

Towards Describing the Extremely Multifunctional Hungarian Discourse Marker *hát*

Anna Szeteli

University of Pécs, Pécs, Hungary

anna.szeteli@gmail.com

Abstract: This paper presents the findings of an experimental research on five frequent functions of the Hungarian spontaneous speech specific discourse marker *hát*. As a first step, the five categories were established based on the initial results of a pilot study – a read-aloud experiment involving nine female native speakers of Hungarian – and were defined in a representationalist, dynamic pragmasemantic framework called ReALIS. The present paper is an extension of the former pilot study, now involving data from fifty-three participants, including both females and males. Since the usage of the discourse marker *hát* can only be adequately interpreted in whole utterances or, better, in discourse context, data collection was built around test situations. The analysis identifies important suprasegmental characteristics, namely, durational ones, of the different uses of the discourse marker in question.

Keywords: pragmasemantics; discourse markers; Discourse Representation Theory; PRAAT

1. Introduction

Section 1 gives a brief introduction to Hungarian discourse marker research in the context of the investigated marker *hát*. Section 2 is devoted to the pragmasemantic framework for the analysis of the discourse marker. Section 3 describes the research methodology and the test situations. Section 4 presents the results of the statistical analysis, on which the main conclusions of the paper are based in Section 5.

1.1 The Hungarian Discourse Marker *hát*

As part of the recent growth of interest in research on Hungarian discourse markers, the investigation of *hát* has been a central topic both in pragmatics (Schirm 2011, Alberti 2016) and in spontaneous-speech-corpus based suprasegmental phonology (Dér and Markó 2017) over the past few years.

The issue of the usage of *hát*, however, in Hungarian spontaneous speech, had long been a topic of discussion – although not from a linguistics perspective. The sentence-initial *hát*, which can indicate uncertainty or hesitation, has long been stigmatized in Hungarian language use, and its usage is still often advised against, especially by teachers in public education (Schirm 2011).

Similar to other discourse markers, *hát*, which shares its adverbial origin with *tehát* “so/thus”, went through a decategorization process over time (Schirm 2011). As Schirm (2017) describes the linguistic history of *hát*, before being used as a discourse marker, it was used as a conjunction with inferential-conclusional functions (similarly to *tehát*). It is this function of *hát* that can now be considered as its “core meaning” (Bell 1998).

The reason why it is extremely difficult to account for the meaning or function of *hát* is because it has several different interpretations depending on the mental state of the speaker who utters it. It can indicate strong-mindedness and hesitation, and it can carry complex suprasegmental tones such as teasing, for example.

The first research to investigate the functions of *hát* with relation to its prosodic realization was that of Dér and Markó in 2017. In their study, the authors set out to give account of the suprasegmental characteristics of four observed monofunctional uses of the marker, sidelining the fact of its otherwise extreme multifunctionality. The authors concluded that further investigation on a larger amount of data was necessary to differentiate between the various usages, however, at the same time, they rejected the possibility of such research – in lack of an adequate method for the automatic identification of monofunctional uses of *hát* in corpora (Dér and Markó 2017, 11).

As opposed to Dér and Markó’s (2017) corpus-based study of *hát* “well/so” in spontaneous speech, in our experimental research we were able to point out significant differences in the temporal characteristics of the different usages – “we” stands for the authors of the former pilot study (Szeteli and Alberti 2017). Nevertheless, there is no conflict between their results and ours. This is due to the fact that the functions they examined from the point of view of their prosodic characters were (essentially) distinct turn positions. They identified a turn-initial, a turn-medial and a turn-final position, and they also had a fourth, purely pragmatic category for *hát* indicating judgment. From these four categories, only the turn-medial *hát* proved to be significantly longer than the others.

In our research, we embedded the monofunctional uses in short dialogues, and we built very explicit contexts around them. We examined five different (but uniformly) turn-initial positions, out of which four were sentence-initial ones, and one was sentence-final. The difference between the various usages lies with what Alberti (2016) calls the “semaphore

effect” of certain discourse markers. According to this idea, discourse markers can signal how easy/difficult it will be for the listener to digest the message at hand. This semaphore effect can be regarded as a peripheral attribution of *hát* (cf. core/periphery model of Bell 1998) to the message.

In the test, various scenarios were described (see 3.2.1 and 3.2.2, e.g.), at the end of which, different alternatives were given as answers to the problem presented in the test situation. These answers were designed to represent the following alternatives: (i) a straightforward answer, (ii) an uncertain answer, (iii) an answer which the speaker considers uneasy or embarrassing, (iv) a teasing response, and (v) a confirmation with a sentence-final *hát*. The differences between the answers are described in the pragmasemantic framework *ReALIS* (Alberti 2011, Alberti and Nöthig 2015, Alberti et al. 2014, 2016) in the following section.

2. Dynamic Representationalist Framework *ReALIS*

The scientific antecedents of the applied framework *ReALIS* (Reciprocal and Lifelong Interpretation System) are, on the one hand, Montagovian formal semantics with its external truth-evaluation process (Dowty et al. 1981), and, on the other hand, the most classical achievements of pragmatics (Cooperative principle of Grice 1975, Works of Austin (1962/1975) and Searle (1979)), cognitive linguistics (Nuyts 2017) and The Theory of Mind (Wimmer and Perner 1983).

There have been other formal theories developing from the Montagovian tradition with a Gricean perspective on language use (e.g.: The Dynamic Pragmatics of Lauer (2013)). Nevertheless, these advocate an eliminative reduction of possible worlds and have preserved the antipsychologist and antirepresentationalist tradition (Groenendijk, Stokhof and Veltman 1996). *ReALIS*, however, belongs to the family of discourse representation theories (DRT: Kamp et al. 2011, Maier 2016) and as such, it can be regarded as the representationalist counterpart of these theories. While the antirepresentationalist Amsterdam school strives to eliminate the level of discourse representation, in *ReALIS* such a conflict is non-apparent. In the ontology of this framework, the level of discourse representation is embedded in the world model, realized through possible-wordlet-representations treated as mental states of the communicating interlocutors in the world.

