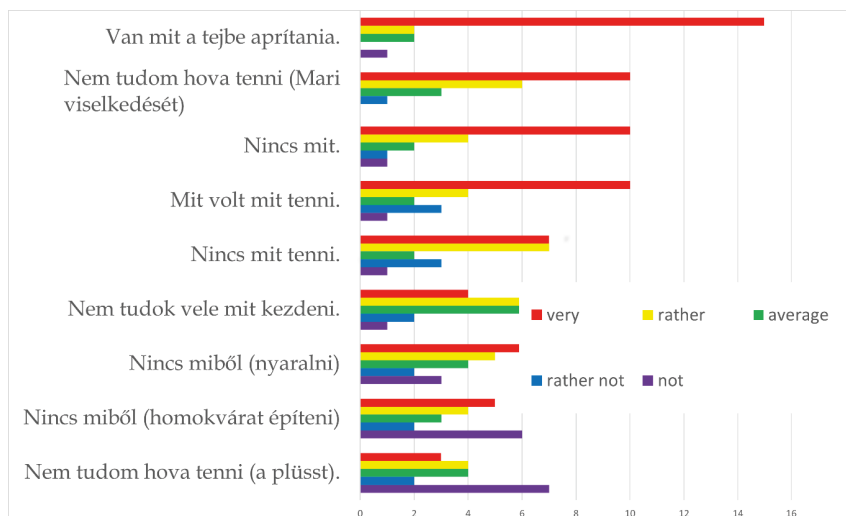


Figure 4 – Native speakers’ idiomaticity judgements (standard deviation)<sup>7</sup>.

<sup>7</sup> See Footnote 2 for the English translations.

## 5. Corpus-assisted teaching of the MEC

Corpora can help us to identify the most frequent examples of lexically specific grammatical constructions, and they can be effectively used to discover a construction's productivity and variability. They can also be helpful in finding appropriate contexts to present the different expressions.

Over thirty years ago Tim Johns (1991) introduced data-driven learning to the classroom. He emphasized that «the task of the language teacher is to provide a context in which the learner can develop strategies for discovery» (ivi, 1). He claims that the use of corpora, including concordances, can stimulate enquiry in language learning since it allows students to formulate and prove their own hypotheses about patterns of the target language. Keeping in mind that form and function are inseparable in language, it is useful to draw the students' attention both to the formal specificities of the MEC and to the fact that the abstract grammatical construction itself carries a meaning.

We have seen above that MECs are complex linguistic signs that combine a particular structural pattern with a particular meaning. Therefore, learners of Hungarian as a second/foreign language should learn them as pairings of form and meaning, where patterns with high relative frequency are associated with particular lexical expressions.

At first, students must get acquainted with the canonical MEC: they should be made aware of the relatively fixed schematic template [existential predicate, *wh*-element, (inflected) infinitive or conjugated subjunctive], and of the default meaning associated with the pattern [the availability of a proposition being possible or not (for someone)]. It is worth starting out with the negated form of the substantive verb since negation is more frequent in canonical MECs than positive statements. To this end, we can use concordance and frequency lists, and we can guide the students in the analysis of the data.

As a first step, students could be presented with the morphologically easier (inflected) infinitive-version. (N.B. MECs with the inflected infinitive can only be taught after the systematic elaboration of the inflectional variations of the Hungarian infinitive.)

We can start with prototypical sequences that contain the negative existential verb *nincs* 'there-is-no' and a *wh*-element: we can do a simple query for <nincs hol> 'there is nowhere'/'someone has nowhere', for instance, and we can also extract the most frequent collocations from the query. This way, we can demonstrate the semantic prosody of the [NINCS HOL + INF] MEC



through its most natural, native-like, and common instances. We may also want to filter the list so that we would have a reasonable number of examples to work with in class, including language that our students can be expected to understand at the given level. We may also want to highlight the target constructions.

# Corpus: MNSZ2

# Hits: 537

# Query word,[word="(?)nincs"|lemma="(?)nincs"]|word="(?)hol"|lemma="(?)hol"] 537

doc#114 - **Ottónak < nincs hol > aludnia** - mondta ki a lényeket. -

doc#327 - Nincs lakása? - < **Nincs hol > aludnom...** - Sirdogál. - Én ebbe belepusztulok

doc#534 parkolóhelyet zsúfolásig elleptek az autók, < **nincs hol > megállni.**

doc#979 Ragika, meghívlak. Sőt ha < **nincs hol > aludnod,** nálunk megleheted.

doc#984 de ettől még a fiataloknak < **nincs hol > lakniuk!** Vajk 02/20/99 13:06:13

doc#1081 nincs megtakarított pénzük, < **nincs hol > lakniuk,** nincs mit csinálniuk. -

doc#1086 Udvariasan elnézést kért, elmondta, < **nincs hol > aludnia,** de rendet fog rakni

doc#1089 iskolák, munkahelyek, ahol befogadnák őket, < **nincs hol > lakniuk,** ha a

doc#1097 ahol van munkalehetőség, viszont < **nincs hol > lakniuk.** Így aztán üres

doc#1107 lakásokat keresnek

doc#2606 a 74 szakembernek pillanatnyilag < **nincs hol > dolgoznia,** mert...

