
*Ilona Csilla Dér –
Alexandra Markó*

**The suprasegmental marking of
Hungarian discourse markers**

Introduction: Pragmatical research (Hungarian language)

Balogh (2000)

Bańczerowski (2000, 2005a, 2005b)

Dér (2005, 2006)

Sz. Hegedűs Rita (2001, 2002)

Keszler (1998, 2000)

Kugler (2000, 2002, 2003)

Péteri (2001)

Simigné Fenyő (2005)

Introduction: Discourse markers

diskurzusjelölők, discourse connectors, discourse connectives, discourse particles, discourse operators, cue phrases, etc.

Def.: Lingual-pragmatic units that provide information about the segmentation and operation of a discourse, that is, they show the discourse structure (cf. Fraser 1999: 931, Louwense–Mitchell 2003: 199).

E. g. They signal the introduction of a new topic or a change in topic:

- (1) **Tényleg**, *hogy sikerült a vizsgád?* 'As a matter of fact, how did your exam go?'
- (2) **Mellesleg** *a baleset mikor történt?* 'By the way, when did the accident happen?'

Introduction: Discourse markers

- Louverson and Mitchell (2003): discourse markers come up ten times more often in spoken discourses than in written ones
- functional group
- heterogeneous regarding their origin:
 - different word classes (adverbs, connectives, verbs etc.)
 - different levels of language (lexemes, syntagms etc.),
 - nonverbal discourse markers (Schiffrin 1987: 328; about the discourse functions of humming in Hungarian see Markó 2005).

Introduction: Discourse markers

- syntactically isolated from the rest of the sentence
- prosodically independent: they are separated from their context by pauses or commas (Jucker–Ziv 1998: 3, cited by González 2004: 43–44; Zwicky 1985: 303–304, cited by Fraser 1999: 933).
- The results of international experimental phonetic research are quite contradicting: Hirschberg and Litman (1993: 516, cited by Aijmer 2002: 33): *well* proved to be a prosodic unit only in 50% of its occurrence.

Introduction: Discourse markers

Hansson (1999):

Pauses come up before the Swedish discourse marker *men* 'but'/'and' and *så* 'so' quite often, but after them very rarely (*men*: 90% and 17%; *så*: 59% and 8%).

This investigation

- concerns only a few elements (n = 42 and 51);
- the difference between short and long pauses were very small.

Hansson says: it is not only one but **several prosodic characteristics together that help the listener in the identification of discourse markers.**

Introduction: Discourse markers

Pause: distinctive feature in the definitions of discourse markers.

Bevezető kifejezések ('introductory expressions'): the term suggests that discourse markers appear at the beginning of discourse segments. → Is it true?

The present research's aims

- i) Checking the statements in literature about
 - the **turn-initial position** of DMs and
 - the **preceding and/or following pause** in Hungarian.
- ii) Investigating some Hungarian **homophones** to find the **cue(s)** which help(s) the listener to distinguish the DM and the S (sentence) function of the same lexical form.
- iii) Listing the Hungarian DMs.

Material and subjects

- Corpus = spontaneous conversation with four people: 114 minutes; recorded directly to PC in a radio studio (mono, 22 050 Hz, 16 bit)
- The informants: four 21-year-old university students: two females and two males (speech time per person is between 12'01" and 36'01")
- Task: to talk about whatever they want – there were no instructions

Methods and tools

- i) Annotation of the corpus (labeling of turns, text, pauses, etc.) – Praat.
- ii) Collecting all the DMs in the corpus.
- iii) Investigation of
 - DM's position in the turns and
 - pauses before and after DM-s – Praat.
- iv) Investigation of homophones (phonologically the same form in DM function and sentence function):
 - position in the clause;
 - temporal parameters – Praat, SPSS;
 - F_0 and sentence accent – Praat.

Method

Excluded:

- commentary pragmatic markers (attitude markers; interjections, pragmatic idioms; cf. Fraser 1999: 942–943).