In that way, *ReALIS* is a formal semantics-based approach, which is able to give a cognitive description of mental states, while applying the one and the same formal apparatus to both linguistic representations and to representations of the mind. Furthermore, the approach makes it possible to consider that the speaker and the listener take the role of the addresser and the addressee while articulating and perceiving a message (Oishi 2017). In this sense the linguistically encoded and conventionalized intensional profile of an utterance and the mental state of the interlocutors can be differentiated from each other and evaluated via a generalized Montagovian pattern-matching mechanism.

We use a finite structure called worldlet to represent some important properties of the mental states, as follows. The holder of the information is given $\langle i, u, o \rangle$, which can stand for

the speaker (I), the listener (you) or another (o) entity. Furthermore, referent *r* stands for an underspecified holder of the information or an attitude. The holder has an attitude toward the fact *e*, which can be a sensorial experience (E), or belief (B), or a desire (D), or an intention (I), or the interlocutor makes a hypothesis about him/her being authorized to achieve the fulfillment of the proposition *e*. Attitudes can also be embedded in each other recursively. The interlocutor's authority, for instance, can be another interlocutor's knowledge or desire. We are using an 11-degree scale $\langle -5, -4, -3, -2, -1, 0, +1, +2, +3, +4, +5 \rangle$ to capture the grades of epistemic modality (cf.: Leiss 2014, 53; Nuyts 2017). The original triplet $\langle +1, 0, -1 \rangle$ expressed three degrees as per true, false and unspecified (0). The third characters mark point of time relative to the utterance time 0 as follows: “+” refers to a later one, and “-” to a previous one. Every worldlet label contains this temporal “stamp” to capture various phenomena, such as the one referred to by the epistemic verb *think* in the past tense in the sentence *Mary thought that I know that fact e does not hold*.

2.1 Conventionalized Intensional Profiles

As striving for explanatory adequacy, it is worth hypothesizing from the perspective of language acquisition that children – on the basis of the meagre data set available to them – should understand the system of intensional profiles of sentence types via guided by certain operations. Only “generator values” should be set and keep in mind, which appear with a black background in the tables. Other values in the profiles are decided by means of general constraints requiring certain values to equal or to stand in complementary distribution. We attempt to base the current model of the profile system on the assumption that the *iB*-value (“what I know about the truth status of the given eventuality”) always serves as a generator. The *iBuB*-values (“the knowledge I attribute to you”) in the general target-oriented mentalization, for instance, are assumed to coincide with the *iB*-value or to be its opposite (α^* is defined as the set consisting of the scale values which are not in set α or $\{\alpha\}$).

Target-oriented mentalization	Declarative	Imperative	Interrogative
For <i>e</i> : <i>iB</i>	$\underline{iB} \in +5\alpha$	$\underline{iB} \in -5\gamma$	$\underline{iB} \in \gamma \bullet \beta \ (-5 \bullet \pm 5)$ $\rightarrow \underline{iB} \in "0"$
$\underline{iBuB} \in " \bullet +5$ $\underline{iBuB} \in \underline{iB}$ or $\underline{iB}^*$	$\underline{iBuB} \in \alpha^*$	$\underline{iBuB} = \gamma$	$\underline{iBuBiB} \in \beta \ (" +5 \bullet \gamma' \bullet \pm 5)$ $\rightarrow \underline{iBuBiB} \in " +5 \bullet (') +5 \bullet "0"$
For <i>e</i> : ... $\bullet W$, $r \in R \subseteq \{i, u, o\}$,	$W = \underline{uB} +$	Default: $e' = \underline{res}_e$	$W = \underline{iB} +$
$(\Sigma \underline{iB} : \underline{rD}) / SR \in " +5$	$\underline{iBrDuB} \in \pm 5\beta$	For <i>e</i> : <i>iBrD</i>	$\underline{iBrDiB} \in \beta$
$\underline{iBuA} \in " +5 \bullet " +5$	$\underline{iBuAuB} \in \beta$	For <i>e</i> : <i>iBuA</i>	$\underline{iBuAiB} \in \beta$ A factor: $\underline{iBuB} \in \beta$
For <i>e</i> : $\underline{iIuI} \in " +5 \bullet +5$	$\underline{iIuI} + \underline{uB} = \alpha$	For <i>e</i> : <i>iIuI</i> +	$\underline{iIuI} + \underline{iB} \in \beta$
$\underline{iAiIuI} \in " +5$	$\underline{iAiIuI} + \underline{uB} = \alpha$	For <i>e</i> : <i>iAiIuI</i> +	$\underline{iAiIuI} + \underline{iB} \in \beta$

Table 1. The Three Basic Conventionalized Intensional Profiles and Their Shared Basis (essentially based on Alberti and Kleiber 2014)

This general conception is presented in a formalized way in Table 1 as an underspecified intensional profile with the name “target-oriented mentalization”. This fictive profile serves as the shared basis for the intensional profiles which define the three major sentence types (to be regarded as well-formalized Gricean felicity conditions and as other social rules), primarily in the process of constructing our comprehensive mental system of conventionalized intensional profiles. Supposing intensional profiles rests on the idea that in the course of language acquisition infants obtain so meagre data, at least compared to the high complexity of the system, that its acquisition requires that they often have recourse to such general methods of creating (lacking) truth values in profiles as compositionality (in the Fregean/Montagovian sense), opposition, and transferring values.