doc#2627 elmennek, mert még dolgozni kell, de < **nincs hol > dolgozni.**

**család**nak nincsen más problémája, mint hogy < **nincs hol > laknia.**

# Collocations

# Corpus:corpname

# Query: word,[word="(?)nincs"|lemma="(?)nincs"]|word="(?)hol"|lemma="(?)hol"] 537

	Freq	T-score	MI	logDice
laknia	51	7.141	18.528	10.938
lakniuk	42	6.481	19.063	10.909
aludnia	22	4.690	17.279	9.711
hajóztatni	14	3.742	20.874	9.694
aludnotok	14	3.742	20.261	9.670
laknom	6	2.449	16.958	8.247
laknunk	4	2.000	17.454	7.798
aludniuk	4	2.000	16.143	7.621
megpihennie	3	1.732	19.570	7.492
parkolniuk	3	1.732	18.570	7.469
aludnom	5	2.236	14.833	7.450
aludnod	3	1.732	16.631	7.342
enniük	3	1.732	14.431	6.847
dolgozniuk	8	2.828	13.261	6.748
tárolni	5	2.235	11.481	5.148
elhelyezkedni	4	1.999	11.324	4.971
lakni	4	1.999	11.158	4.825
dolgoznia	3	1.731	10.980	4.617
elhelyezni	4	1.998	10.203	3.954
aludni	5	2.232	9.101	2.911
dolgozni	10	3.154	8.517	2.351
játszani	3	1.722	7.470	1.296

Figure 5 – Concordance list of <nincs hol> (there-is-no where)<sup>8</sup>.

<sup>8</sup> The English translations of the items in bold: doc#114: Ottó.Dat there-is-no where sleep. Inf.Sg3; 'Ottó has nowhere to sleep'. doc#327: there-is-no where sleep. Inf.Sg1; 'I have nowhere

As a second step, we can guide our students to analyze the tokens to infer the abstract general template, and to identify the most prototypical verbs that can occur in the [NINCS HOL INF] MEC. We can ask questions like the following ones: What comes after the <nincs hol> string? In what form does the infinitive stand? Which verbs (infinitives) feature most frequently after the <nincs hol> string? Which person-markings are the most frequent ones? What does the [NINCS HOL INF] construction mean? In conclusion, we can establish a template like the one in Figure 6. At this point, students can also be encouraged to look for equivalent constructions in their own languages.

NINCS HOL + INF			
AVAILABILITY NOT POSSIBLE	PROPOSITION		PERSON
main existential predicate (negative) NINCS	interrogative indefinite pronoun HOL	subordinate verb in the (CONJUGATED) INFINITIVE FORM	dative case marking -NAK/NEK
NINCS	HOL	LAKNI	–
		ALUDNI	
		DOLGOZNI	
		aludnia	Ottónak/neki
		aludnod	neked
		aludnom	nekem
		laknia	neki
		laknod	neked
		laknom	nekem

Figure 6 – Template for the [NINCS HOL INF] MEC (there-is-no where Inf)<sup>9</sup>.

to sleep'. doc#534: there-is-no where stop.Inf; 'There's nowhere to stop'. doc#979: there-is-no where sleep.Inf.Sg2; 'You have nowhere to sleep'. doc#884: the young.Pl.Dat there-is-no where live.Inf.Pl3; 'The young have nowhere to live'. doc#1081: there-is-no where live.Inf.Pl3; 'They have nowhere to live'. doc#1086: there-is-no where sleep.Inf.Sg3; 'She/He has nowhere to sleep'. doc#1089: there-is-no where live.Inf.Pl3; 'They have nowhere to live'. doc#1097: there-is-no where live.Inf.Pl3; 'They have nowhere to live'. doc#1107: there-is-no where work.Inf.Sg3; 'She/He has nowhere to work'. doc#2606: there-is-no where work.Inf; 'There is nowhere to work'. doc#2627: family.Dat there-is-no where live.Inf.Sg3; 'The family has nowhere to live'.

The English translations of the collocating infinitives: *laknia* (live.Inf.Sg3); *lakniuk* (live.Inf.Pl3); *aludnia* (sleep.Inf.Sg3); *hajózzatni* (take-a-boat-ride.Inf); *aludnotok* (sleep.Inf.Pl2); *laknom* (live.Inf.Sg1), *laknunk* (live.Inf.Pl1); *aludniuk* (sleep.Inf.Pl3); *megpihennie* (vp-rest.Inf.Sg3); *parkolniuk* (park.Inf.Pl3); *aludnom* (sleep.Inf.Sg1); *aludnod* (sleep.Inf.Sg2); *enniük* (eat.Inf.Pl3); *dolgozniuk* (work.Inf.Pl3); *tárolni* (store.Inf); *elhelyezkedni* (vp-be-situated.Inf); *lakni* (live.Inf); *dolgozni* (work.Inf.Sg3); *elhelyezni* (vp-place.Inf); *aludni* (sleep.Inf); *dolgozni* (work.Inf); *játszani* (play.Inf).

<sup>9</sup> NINCS: there-is-no; HOL: 'where'; LAKNI: live.Inf; ALUDNI: sleep.Inf; DOLGOZNI: work.Inf; *aludnia Ottónak/neki*: sleep.Inf.Sg3 Ottó.Dat/for-him/her; *aludnod neked*: sleep.Inf.Sg2

In addition, we can use concordance lists like the one in Figure 7. We can ask our students to find the three major elements (the verb, the wh-word, and the infinitive) of the construction, and to observe the formal variability of the infinitive form in the canonical MEC.

The examples in Figure 7 are selected not only with regard to form – i.e., the present tense negative form of the main existential predicate (substantive verb) and the infinitival form of the subordinate verb –, but also with respect to the environment of the construction. That is because we can only expect the students to understand the basic meaning of the construction if they are provided with sufficient context to infer its meaning. The small contexts which precede the MEC-forms contain elements that can help learners to work out the general meaning of the grammatical construction. For instance, in the first example, *hajléktalan* ‘homeless’ is the key word on the basis of which students can infer the meaning of *nincs hol aludnia* ‘has nowhere to sleep’. In the second example, the noun *vétkem* ‘my sin’ is the prompt word.