Kept:

- hesitation elements (which has a function of saving time or keeping the right of speaking)
- all representatives of turn-taking

Method

7 homophones:

így 'so'

meg 'and'

most 'now/FILLER WORD'

tehát 'so, thus'

tényleg 'really, actually'

úgyhogy 'so that, so'

vagy 'or'

- **high frequency** (*így, most, tehát*) – left out DMs
egyébként 'by the way, however', *érted* 'you know, you see', *mondjuk* 'let's say', *szóval* 'well, that is', *tudod* 'you know' and *úgy* 'so, that way/FILLER WORDS'
- **various sources:** adverbs, connectives, pronouns (finite verbs also fell out because of their small number)

Method

Problem 1: segmentation of discourse: turns, sentences, clauses, syntagms

Position

B = at the beginning of the clause

I = inside of the clause

E = at the end of the clause

predicate → clause

meg ilyesmi ‘and things like that’, *meg ilyenek* ‘and such as these’, *meg mit tudom én* ‘and I don’t know what’, *vagy mikor* ‘or whenever’, *vagy nem tudom* ‘or I don’t know’ → clause

Method

Suspected omission of verb as predicate (→ clause):

Piros óvodát láttam, meg piros iskolát.
red nurseryACC see PAST. Sg. 1 and red schoolACC
'I saw a red nursery, and a red school'

But:

Láttam piros óvodát meg piros iskolát.
see PAST Sg. 1. red nurseryACC and red schoolACC
'idem'

Method

Problem 2:

Position B?

most 'now/FILLER WORD'

elhatároztuk akkor ezt kipróbáljuk, | [hiszen]

***most** annyira úgysem ciki az egész*

'we decided to give it a try | [because] **DM** the whole thing is not so embarrassing'.

Method

Problem 3: identification of connectives as discourse markers

- common linking function
- Polanyi (2001: 266–267): discourse segments = clauses; all connectives are discourse markers at the same time
- Fraser (1999: 938–942): only those connectives are discourse markers which join topics and messages.

We agreed with the latter view, and we set a high value on investigating only those markers that clearly qualify as discourse markers.

Method

Problem 4:

Turn- and clause-**closing** elements: discourse markers or connectives?

Speaker's intention:

- stopped talking?
- wanted to give the word to another speaker?
- wanted to end her/his sentence?

E. g.: *úgyhogy* 'so that, so'; *tehát* 'so, thus' in E position

Method

Problem 5:

- different stages of the process of becoming a discourse marker
- many of them serve more, diachronically developed discourse functions

For example:

így 'so, this way'; *úgy* 'so, that way' (both of them are filler words as discourse markers)

Method

Problem 6: Collocations

Discourse markers which appear after each other:

- two different markers?
- a new collocation?

E. g.:

hát meg 'well and', *és akkor* 'and then', *tehát így* 'so [+filler word]', *és így* 'and so', *de egyébként* 'but otherwise', *most így* '[filler word+] so')

Method

Decision:

we did **not** investigate:

- collocations ~ markers appearing after each other
- markers that do not stand after each other but appear in the same clause **and** have the same discourse function

Explanation:

- the same function and the subsequent position can influence the analysis of prosody
- we are not sure whether we are dealing with a real discourse marker collocation or not

Homophone *meg* ‘and’

Excluded cases: second discourse marker (a former connective) not directly next to *meg* but in the same clause and with the same function:

s	<i>Andrisba</i>	meg	<i>megbízok</i>
AND	AndrewINFL	AND	trust PRES Sg. 1.
'And I trust Andrew/And in Andrew, I trust'			

Homophone *tehát~tát* ‘so, therefore, namely’

Similar to *vagy* ‘or’: can be a reformulation marker, a connective, but it has other discourse functions

Whether the speaker

- was correcting her-/himself?
- tried to gain time to think?
- used *tehát* in a linking function?

Homophone *tehát~tát* ‘so, therefore, namely’

Examples:

*mert ott kell írni, **tát** mellette*

‘because we have to write it there, DM beside it’

*viszont nem írhattunk máshova **tehát** csak jegyzőkönyvbe*

‘but we were not allowed to write it elsewhere

CONNECTIVE only into the minutes’

*mert nyolcadiktól fölfele tanít **tehát** kilencedikeseket*

‘because he teaches upwards from the eighth grade

CONNECTIVE nine graders’

*mindenkinek a nevére is emlékszem meg ilyesmi **tát**’*

‘I remember everyone’s name and things like that DM’