The profile of target-oriented mentalization describes a person concentrating on a state-of-affairs e , whose $[-5, +5]$ scale in the iB -dimension is exactly distributed into three disjoint intervals by the three major sentence types. Value -5 provides profile for the speaker who, aware of the fact that e does not hold, intends to change that state of things, by calling the listener for help using an imperative sentence. The situation in which $iB \in +5$ may stimulate readiness for cooperation: as information is valuable, supplying the listener with e , which the speaker knows to be true, is likely to serve the listener’s interest. The situation in which $iB \in [-4, +4] = “0”$ can be construed as follows: the speaker is not in a position to carry out the former two types of action so their obvious aim can be to reach one of these states ($iB \in +5$ or $iB \in -5$); this can be initiated by taking the addresser role of a yes-or-no question.¹

The second step in the profile of target-oriented mentalization concerns the addressee. The knowledge that belongs to the addressee role can be either the same as, or the opposite of, the knowledge of the addresser. The imperative will specify the background in the former way: it is on this – shared negative – basis that the speaker can call for joining forces in order to change the (unwanted) state of things. As for the latter way, what makes sense of the declarative type is exactly the listener’s uninformedness: $iBuB \neq +5$, with a speaker informed.² A similar opposition makes sense of the interrogative type, too: now it is the listener who is assumed to be informed: $iBuB \in \pm 5$, with the speaker uninformed. The informed status of u can mean both knowing that e is true or that it is false. The iB component in label $iBuB$ is responsible for mentalization (“What I think about you is that . . .”). Its value is “ $+5$ (i.e., the left half of a bell-shaped distribution), in

1 The symbols ‘ n ’ and ‘ n ’ denote a narrow and a broader interval around n in the following precise sense: the former symbol means a (bell-shaped) normal distribution over interval $[n-2, n+2]$ while the latter one a flatter normal distribution over interval $[n-4, n+4]$. Symbols ‘ n ’, ‘ n ’, ‘ n ’ and ‘ n ’ denote the left/right half of the corresponding normal distributions.

2 The given value pertains to the underlined part of the complex label in question (but underlining is omitted if this can cause no misunderstanding). If a pair or triplet belongs to a complex part of a label, the corresponding values are connected by the symbol ‘•’.

the case of worldlets $iBuD$ and $iBuA$, too. That is, the speaker's ideal position is to have sure knowledge on the listener's given attitude (+5). The worst (still acceptable) case can be formulated as follows: it arises in the speaker's mind as a possibility ($iBuX=+1$) that the given attitude X is such that is prescribed in the given intensional profile as the value of $iBuX$.³

The next question is as to what desire moves the speaker to the given speech act. It can generally be claimed that this desire pertains to an eventuality e' which has to be construed on the basis of e . Hence, in worldlets iD and $iBuD$, it is e' that is there to be evaluated. As for authority, the listener has entire authority (+5) over e' in an ideal case while in the worst case (+1) they might be able to execute e' ($iBuA\epsilon=+5$). In the case of an imperative, e' essentially coincides with e ; only the truth value of e should be reversed. As declaration and interrogation aim at transmitting some knowledge on e , e' should be defined on the basis of output information states $uB+$ and $iB+$. If, for instance, e is the state that someone is vegan, then e' is the event that [the appropriate interlocutor learns that s/he is vegan].

It has not been discussed yet which interlocutor's interest is to be served in the case of the major sentence types. It is hypothesized that the "discourse-markerless" basic case is when the decision has not been made but the speaker enforces some kind of summarized interest by using the major intensional profiles. In the table, the formula with summation is devoted to the formulation of this approach.

The general formulation of addresser's intention should be related to an eventuality e'' , also to be calculated on the basis of e , whose achievement is assumed to require the addressee's aid ($iIuI+$: [I intend you to intend e'']). In the imperative the ultimate intention will pertain to the resetting of the truth value of e in the external world, whilst declaratives and interrogatives serve the purpose of resetting the generalized truth value of e in certain interlocutors' output information states.

Another generalization formulates when a speaker can take the addresser role of an intensional profile. They are in a position to take it if their intention to influence the partner does not violate any criterion of authority ($iAiIuI+=+5$), or at least they think to have some argument for having this authority ($iAiIuI+=+1$).

3. The Experiment

3.1 Research Methodology

The data for the present research came from 53 voluntary participants (28 women and 25 men), all native speakers of Hungarian, and all university students at the same institution, aged between 18–24. The data were elicited through short pre-written dialogues that the

3 This approach to mentalization is very permissive and uniform; it expresses our experiences gained so far in the course of our research activity in framework $\mathfrak{ReALIS}$ which offer no support for the idea that the speaker should monitor the attitudinal dimensions uB , uD and uA differently.

participants had time to get familiar with before the recording. Each situation was read out loud in the form of a dialogue between each participant and the author of the present paper. The dialogues were recorded with a dictaphone (44,1 kHz/16 bit). The relevant data (extracts from the recordings) were analyzed in PRAAT.

Before the recordings were made, the participants had around ten minutes to read and comprehend the instructions and the situations. Then the leader of the experiment read out every situation again and began to act out role **A**. The 53 participants were asked to play role **B** in the situations. They were prompted to read out every single word and to act out their role, refraining from monotonous speech. The written texts did not contain any punctuation so as to not influence the intonation patterns of the participants.

3.2 The Five Functions Embedded in the Situations

The story had the following common frame with four different outcomes, depending on the answer of character **B**:

A and **B** are twenty-year-old university students in love with each other, who have been going to the cinema for a year on a weekly basis. They take turns in choosing the movies. They know each other's tastes quite well. This time it is **B**'s turn to make a decision. This is an excerpt from their dialogue. **B** has already browsed the cinema program and she has already made her decision, but she has not told **A** yet, who is eagerly waiting for the "announcement of the result". There are the following three movies competing:

- an English detective story nothing out of the ordinary,
- an Icelandic drama which seems depressive,
- an American comedy presumably full of dirty jokes.