We can also do further queries for specific wh-elements (e.g., <nincs mit>, <nincs kivel>, <nincs hova>, <nincs miből> etc.), and for other forms of the main existential predicate (e.g., <van hol>, <nem volt hova> etc.), and we may come up with further templates as a result. Using these, we can infer the most general abstract schema as presented in Figure 1.

During these analyses, learners can discover and observe various instances of the MEC-construction together with their most frequent lexical-grammatical realizations. In this way, students meet authentic language data that they can use to learn ‘real’ language, while their attention is focused on the patterning of language. Here, they can realize that the MEC is a relatively strict grammatical construction with a given meaning, where typical <main-existential-predicate + wh-element> strings are associated with particular subordinate verbs, but it is also a grammatical construction which is open to certain modifications (e.g., to person-marking by inflecting the infinitive). In fact, as a next step, learners can be asked to make similar sentences by analogy – e.g., *Nincs hol úszni*; *Van hol vásárolni*; *Nincs mit olvasnom*; *Nem volt kivel beszélgetnie*, etc.).

---

for-you; *aludnom nekem*: sleep.Inf.Sg1 for-me; *laknia neki*: live.Inf.Sg3 for-him/her; *laknod neked*: live.Inf.Sg2 for-you; *laknom nekem*: live.Inf.Sg1 for-me.

#1225446795	Egy hajléktalant nem azért kellene vagy lehetne büntetni, mert	<b>nincs</b>	<b>hol</b>	<b>aludnia</b>	vagy melegegdnie			
#948777887	Kevés a vétkem,	<b>nincs</b>	<b>mit</b>	<b>bevallanom</b>				
#1000231272	Jobban élvezi a semmittevést akkor, ha előre tudja, hogy	<b>nincs</b>	<b>mit</b>	<b>csinálnia</b>				
#1088101280	Sztárok után most egy - egyelőre - ismeretlen nővel jár Jake Gyllenhaal, és	<b>nincs</b>	<b>mit</b>	<b>csodálkozni</b>	a választásán			
#176010555	Most eltemettem a feleségemet is. Nincs már családom, nincs már házam, lassan	<b>nincs</b>	<b>mit</b>	<b>ennem</b>	sem.			
#413359430	Nincs kenyér,	<b>nincs</b>	<b>mit</b>	<b>enni,</b>	nincs	mi között	válogatni	
#599831031	Ha bejön valami kis pénz, van csoki, chips, rágó, ha pedig elfogyott, akkor	<b>nincs</b>	<b>mit</b>	<b>enni,</b>				
#1232794418	Élelem nincs, a mentett állatokat	<b>nincs</b>	<b>miből</b>	<b>etetni!</b>				
#496112387	Ezeknek az embereknek az a legfőbb "bűnük", hogy szegények, és	<b>nincs</b>	<b>hol</b>	<b>lakniuk.</b>				
#955170205	Hova menekülhetnénk? -Sehova,	<b>nincs</b>	<b>hova</b>	<b>menni</b>				
#1220530146	Bár nem lehet becsomagolni, mégse jó úgy e-könyv-olvasót ajándékozni, hogy	<b>nincs</b>	<b>mit</b>	<b>olvasni</b>	rajta.			

Figure 7 – Concordance list for the [NINCS wh-pronoun INF] canonical MEC<sup>10</sup>.

Simple language games can help learners commit the construction to memory. In the word rotator illustrated in Figure 8, students can change

<sup>10</sup> The English translations of the items in bold: *nincs hol aludnia*: there-is-no where sleep.Inf.Sg3; 'she/he has nowhere to sleep'; *nincs mit bevallanom*: there-is-no what.Acc vp-admit.Inf.Sg1; 'I have nothing to admit'; *nincs mit csinálnia*: there-is-no what.Acc do.Inf.Sg3; 'she/he has nothing to do'; *nincs mit csodálkozni*: there-is-no what.Acc be-surprised.Inf; 'there is nothing to be surprised at'; *nincs mit ennem*: there-is-no what.Acc eat.Inf.Sg1; 'I have nothing to eat'; *nincs mit enni*: there-is-no what.Acc eat.Inf; 'there is nothing to eat'; *nincs mit enni*: there-is-no what.Acc eat.Inf; 'there is nothing to eat'; *nincs miből etetni*: there-is-no what.Elat feed.Inf; 'there is nothing to feed it/him/her from'; *nincs hol lakniuk*: there-is-no where live.Inf.Pl3; 'they have nowhere to live'; *nincs hova menni*: there-is no where-to go.Inf; 'there is nowhere to go'; *nincs mit olvasni*: there-is-no what.Acc read.Inf; 'there is nothing to read'.

each of the three elements of the bare-infinitive type MEC in a rotating calendar to build MECs with the canonical meaning. The students' task is to come up with as many MECs as possible, and, possibly, to think of a simple speech situation in which the given MEC can appear. Here, particular attention must be paid to the *wh*-element since it must be compatible with the infinitive: only those *wh*-elements can be selected for a subordinate verb that can occur in a verb-argument structure construction with it (it is best to use verbs the argument structures of which the students are familiar with, and it is important to recall this information before working with the word rotator).