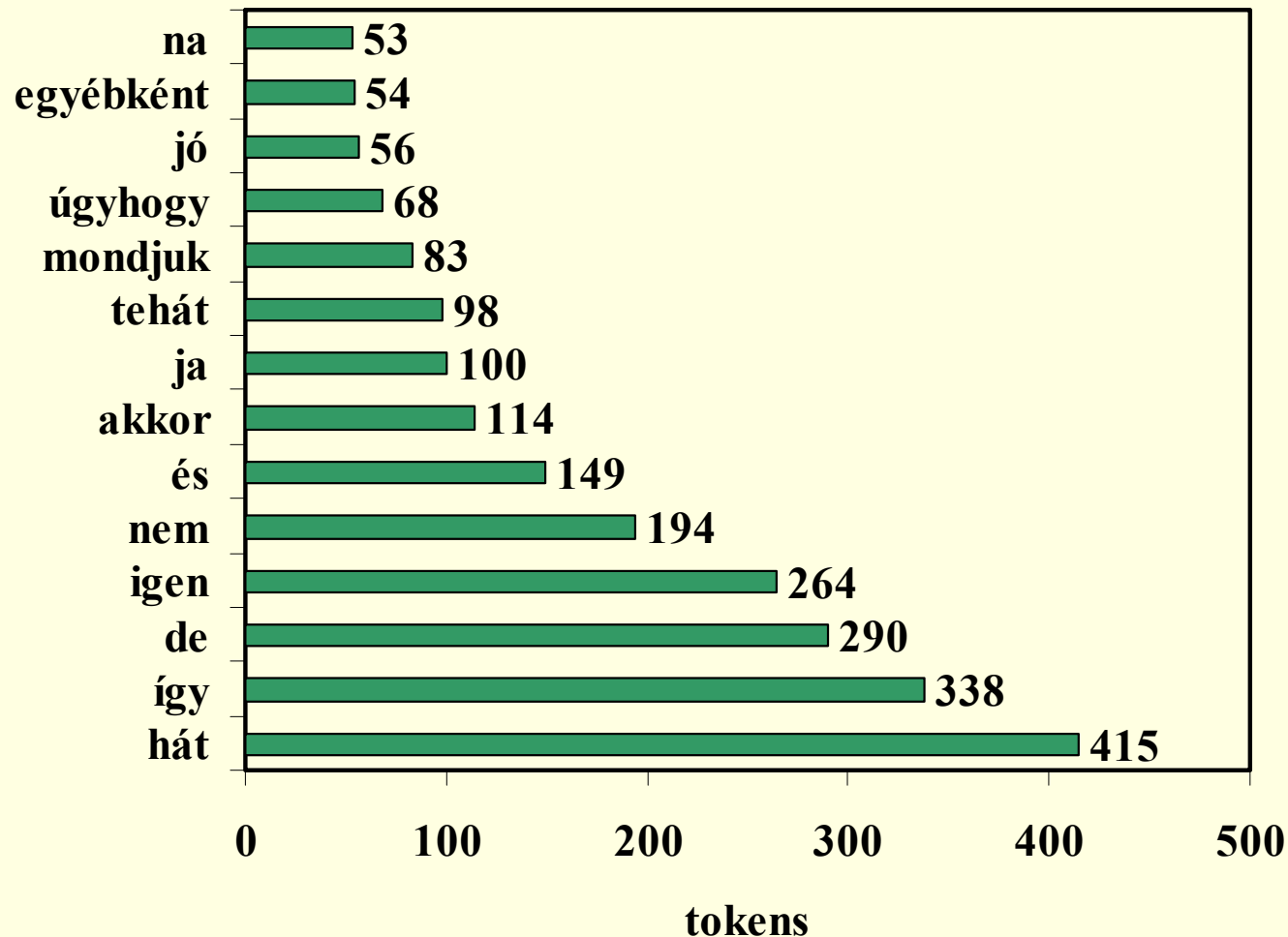
Results: DMs in the corpus

73 types, 2853 tokens, average 25 per minute:

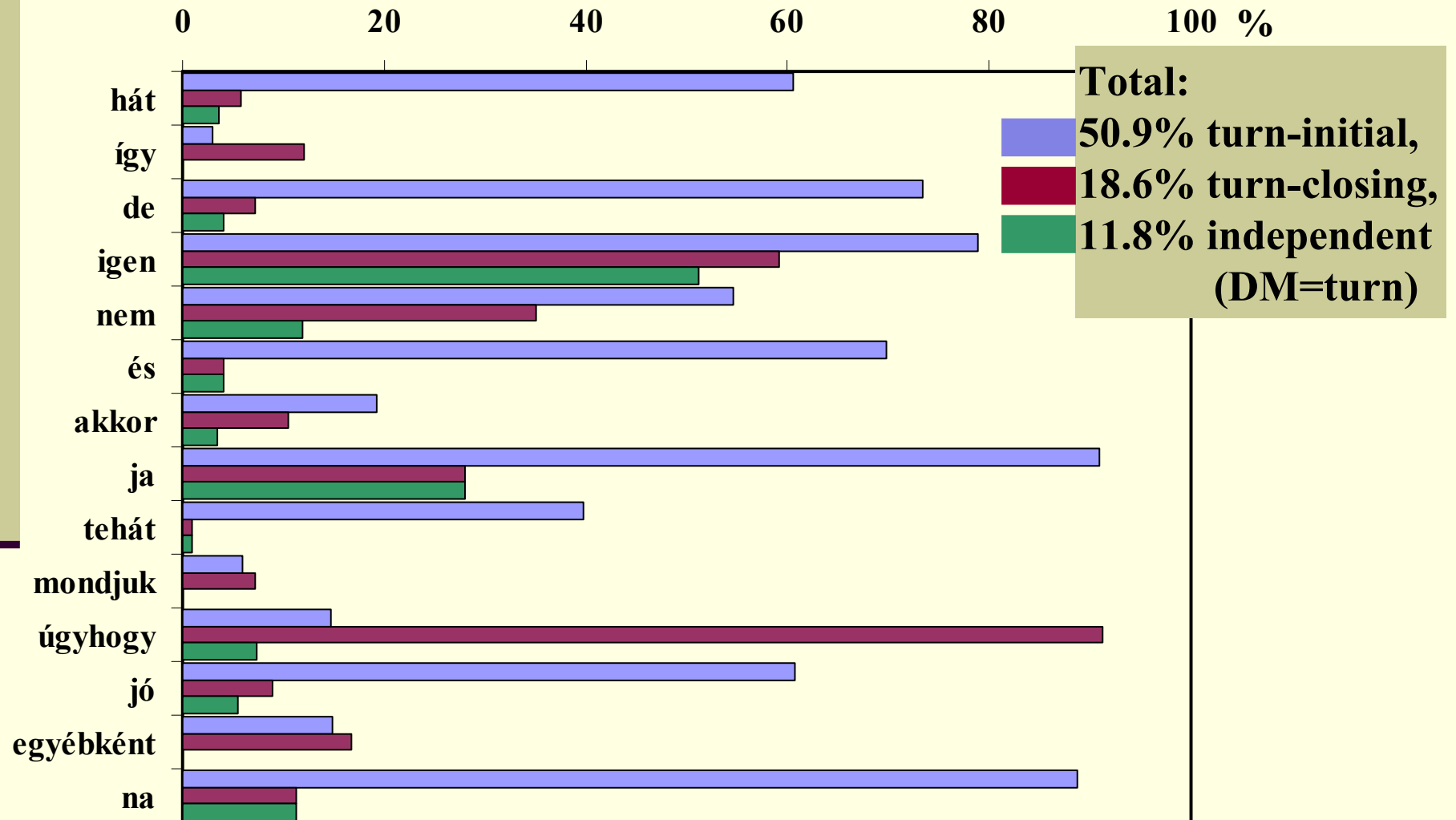
á; a másik meg; aha; akkor; azért; aztán; bár; bocsánat, hogy beleszólok; de; de hát; de igen; egyébként; egyik; érted; értem; és; ha; hát; hm; hm?; hogy; hogyha; igen; így; illetve; inkább azt mondom; itt; izé; ja; ja igen; jaj; jó; látod; mármint; mármint hogy; meg; még azért visszakerdezek, hogy; még így visszatérve; mert; mit is akartam; mmm; mmm ('not' humming); mondjuk; most; na; nem; nézd; oké; pedig; például; pláne; s; sőt; szeretnék szólni, hogy; szóval; tehát; tényleg; tessék; tudniillik; tudod; úgy; úgy értem, hogy; ugye; úgyhogy; ühüm; üüm ('not' humming); vagy; vagy hogy mondjam; vagyis; vagyis hogy; várj/várjál; viszont; visszatérve.