The first movie stands as an expected answer. The speaker thinks that they have a common ground with the listener, as it will be explicated in the next two sections. The Icelandic drama is the test for the uneasy answer, detailed in section 3.2.3, with the speaker thinking that the listener does not want to see it. The last movie represents a generous-kind decision on the speaker's part, inasmuch s/he knows that the listener thinks that they have different preferences. Here the speaker will even have to give confirmation.

3.2.1 *Straightforward Answer* (Hát)

B is sure that s/he has made the single good decision and even **A** could not have made a better one. This opinion becomes clear from **B**'s argumentation itself which consists of true facts without any lies, distortions or pleasantries.

- (1) (a) A: Na, melyik filmet választottad?
 A: So, which movie.ACC select.PAST.2SG
 "A: So, which movie have you opted for?"

- (b) B: Hát a krimit
 B: Well/So the detective-story. ACC
 “B: The detective story, surely!”

- (c) Az mindkettőnknek be szokott jönni
 Az ízléstelen vígjátékokból a múltkor végképp elégünk lett
 a nyomasztó északi drámákért pedig még én sem rajongok
 bár én alapjában véve kedvelem a komolyabb műfajokat

“That (kind of movie) works for both of us.
 Last week we got utterly fed up with these tasteless comedies.
 As for depressive dramas, even I am not very keen on them.
 Although I am fundamentally fond of serious genres.”

3.2.2 *Uncertain Answer* (Háát)

B is not sure if s/he has made the best decision. S/he feels that based on the current cinema program there is no decision which could be argued for enthusiastically. S/he thinks that their mutual experience is such: In the past, several detective stories proved boring but acceptable. **A** tends to choose comedies, but last week even s/he himself/herself was shocked by the tastelessness of the movie s/he had opted for. Finally, **A** cannot tolerate dramas.

- (2) (a) A: Na, melyik filmet választottad?
 “A: So, which movie have you opted for?”

- (b) B: Hát a krimit
 B: Well/So the detective-story.ACC
 “B: Well, the detective story . . . ”

- (c) Az azért többé-kevésbé mindkettőnknek be szokott jönni
 Az ízléstelen vígjátékokból a múltkor már neked is eléged lett
 a nyomasztó északi drámákért pedig még én sem rajongok
 Vagy nagyon unod már a krimiket?

“That more or less works for both of us.
 Last week even you got utterly fed up with the tasteless comedies.
 As for depressive dramas, they do not enthuse even me.
 Or are you very tired of watching detective stories?”

At this point another important difference between this study and the previous corpus based one by Dér and Markó (2017) should be mentioned. Dér and Markó were looking for monofunctional uses of the discourse marker *hát* (Dér and Markó 2017, 11), as it was mentioned in the introduction. One of my most important arguments for using experimental data instead of corpus data in the investigation of the prosodic features of discourse markers is that it is impossible to obtain sufficient data if we want to control or synchronize as many variables as are generally needed in pragmatics research. More specifically, in this experiment based on a representationalist theoretical framework, Uncertainty could be controlled in the sense that it was not directed at the proposition but at the choice, which is a performative speech act (Austin 1975): the speaker makes a decision about the movie to watch. In this experimental case *hát* cannot be a hesitation element of the speech processing, which is a frequent ambiguity in corpus data. If the uncertain type is pronounced significantly longer than the straightforward one, it is only a marker of the difficulty of making a choice, but not that of processing. In spontaneous-speech-corpora, however, *hát* can be the marker of uncertainty about the propositional content of the sentence (e.g., “How old is your grandmother?” – “Well . . . she was born in 1928, sooo . . . – in which case the speaker is thinking while speaking, but it sounds the same as an uncertain *hát*).

3.2.3 Anxiety (Hátöö)

A further difficulty for corpus-data analysts comes from the fact that it is very difficult to separate the following two types of uncertainty in spontaneous speech: uncertainty with a common ground about the judgment of the proposition, and uncertainty resulting from the other one’s opinion (see example 3). In the situations mentioned above (3.2.1 and 3.2.2), the speaker and the listener shared the (un)satisfaction about the chosen movie, at least as far as the speaker could see.

B feels that detective stories are getting more and more boring. As for comedies, even **A** has recognized that they are no longer funny but rather disgusting; but what s/he (**A**) truly and deeply hates are depressive Northern dramas. **B**, however, has heard about this Icelandic drama from a university friend, a philologist. **A** is jealous of the guy, maybe not without a reason. **B** finally decides to exercise his/her right but with a deep concern . . . What if **A** wants to know who suggested that movie . . .

- (3) (a) A: Na, melyik filmet választottad?
 “A: So, which movie have you opted for?”
- (b) B: Hát a drámát
 B: Well/So the drama.ACC
 “B: Well . . . the drama.”

- (c) Tudom, hogy nem nagyon rajongsz ezért a műfajért
de mintha ezt az izlandi filmet valahol nagyon dicsérték volna
Asszem valami díjat is nyert valahol
A krimiket már kissé unom
a mostanában futó vígjátékok alpáriságából pedig a múltkor már neked is
eleged lett, úgy emlékszem

“I know that you are not very keen on this kind of movies
but this Icelandic one was praised . . . somewhere . . .
and I also think it received some kind of award.
I’m getting a little bored with detective stories
and as far as I can remember, last time even you got fed up with the
disgusting comedies.”