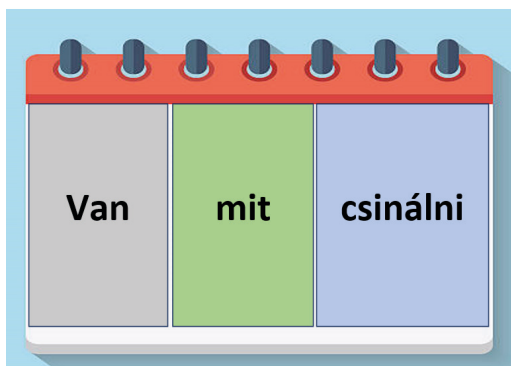
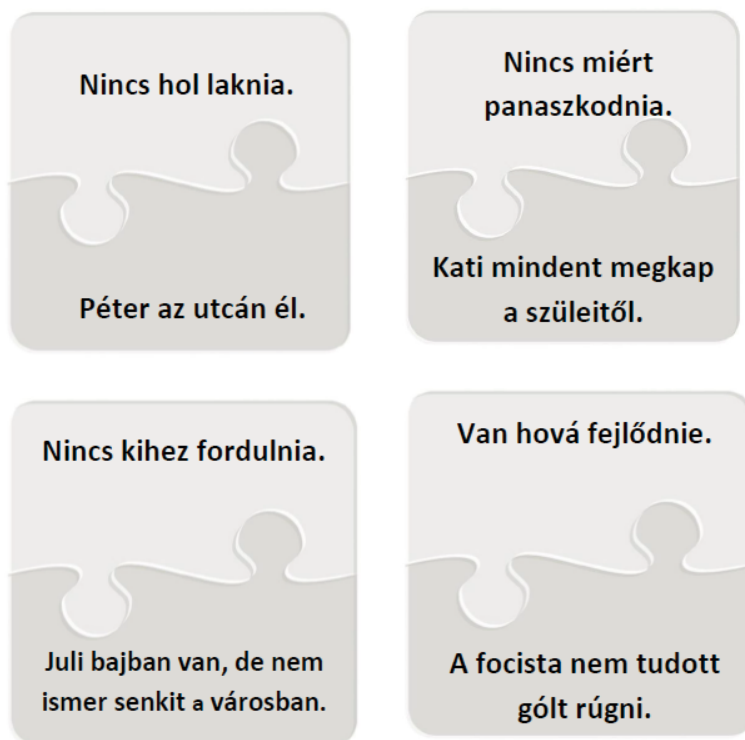


Figure 8 – The MEC Word Rotator<sup>11</sup>.

We can help our students to build appropriate narrow contexts around specific MECs with puzzle pieces (Figure 9). Each student gets one puzzle piece with either a MEC or with an explanatory sentence in it that provides context to a given MEC. Students must find the corresponding piece to their own puzzle piece.

<sup>11</sup> VAN: there-is; MIT: what.Acc; CSINÁLNI: do.Inf. 'There are things to do'.

Figure 9 – MEC puzzle pieces<sup>12</sup>.

Once the learners are familiar with the MEC construction, we can create broader contexts in which they can use the construction in meaningful interactions. For instance, we can use simulation tasks to elicit the use of the construction in conversation (Figure 10). Students work in pairs. They

<sup>12</sup> Puzzle piece 1: *Nincs hol laknia* 'She/He has nowhere to live'. *Péter az utcán él* 'Péter lives on the street'. Puzzle piece 2: *Nincs miért panaszkodnia* 'She/He has no reason to complain'. *Kati mindent megkap a szüleitől* 'Kati gets everything from her parents'. Puzzle piece 3: *Nincs kihez fordulnia* 'She/He has no one to turn to'. *Juli bajban van, de nem ismer senkit a városban* 'Juli is in trouble, but she doesn't know anyone in the city'. Puzzle piece 4: *Van hová fejlődnie* 'She/He has plenty to improve'. *A focista nem tudott gólt rúgni* 'The footballer couldn't score a goal'.

are given leaflets, booklets, web sites etc. about different towns and cities, and they must choose a holiday destination for themselves, or for certain people, families etc. First, they list different facilities and talk about the available opportunities (or the lack thereof). Based on the advantages and disadvantages, they finally decide on the best holiday destination.

## VESZPRÉM

<https://www.veszpreminfo.hu/>

<https://blog.hovamenjek.hu/a-honap-varosa-veszprem>

<https://pixabay.com/de/photos/veszpr%C3%A9m-veszpr%C3%A9m-stadt-burg-2912431/>



FACILITIES	AVAILABLE OPPORTUNITIES
<i>Vannak múzeumok és galériák.</i>	<i>Van hova menni, ha érdekel a kultúra és a történelem.</i>
<i>Van egy nagy állatkert.</i>	<i>Van hol eltölteni az időt a gyerekekkel.</i>
<i>Kisváros.</i>	<i>Este nincs hol szórakozni.</i>
<i>Nincs tenger.</i>	<i>Nincs hol úszni.</i>
...	<i>Nincs hol friss tengeri halat enni.</i>
...	...

Figure 10 – Simulation task<sup>13</sup>.

Another technique that can be used to generate interaction around the MEC-construction is problem-solving activities (Figure 11). Students should list problems that are likely to present themselves in given situations (e.g., complaints of a lonely woman living in a small village), and they should offer advice on how to solve the problems and overcome the difficulties.

<sup>13</sup> Facilities: *Vannak múzeumok és galériák* 'There are museums and galleries'; *Van egy nagy állatkert*: 'There is a big zoo'; *Kisváros* 'It's a small town'; *Nincs tenger* 'There isn't a sea'. Available opportunities: *Van hova menni, ha érdekel a kultúra és a történelem* 'There are places to go if you're interested in culture and history'; *Van hol eltölteni az időt a gyerekekkel* 'There are places to spend your time with the kids'; *Este nincs hol szórakozni* 'There is nowhere to go out in the evening'; *Nincs hol úszni* 'There is nowhere to swim'; *Nincs hol friss tengeri halat enni* 'There is nowhere to eat fresh seafish'.