Again, the considered kind of uncertainty can be confused with uncertainty deriving from other difficulties in attempt to produce a formally ideal utterance. In our experiment, the anxiety of the speaker due to the listener’s contrary desire was considered, as an idealized basic case. In real discourse, however, this type of *hát* can be used even when the speaker is only anxious about word choice or concept usage, or the informativeness of the answer (cf. Gricean cooperative principle). Let us consider, for instance, the sentence “Why are places near the sea cooler in summer and warmer in winter than farther inland?” In the Hungarian version of the answer “Well . . . because of the specific heat capacity of water,” the pronunciation of *hát* (*Hátööö*) tends to coincide with that of the anxious *hát* discussed here. In which case the speaker is anxious about the concepts used and/or informativeness.)

“I assume that we qualify the decision in the same manner: both of us find it either straightforward (+5) or not (‘0’); and we know this about each other.”		
1. straightforward answer	2. uncertain answer	3. answer which the speaker considers uneasy or embarrassing
$iBuB+ = iBuB+iB$, and its value is in negative correlation with the duration of the vowel in the discourse marker <i>hát</i>		$iBuD+ = ''+5, -5'$ “I am aware of the fact that my choice will not make you happy (-5’).” “I am sure that you will realize whose interest (r) will be served by this choice (“+5).”
$'5$	$'0'$	
$iBuB+ = ''+5, \iota$; $iBuB+iB = ''+5, ''+5, \iota$		$iBuB+rD = ''+5, ''+5 \ r \in R$ “I want you to know (instead of pretending naïveté) that I am aware of the fact that my choice will not make you happy”
with values of ι the same (synchronizing minds)		$iIuB+iBuD+=+5, +5, '5, -5'$

Figure 1. The pragmasemantic analysis of the first three types of *hát*

3.2.4 Teasing and Confirmation (HáÁÁÁát and hát!)

B feels that detective stories are getting more and more boring but what **A** truly hates are depressive Northern dramas. S/He would get depressed, and s/he would probably badger her/him the whole evening that the freaking movie had been recommended by the hot friend from the university, the charming philologist. Therefore, **B** decides to choose the comedy (it cannot be as tasteless as last week's one), but now s/he (**B**) can afford to tease her/him (**A**) in exchange for her generous decision.

Her/him (**B**'s) goal is as follows: At first, **A** should not believe that s/he is prepared to watch a comedy after last week's nightmare. This should make **A** even happier and the evening will be great.

- (4) (a) A: Na, melyik filmet választottad?
 "A: So, which movie have you opted for?"
- (b) B: Hát a vígjátékot
 B: Well/So the comedy. ACC
 "B: Why, the comedy!"
- (c) A: A vígjátékot?!
- A: The comedy. ACC
 "A: Have you opted for the comedy?!"
- (d) B: A vígjátékot Hát
 B: The comedy. ACC well/so
 "B: The comedy, for sure!"
- (e) Tudom, hogy mennyire szereted ezt a műfajt és úgy szeretem, ha vidám vagy este
 "I know how much you enjoy Hollywood comedies, and you know how much I like it when you are happy in the evening."

4. teasing/badinage	5. confirmation, sentence-finally
iBuD+= '5,+5	
↑ "I am sure that my choice will make you happy." ↓ "At first, you will not believe that I have factually opted for the given version." ↓↓ "Later you will accept my choice (with pleasure)."	↓ "I am sensing that at the moment you are not absolutely sure that I was serious when I have opted for the given version." ↓↓ "I want you not to have doubts as to the choice."
iBuB+= ''+5, '-3''	iBuB=' +5 (and it also happens to be different from 0)
iluB++ = +5,+5	iluB+ = +5,+5

Figure 2. The pragmasemantic analysis of the last situation types of *hát*: the generous-kind decision and the confirmation

4. Statistics

4.1 Normal Distribution

From the duration of the five measured categories of *hát* for all 53 participants, only the sentence-final category turned out to have a normal distribution (see Table 2). Since the sentence-initial types stood in the same position in the sentence, the aim was to investigate for any statistically significant difference in their prosodic characteristics – especially concerning duration (measured in milliseconds), a reliably measurable property. The lack of normality reduces the range of tests which can be applied.

	Kolmogorov-Smirnov*			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
h1	0.205	53	0.000	0.784	53	0.000
h2	0.147	53	0.006	0.928	53	0.003
h3	0.186	53	0.000	0.845	53	0.000
h4	0.268	53	0.000	0.653	53	0.000
hf	0.077	53	0.200**	0.976	53	0.368

* Lilliefors Significance Correction

** This is a lower bound of the true significance.

Table 2. Tests of Normality

4.2 Exploratory Data Analysis

Table 3 and Figure 3 show that the two uncertain answers (h2, h3) were longer in respect of *hát* than the straightforward and the teasing one (h1, h4).

N		h1	h2	h3	h4	hf
	Valid	53	53	53	53	53
	Missing	0	0	0	0	0
	Mean	217.607	394.016	358.946	259.941	336.032
	Median	203.038	355.018	286.802	213.670	340.233
	Std. Deviation	61.0431	173.61	188.327	154.78902	65.876
	Range	338.16	652.546	871.811	786.615	385.419
	Minimum	140.465	156.071	154.184	121.728	168.856
	Maximum	479.281	808.617	1025.995	980.343	554.275

Table 3. Descriptive Statistics of the five measured categories of *hát*

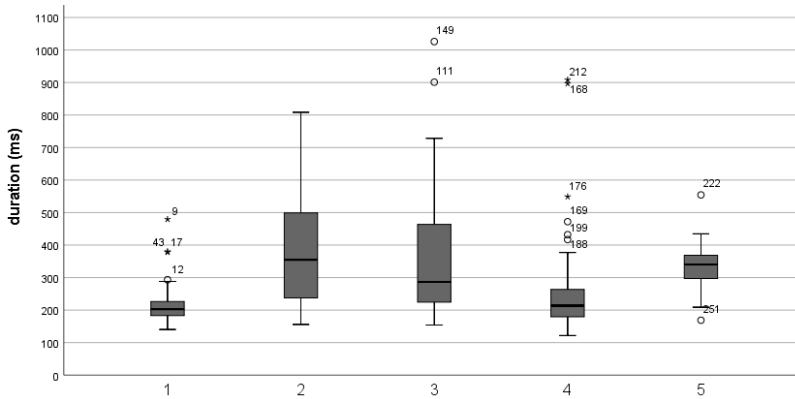


Figure 3. Simple boxplot of five measured categories of *hát*

The confidence intervals in Figure 4 show that h2 and h3 are hard to separate so the hypothesis that they have different distributions is to be rejected. Variants h1 and h4 are also very similar in respect of their means; the radically different standard variants, however, suggest that they have also different distribution functions (see also Table 3). Our interpretation is that it is only for a subgroup that h4 serves as the holder of the teasing intonation in the experiment. Many participants put the teasing intonation on the other part of the sentence. This leads to a binominal distribution with a high standard deviation in the case of h4.

This phenomenon is shown also in Figures 5 and 6, which illustrate the relation of h1 and h4, and the relation of h1 and h2. In the first case some participants who produce a regular, very short h1, produce then a longer h4, which stands for teasing, which is, however, is not the typical case, because a lot of them teased instead while saying the sentence fragment *a vígjátékot!* “the comedy”.

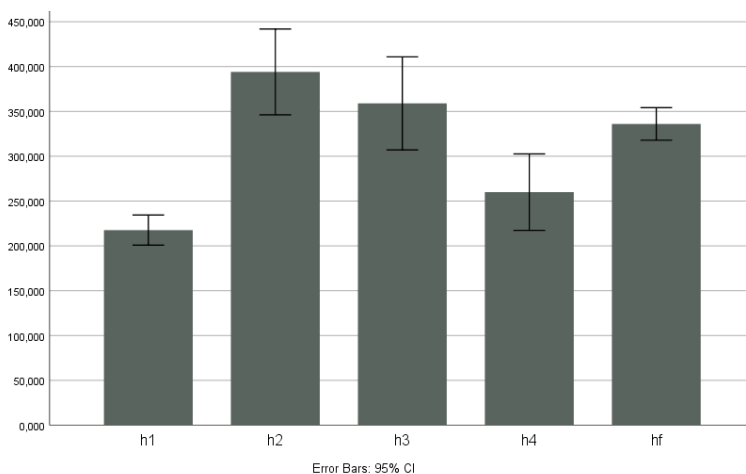


Figure 4. Bar charts with 95% confidence intervals

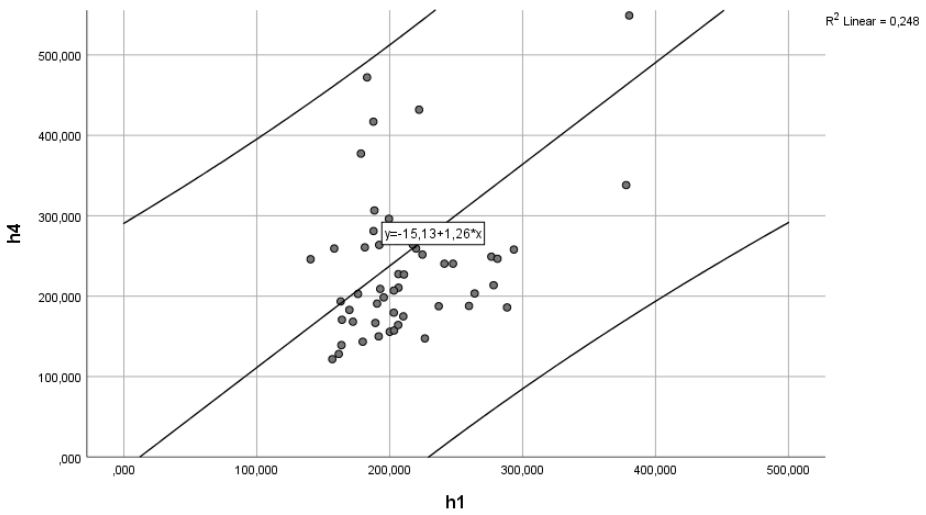


Figure 5. The duration of h4 in the function of that of h1

In contrast, the uncertain *hát* (h2), as present in Figure 6, functions as the regular holder of the uncertainty marker in the sentence: it was produced spectacularly longer by the participants than the short h1 values.