PROBLEMS/COMPLAINTS	SOLUTIONS/ADVICE
<i>Nincs hol dolgoznia.</i>	?
<i>Nincs kivel barátkoznia.</i>	?
<i>Nincs hova mennie szórakozni.</i>	?
...	...

Figure 11 – Problem-solving task<sup>14</sup>.

We can also use tasks where learners use the MEC to share personal information about themselves. Students work in pairs. Each pair is given a number of verbs such as *örül* ‘is happy’, *panaszodik* ‘complain’, *aggódik* ‘worry’, *fél* ‘fear’ etc. They build MECs with these verbs to ask questions of each other: e.g., *Van minek örülnöd?* ‘Do you have something to be happy about?’ The questions should be used to generate conversation: students can elaborate on their answers and discuss further details.

In addition to the schematic grammatical MEC-construction, learners must also be presented with grammatically and lexically fully fixed idiomatic instances of the MEC. At the beginning of their studies, students learn the phrase *Nincs mit* (there-is-no what.Acc) ‘not at all/you’re welcome’ as a conventional response to *Köszönöm* (thank.Sg1Def) ‘thank you’ – most probably holistically, without any recourse to its inner structure. Now students might find out that this phrase is actually a MEC, but it is very restricted both as for form and for usage. It is a situation bound utterance (Kecskes 2010) which is used as a formulaic response to thank you. In this function, the string is totally fixed, it cannot be modified or extended in any way: one cannot say, for example, *Nincs miket* (there-is-no what.Pl.Acc) or *Nincs mit (meg)köszönni* (there-is-no what.Acc (vp)thank.Inf).

Upper level learners may be presented with corpus data for other idiomatic instances of MEC such as *Nincs mit tenni* (there-is what.Acc do.Inf; ‘There is nothing you can do about it’/‘There is no way around it’) or *Ezzel nem tudok mit kezdeni* (with-this no I-can what.Acc start.Inf; ‘I can’t do anything with/about it’/‘I can’t handle this’). Here, we will look at the highly

<sup>14</sup> *Nincs hol dolgoznia* ‘She/He has nowhere to work’. *Nincs kivel barátkoznia.* ‘She/He has noone to make friends with’. *Nincs hova mennie szórakozni* ‘She/He has nowhere to go out’.



idiomatic expression *Van mit a tejbe aprítani* (there-is what.Acc the milk. into chop.Inf) (Figure 12).

```
# Corpus: MNSZ2
# Hits: 48
#
word,[word="(?)van"|lemma="(?)van"] [word="(?)mit"|lemma="(?)mit"] [word="(?)a"|lemm
a="(?)a"] [word="(?)tejbe"|lemma="(?)tejbe"] [word="(?)apritani"|lemma="(?)apritani"] 48

doc#699      Drága porcelánok! Képek! A szép bútorok... < Van mit a tejbe aprítani >!
doc#1432     minden pénzt neki adtam. Boldog vagyok, hogy < van mit a tejbe aprítani >
doc#2251     hiszen ő nem azért rohant ide, mert nem < volt mit a tejbe aprítani >. Nem élt
              ő rosszul Magyarországon
doc#2288     egyéni vállalkozó lett, ment a buli, < volt mit a tejbe aprítani >
doc#2750     ne sírjunk, mert nekünk még < van mit a tejbe aprítani >, sok szabadság,
              utazgatás, hobbik stb...
doc#2868     <Van mit a tejbe aprítani >! Két szomszédos villát vásárolt az énekesnő
doc#2877     egy cégnél voltam vezető beosztásban, tehát < volt mit a tejbe aprítani >.
```

Figure 12 – Corpus examples for the [VAN MIT A TEJBE APRÍTANI] MEC (there-is what.Acc the milk-into chop.Inf; ‘can put meat on the table’)<sup>15</sup>.

With sufficient guidance from the teacher, students may guess the meaning of the expression from the provided contexts. They can realize that it has something to do with money and being well-off: it means ‘have plenty of money’, ‘be loaded’, or, in an idiomatic translation, ‘can put meat on the table’. By looking at a concordance list for *tejbe aprítani* (Figure 13), students can also conclude that this MEC tolerates minimal modification: only the main existential predicate (*van* ‘is’) is likely to be changed (to *volt* ‘was’, *lesz* ‘will be’, *lett* ‘got/became’, *legyen* ‘let there be’, and to negative forms of these), and person-marking happens by adding the dative form of the personal pronoun: the infinitive is typically not inflected. This dative pronoun may occasionally appear between the main existential predicate and the wh-element.

<sup>15</sup> The English translations of the items in bold: *van mit a tejbe aprítani* there-is what.Acc the milk-into chop.Inf; ‘there is meat to put on the table’; (*nem*) *volt mit a tejbe aprítani* (no) there-was what.Acc the milk-into chop.Inf; ‘there was (no) meat to put on the table’.

# Corpus: MNSZ2

# Hits: 64

#

word,[word="(?)a"|lemma="(?)a"]][word="(?)tejbe"|lemma="(?)tejbe"]][word="(?)aprítani"|lemma="(?)aprítani"] 64

doc#369 Pedig **van neki mit** < **a tejbe aprítani** >. Nincs rászorulva, elhiheted... -  
 doc#699 mindent megteszek, hogy **legyen mit** < **a tejbe aprítani** >!  
 doc#2623 visszaadnám, mert amúgy is **van neki mit** < **a tejbe aprítani** > amit tőlünk  
 nyugdíjasoktól elvettek hosszú  
 doc#2856 akkor a hatalom éhes embernek **nem lesz mit** < **a tejbe aprítani** >.