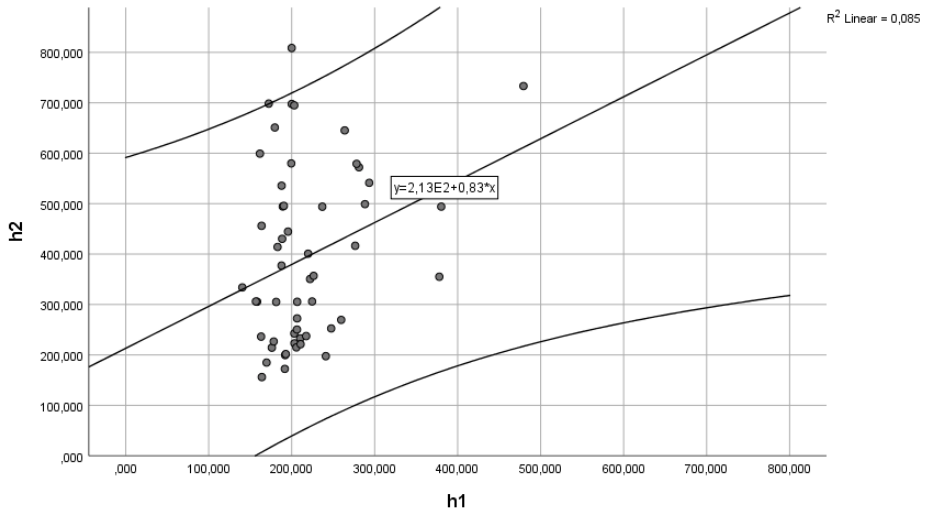


Figure 6. The duration of h2 in the function of that of h1

4.3 Significance Tests

Because of the lack of normal distribution except for the sentence-final *hát*, Friedman's ANOVA was used, which is essentially the non-parametric version of the one-way repeated-measures ANOVA. Also, commonly z-score method was used to identify justifiable outliers. Friedman's test proved the significant difference between the groups by 46 interlocuters, without 7 outliers, so the exact differences were pointed out through Wilcoxon's post hoc test with Bonferroni's correction that the data sets were significantly different in the case of the pairs $h2 - h1$, $h4 - h2$, $h4 - h3$ and $h3 - h1$.

	h4 - h1	h2 - h1	h3 - h2	h4 - h2	h4 - h3	h3 - h1
Z	-,366 ^b	-5,610 ^b	-1,983 ^c	-4,529 ^c	-4,190 ^c	-4,998 ^b
Asymp. Sig. (2-tailed)	,714	,000	,047	,000	,000	,000

Table 4. Wilcoxon Signed Ranks Test (b. Based on negative ranks; c. Based on positive ranks)

4.4 The Problematic Pairs

Two pairs were not significantly different according to Friedman's ANOVA, so the method of Loftus and Masson (1994) was used to equalize the means between participants. The new error bars were created in consideration of this factor, and $h1$ and $h4$ turned to have different confidence intervals, as shown in Figure 7.

Furthermore, the variance of homogeneity was also significantly different for the pairs $h1 - h4$, $h5 - h2$ and $h5 - h3$.

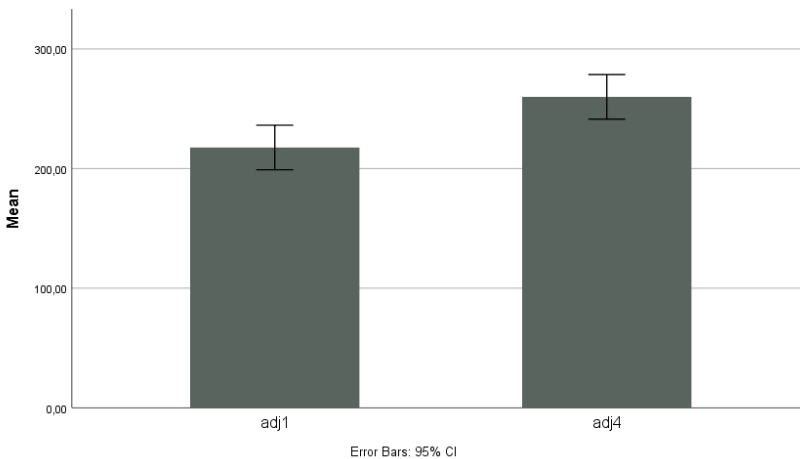


Figure 7. Adjusted confidence intervals for $h1$ and $h4$, based on the method of Loftus and Masson (1994)

Finally, concerning the difference between the properties of the uncertain and the anxious type of *hát* (h2 and h3), it was problematic to pinpoint a sharp difference between the two functions. Only a tendency was found for the anxious type of the discourse marker to be pronounced shorter and to have a longer pause following it. Although we could identify this difference in the pilot-study, we found no significant difference in this respect across the data in the present research. Nevertheless, another qualitative tendency first explored in the pilot-study could be observed in the big sample. Filled pauses were often found to follow the anxious type of *hát*, while these were not a typical property of the uncertain type. Further research is needed in this direction.

5. Conclusions

As for the research methodology employed in the study, it can be concluded that the read-aloud protocol is a useful method to elicit and imitate the general functions of the Hungarian discourse marker *hát* – at the same time, it eliminates the annoying effects of spontaneous speech processing where monofunctional uses of *hát* are very hard to find.

The findings of the present study point to the same kind of differences that were identified by Dér and Markó (2017). There is a measurable function of the discourse marker – with the inferential-conclusional core meaning – which indicates uncertainty, but the filled pause after it may belong to another type, one that indicates that the speaker is anxious.

Furthermore, the results confirmed that the discourse marker *hát* is able to express complex attitudes such as, for example, teasing. Based on the two strategies identified in the research (teasing by using *hát* and/or using the last syllable of the sentence), it should be noted, however, that *hát* does not obligatorily carry such attitudes; it is a rather peripheral phenomenon when it does.

According to my claim, the same applies for the anxious function of *hát*. Here, the filled pause following the discourse marker is the main indicator of the anxiety, while the extremely multifunctional discourse marker *hát* – or any discourse marker, in general – tends to consider their core meaning more while being multifunctional.

The suprasegmental characteristics pointed out by this research can be useful to explore later even by auto-segmented corpus-data, though the data of my experiment was segmented manually.

Funding Acknowledgement

The present scientific contribution is dedicated to the 650th anniversary of the foundation of the University of Pécs, Hungary. I am grateful for the financial support of the University of Pécs (EFOP343). Special thanks are due to the members of Research Team 9ReALIS for Theoretical, Computational and Cognitive Linguistics, especially to Gábor Alberti, Ákos Gocsál and Mónika Dóla for their useful remarks.