Figure 13 – Variants for the [VAN MIT A TEJBE APRÍTANI] MEC  
 (there-is what.Acc the milk-into chop.Inf; ‘can put meat on the table’)<sup>16</sup>.

In conclusion, the class can summarize their inferred information for this particular MEC as presented in Figure 14.

VAN MIT A TEJBE APRÍTANI (~ van pénze, jól megy neki)			
PERSON	AVAILABILITY (NOT) POSSIBLE	PERSON	“MIT A TEJBE APRÍTANI”
dative case personal pronoun / noun	main existential predicate (negative)	dative case personal pronoun/noun	<b>mit a tejbe aprítani</b>
<b>neki</b> N.-nak/nek  nekem neked ...	<b>VAN</b> <b>(NEM) VOLT</b> <b>(NEM) LESZ</b> <b>LETT</b> <b>LEGYEN</b>	<b>neki</b> N.-nak/nek  nekem neked ...	

Figure 14 – Template for the [VAN MIT A TEJBE APRÍTANI] MEC  
 (there-is what.Acc the milk-into chop.Inf; ‘can put meat on the table’)<sup>17</sup>.

<sup>16</sup> The English translations of the items in bold: *van neki mit a tejbe aprítani* there-is for-him/her what.Acc the milk-into chop.Inf; ‘she/he has can put meat on the table’; *legyen mit a tejbe aprítani* there-is.Imp.Sg3 what.Acc the milk-into chop.Inf; ‘(so that) there is meat on the table’; *nem lesz mit a tejbe aprítani* no there-will-be what.Acc the milk-into chop.Inf; ‘there won’t be any meat on the table’.

<sup>17</sup> N.-nak/nek: N.Dat; *neki* ‘for him/her’; *nekem* ‘for me’; *neked* ‘for you’; **VAN** ‘there-is’; **(NEM) VOLT** ‘(no) there-was’; **(NEM) LESZ** ‘(no) there-will-be’; **LETT** ‘there got to be’ (change of state); **LEGYEN** ‘there should be’ (imperative); *mit a tejbe aprítani* what.Acc the milk-into chop.Inf.

Following the corpus-assisted analysis, students can be given tasks that help them memorize the given fixed idiomatic expression in meaningful contexts – for instance, in the form of discourse-completion tasks, the creation of mini-situations, and role-plays.

By using corpus data in the teaching of MECs, both teachers and learners can become attentive to the pattern of the construction. By analyzing real language data, they can discover language for themselves, and they can explore the idiomaticity and (non-)variability of linguistic patterns.

## 6. Conclusions

In this paper we argued why and how the use of corpora can be beneficial in language teaching. Corpora can help us to identify frequent patterns of the target language which could otherwise remain hidden, and they can also be used to discover the formal and functional aspects of abstract grammatical constructions as well as those of the frequent lexically particular instances of the schematic construction. We chose the Hungarian modal existential *wh*-construction (MEC) to demonstrate our points. The reason why we chose this particular construction is partly because although the MEC is a relatively frequent and highly productive infinitival construction in Hungarian today, it is rarely taught explicitly in Hungarian as a second/foreign language (HFL) coursebooks. The MEC is a schematic grammatical construction: a learned pairing of the [matrix predicate + *wh*-word + infinitival or subjunctive subordinate verb] form and the canonical meaning expressing that the possibility of a proposition is available or not.

After describing the construction, we presented relevant corpus data from Old to Modern Hungarian. We showed how the frequency and productivity of the construction increased over time, and how, at the same time, the pattern became associated with particular lexical expressions, which now display a higher degree of invariability than the schematic MEC, and which are associated with distinct functions and meanings. We proposed to place various types of the MEC along a scale of idiomaticity, which proposition we also tested for in the form of a small-scale pilot study. We argued that in the teaching of HFL, attention must be paid both to the abstract, more schematic, and highly productive MEC-template and to its lexically fully

fixed, more idiomatic instances. We claim that if learners are ‘only’ made familiar with the lexical-grammatical buildup of the MEC and its default canonical meaning – that is, if we only treat the MEC as ‘grammar’ –, they will not have sufficient knowledge to use the construction in a native-like fashion. Learners must be made aware that in today’s Hungarian the abstract grammatical construction is most frequently used in the following form: [the negative form of the substantive verb + wh-pronoun + infinitive], and they should use this template to form expressions by analogy. They must, however, also be made aware of the lexical expressions that are most frequently co-selected in the prototypical MEC. Finally, they must be presented with lexically fixed highly idiomatic instances of the construction with little or no variability, where the expression as a whole takes on a distinct meaning that cannot be inferred compositionally.

Finally, we put forth ideas and ways for the teaching of various MEC-constructions in the language classroom with the help of corpora and follow-up activities. Providing that they get sufficient training, assistance, and guidance, we expect that the activities we suggested can help teachers and even learners to identify the most frequent examples of lexically particular grammatical constructions, and to discover a construction’s productivity and variability. Corpora can also be helpful for teachers to find appropriate contexts to present the different expressions in class. Further research is required to evaluate the applicability and effectiveness of the proposed activities in actual classroom settings both from the teachers’ and from the learners’ end.

#### *Works cited*

- Baumann, Tímea *et al.* «Bemutatózik a Korpusznyelvészeti és Szakmódszertani Munkacsoport» [Introduction of the Work Group for Corpus Linguistics and Didactics]. In *Hungarológiai Évkönyv* vol. 21, nn. 1-2 (2020): 32-41. URL: <<https://epa.oszk.hu/02200/02287/00021/pdf/>> (*open access*).
- Dékány, Éva. 2014. «A nem véges alárendelés (az igenevek) története» [History of non-finite subordination (participles)]. In *Magyar generatív történeti mondattan* [Hungarian generative historical syntax], edited by Katalin É. Kiss, 179-239. Budapest: Akadémiai Kiadó. DOI: <<https://doi.org/10.1556/9789630597685>> (*open access*).

- Diessel, Holger. «Usage-based linguistics». In *Oxford Research Encyclopedia of Linguistics*, edited by Mark Aronoff. New York: Oxford University Press, 2017. DOI: <<https://doi.org/10.1093/acrefore/9780199384655.013.363>> (open access).
- Dömötör, Adrienne *et al.* «Kiútkeresés a morfológiai labirintusból – korpuszépítés ó- és középmagyar kori magánéleti szövegekből» [A way out of the labyrinth of morphology – corpus-building from Old and Middle Hungarian personal texts]. *Nyelvtudományi Közlemények* vol. 113 (2017): 85-110. DOI: <<https://doi.org/10.15776/NYK.2017.113.3>> (open access).
- Endresen, Anna, Laura A. Janda. «Taking Construction Grammar One Step Further: Families, Clusters, and Networks of Evaluative Constructions in Russian». *Frontiers in Psychology* vol. 11 (2020). DOI: <<https://doi.org/10.3389/fpsyg.2020.574353>> (open access).
- Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago, IL: University of Chicago Press.
- . 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press. DOI: <<https://doi.org/10.1093/acprof:oso/9780199268511.001.0001>>.
- . 2019. *Explain me This: Creativity, Competition, and the Partial Productivity of Constructions*. Princeton: Princeton University Press. DOI: <<https://doi.org/10.2307/j.ctvc772nn>>.
- Johns, Tim. «Should you be persuaded. Two samples of data-driven learning materials». *ELR Journal* vol. 4 (1991): 1-16. URL: <[https://lexically.net/wordsmith/corpus\\_linguistics\\_links/Tim%20Johns%20and%20DDL.pdf](https://lexically.net/wordsmith/corpus_linguistics_links/Tim%20Johns%20and%20DDL.pdf)> (open access).
- Keck, Casey Mari. «Corpus linguistics and language teaching research: bridging the gap». *Language Teaching Research* vol. 8, n. 1 (2004): 83-109. (open access).
- Kecskes, Istvan. «Situation-bound utterances as pragmatic acts». *Journal of Pragmatics* vol. 42, n. 11 (2010): 2889-2897. DOI: <<https://doi.org/10.1016/j.pragma.2010.06.008>>.
- Lipták, Anikó. 2003. *Hungarian modal existential wh-constructions*. Manuscript. Leiden: University of Leiden.
- McNally, Louise. «Existential Sentences Crosslinguistically: Variations in Form and Meaning». *Annual Review of Linguistics* vol. 2 (2016): 211-231. DOI: <<https://doi.org/10.1146/annurev-linguistics-011415-040837>> (open access).
- Michaelis, Laura A. «Meanings of Constructions». In *Oxford Research Encyclopedia of Linguistics*, edited by Mark Aronoff. New York: Oxford University Press, 2017. DOI: <<https://doi.org/10.1093/acrefore/9780199384655.013.309>> (12/2022).

- Mukherjee, Joybrato. 2016. «Corpus linguistics and language pedagogy: The state of the art – and beyond». In *Corpus Technology and Language Pedagogy: New resources, New Tools, New Methods*, edited by Sabine Braun, Kurt Kohn, and Joybrato Mukherjee, 5-24. Frankfurt am Main: Peter Lang. URL: <[https://www.uni-giessen.de/faculties/f05/engl/ling/staff/professors/jmukherjee/publications/pdfs/Mukherjee2006\\_pedagogy.pdf](https://www.uni-giessen.de/faculties/f05/engl/ling/staff/professors/jmukherjee/publications/pdfs/Mukherjee2006_pedagogy.pdf)> (11/2022 open access).
- Novak, Attila et al. 2017-. *A Történeti magánéleti korpusz (TMK)* [The Old and Middle Hungarian corpus of informal language use]. URL: <<https://tmk.nytud.hu/>>.
- . «Creation of an annotated corpus of Old and Middle Hungarian court records and private correspondence». *Language Resources and Evaluation* vol. 52 (2018): 1-28. DOI: <<https://doi.org/10.1007/s10579-017-9393-8>>.
- Oravecz, Csaba, Tamás Váradi, Bálint Sass. 2014. «The Hungarian Gigaword Corpus». In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC '14)*, edited by Nicoletta Calzolari et al., 1719-1723. Reykjavik-Paris: European Language Resources Association. URL: <[http://www.lrec-conf.org/proceedings/lrec2014/pdf/681\\_Paper.pdf](http://www.lrec-conf.org/proceedings/lrec2014/pdf/681_Paper.pdf)> (open access).
- Preradović, Nives Mikelić, Kristina Posavec, Danijela Unić. «Corpus-Supported Foreign Language Teaching of Less Commonly Taught Languages». *International Journal of Instruction* vol. 12, n 4. (2019): 335-352. URL: <[https://www.e-iji.net/dosyalar/iji\\_2019\\_4\\_22.pdf](https://www.e-iji.net/dosyalar/iji_2019_4_22.pdf)> (open access).
- Prohászka, Zsolt. «Van Marinak csak Pétertől mit ellopni – a modális egzisztenciális kérdőszói konstrukció a magyarban» [Van Marinak csak Pétertől mit ellopni. The Modal existential wh-construction in Hungarian]. Presented at the National Scientific Students' Associations Conference, Pécs: University of Pécs, 2019.
- . «A modális egzisztenciális kérdőszói konstrukció a magyar nyelvhasználatban» [The modal existential wh-construction in Hungarian language use]. In *Szemelvények* [Excerpts]. *Specimina Operum Iuvenum 7. Válogatás a PTE BTK 35. Országos Tudományos Diákköri Konferencián (2021) I–III. helyezést elért hallgatóinak pályaműveiből* [Selection from the papers presented at the 35<sup>th</sup> National Scientific Student Conference (2021) organized by the Faculty of Humanities of the University of Pécs that ranked first, second, third], edited by Fanni Balikó et al., vol. 6 (2022): 58-86. URL: <[https://btk.pte.hu/sites/btk.pte.hu/files/Kari\\_tud\\_diakkori\\_tanacs/specimina\\_operum7\\_beliv.pdf?fbclid=IwAR3HDFa3mKW7ZLc0rKarVZqHYN9hEvvbzCE09w-kXKpR0gSSAuNMcv22Hs](https://btk.pte.hu/sites/btk.pte.hu/files/Kari_tud_diakkori_tanacs/specimina_operum7_beliv.pdf?fbclid=IwAR3HDFa3mKW7ZLc0rKarVZqHYN9hEvvbzCE09w-kXKpR0gSSAuNMcv22Hs)> (open access).
- Prohászka, Zsolt et al. 2022. «Modal Existential Constructions in Hungarian». In *Language Use and Linguistic Structure: Proceedings of the Olomouc Linguistics Colloquium 2021*, edited by Markéta Janebová, Joseph Emonds, and Ludmila