Works Cited

- Alberti, Gábor. 2011. “*ReALIS: Interpretálók a világban, világok az interpretálóban*” [*ReALIS: Interpreters in the World, Worlds in the Interpreter*]. Budapest: Akadémiai Kiadó.
- Alberti, Gábor. 2016. “Hát a *meg* meg a *hát*” [Well, the *and* and the *well*]. Kas (ed.): In “*Szavad ne feledd!*”: *Tanulmányok Bánréti Zoltán tiszteletére* [“Don’t forget your word!” Papers in honour of Zoltán Bánréti], edited by Bence Kas, 17–27. Budapest: Research Institute for Linguistics of the Hungarian Academy of Sciences.
- Alberti, Gábor, and Judit Kleiber. 2014. “*ReALIS: Discourse Representation with a Radically New Ontology*.” In *Complex Visibles Out There. Olomouc Modern Language. Proceedings of the Olomouc Linguistics Colloquium 2014: Language Use and Linguistic Structure*, edited by Ludmila Veselovská and Markéta Janebová, 513–528. Olomouc: Palacký University.
- Alberti, Gábor, Judit Kleiber, Zsuzsa Schnell, and Veronika Szabó. 2016. “Intensional Profiles and Different Kinds of Human Minds: ‘Case studies’ about Hungarian Imperative-Like Sentence Types.” *Linguistics Beyond and Within* 2: 6–26.
- Alberti, Gábor, and László Nőthig. 2015. “*ReALIS2.1: The Implementation of Generalized Intensional Truth Evaluation and Expositive Speech Acts in On-going Discourse*.” *International Journal on Advances in Intelligent Systems* 8 (1&2): 85–106.
- Alberti, Gábor, Noémi Vadász, and Judit Kleiber 2014. “Ideal and Deviant Interlocutors in a Formal Interpretation System.” In *The Communication of Certainty and Uncertainty* edited by Andrzej Zuczkowski, 59–78. Amsterdam: Benjamins.
- Austin, John L. (1962) 1975. *How to Do Things with Words*. Oxford: Clarendon Press.
- Bell, David M. 1998. “Cancellative Discourse Markers: A Core/Periphery Approach.” *Pragmatics* 8 (4): 515–541.
- Dowty, David R., Robert E. Wall, and Stanley Peters. 1981. *Introduction to Montague Semantics*. Dordrecht: Reidel.
- Dér, Csilla Ilona, and Alexandra Markó. 2017. “A *hát* funkciói a prozódiai megvalósulás függvényében” [Functions of *hát* “well, so” with Relation to Its Prosodic Realization]. *Beszédkutató* [Researching Speech Production] 25: 115–117.
- Grice, Paul. 1975. “Logic and Conversation.” In *Syntax and Semantics*, Vol. 3, ed. by Peter Cole and Jerry L. Morgan, 41–58. New York: Academic Press.
- Groenendijk, Jeroen, Martin Stokhof, and Frank Veltman. 1996. “Coreference and Modality.” In *The Handbook of Contemporary Semantic Theory*, edited by Sh. Lappin, 179–213. Oxford: Blackwell.
- Kamp, Hans, Josef van Genabith, and Uwe Reyle 2011. “Discourse Representation Theory.” In *Handbook of Philosophical Logic* 15, edited by Dov M. Gabbay and Franz Guenther, 125–394. Berlin: Springer.

- Leiss, Elisabeth. 2014. "Modes of Modality in an Un-Cartesian Framework." In *Certainty–Uncertainty – and the Attitudinal Space in between*, edited by Sibilla Cantarini, Werner Abraham, and Elisabeth Leiss, 47–62. Amsterdam: Benjamins.
- Lauer, Sven. 2013. "Towards a Dynamic Pragmatics." PhD diss., Stanford University.
- Loftus, G. R., and M. E. J. Masson. 1994. "Using Confidence Intervals in Within-Subject Designs." *Psychonomic Bulletin and Review* 1 (4): 476–490.
- Maier, Emar. 2016. "Attitudes and Mental Files in DRT." *Rev. Phil. Psych* 7: 473–490.
- Nuyts, Jan. 2017. "Evidentiality Reconsidered." In *Evidentiality Revisited: Cognitive Grammar, Functional and Discourse-Pragmatic Perspectives*, edited by Juana Isabel Marín Arrese, Gerda Haßler, and Marta Carretero, 57–83. Amsterdam: Benjamins.
- Searle, John. 1979. *Expression and Meaning – Studies in the Theory of Speech Acts*. Cambridge: Cambridge University Press.
- Schirm, Anita. 2007–2008. "A *hát* diskurzuszjelölő története" [History of the Discourse Marker *hát*]. *Nyelvtudomány* III–IV: 185–201.
- Schirm, Anita. 2013. "A diskurzuszjelölők funkciói: a *hát*, az *-e* és a *vajon* elemek története és jelenkori szinkrón státusa alapján" [The Function of Hungarian Discourse Markers: The history and Present Synchronic Status of the Hungarian Elements *hát*, *-e* and *vajon*]. PhD diss., University of Szeged.
- Szeteli, Anna, and Gábor Alberti. 2017. "The Pragmasemantic Contribution of the 'Noise Word' *hát* in Hungarian, Which Can Mean Everything and Its Opposite." Paper presented at the Budapest Linguistics Conference, Hungary, June 2017.
- Oishi, Etsuko. 2017. "Austin's Speech Acts and Mey's Pragmemes." In *Pragmemes and Theories of Language Use*, edited by Keith Allan, Alessandro Capone and Istvan Kecskes, 335–350. Cham: Springer.
- Wimmer, Heinz, and Josef Perner. 1983. "Beliefs about Beliefs: Representation and Constraining Function of Wrong Beliefs in Young Children's Understanding of Deception." *Cognition* 13: 103–128.