- Veselovská, 202-223. Olomouc: Palacký University Olomouc. DOI: <<https://doi.org/10.5507/ff.22.24461489>>.
- Reppen, Randi. 2010. *Using Corpora in the Language Classroom*. New York: Cambridge University Press. DOI: <<https://doi.org/10.3998/mpub.371534>>.
- Sass, Bálint. «Keresés a korpuszban: A kibővített Magyar történeti szövegtár új keresőfelülete» [Searching in the corpus. The new search interface of the extended Hungarian Historical Corpus]. In *A nyelvtörténeti kutatások újabb eredményei* [Latest findings in language history research], edited by Tamás Forgács, Miklós Németh, and Balázs Sinkovics, vol. 9 (2017): 267-277. URL: <[http://real.mtak.hu/72475/1/sb\\_kereses\\_korpuszban.pdf](http://real.mtak.hu/72475/1/sb_kereses_korpuszban.pdf)> (open access).
- Sass, Bálint et al. 2011. *Magyar igei szerkezetek. A leggyakoribb vonzatok és szókapcsolatok szótára* [Hungarian verb phrases. Dictionary of the most frequent arguments and collocations]. Budapest: Tinta Könyvkiadó.
- Šimík, Radek. 2011. *Modal existential wh-constructions*. Utrecht: LOT. URL: <[https://www.lotpublications.nl/Documents/269\\_fulltext.pdf](https://www.lotpublications.nl/Documents/269_fulltext.pdf)> (open access).
- Simon, Eszter. 2019. «Magyar nyelvű történeti korpuszok» [Hungarian historical corpora]. In *Beszéd- és nyelvelemző szoftverek* [Speech and language analyzing softwares], edited by Sulyok Hedvig, Juhász Valéria, and Erdei Tamás, 31-43. Szeged: SZTE JGYPK. URL: <[http://acta.bibl.u-szeged.hu/72106/1/beszed\\_es\\_nyelvelemzo\\_szoftverek\\_tanulmanykotet\\_2019\\_031-043.pdf](http://acta.bibl.u-szeged.hu/72106/1/beszed_es_nyelvelemzo_szoftverek_tanulmanykotet_2019_031-043.pdf)> (open access).
- Simon, Eszter, Sass Bálint. 2012. «Nyelvtechnológia és kulturális örökség, avagy korpuszépítés ómagyar kódexekből» [Human language technology and cultural heritage: corpus-building from Old Hungarian codices]. In *Általános Nyelvészeti Tanulmányok* [Studies in General Linguistics], edited by Gábor Prósztéký and Tamás Váradi, vol. 24 (2012): 243-264. URL: <[http://omagyarkorpusz.nytud.hu/documents/papers/chapters/simonsass-chapter\\_-\\_nyelvtech\\_es\\_kult\\_orokseg.pdf](http://omagyarkorpusz.nytud.hu/documents/papers/chapters/simonsass-chapter_-_nyelvtech_es_kult_orokseg.pdf)> (open access).
- Sinclair, John. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- . 2008. «The phrase, the whole phrase and nothing but the phrase». In *Phraseology. An interdisciplinary perspective*, edited by Sylviane Granger and Fanny Meunier, 407-410. Amsterdam-Philadelphia: John Benjamins. DOI: <<https://doi.org/10.1075/z.139>>.
- Szabó, Veronika. «Korpusznyelvészeti adalékok az infinitívuszt tartalmazó szerkezetek tanításához» [How corpus-linguistics can contribute to the teaching of constructions including infinitives]. *THL2* nn. 1-2 (2020): 7-36. URL: <<https://epa.oszk.hu/01400/01467/00018/pdf/>> (12/2022 open access).

Szabó, Veronika, Prohászka Zsolt. «Hungarian modal existential wh-constructions in written language use – diachronic perspectives». Pécs: University of Pécs, 26. August 2021. URL: <[https://icsh15.netlify.app/icsh15\\_abstracts/MECs\\_0825\\_finalV.pdf](https://icsh15.netlify.app/icsh15_abstracts/MECs_0825_finalV.pdf)> (*open access*).