Results: the most frequent DMs in the corpus

Types ≥ 50 = 80% of the total amount of DMs



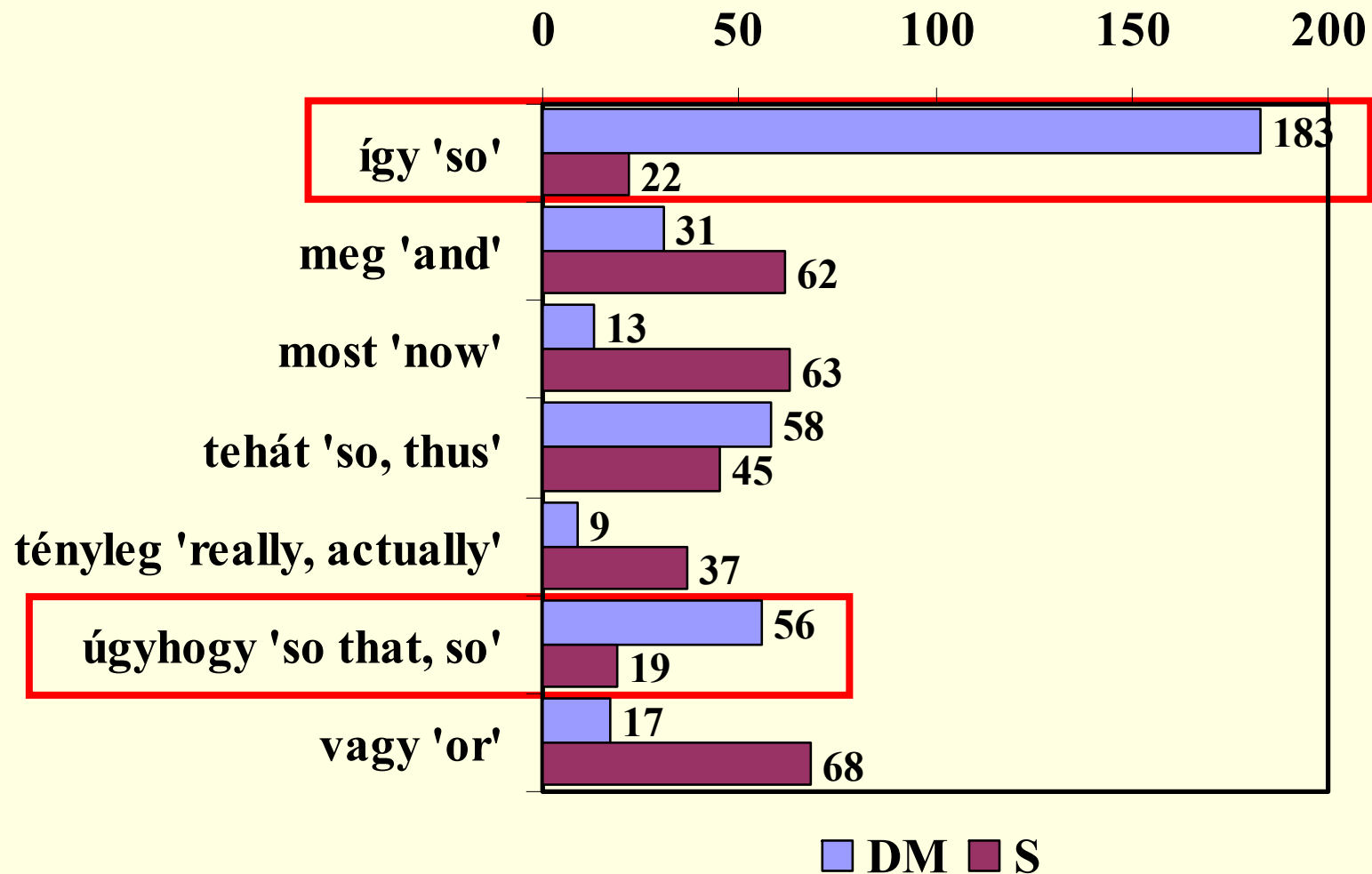
Results: Position in the turns



Results: Preceding and following pauses

- 5.5% of DMs is preceded by pause (from 0.0 to 23.5% – *tehát*).
- 8.2% of DMs is followed by pause (from 0.0 to 41.7% – *ugye*).

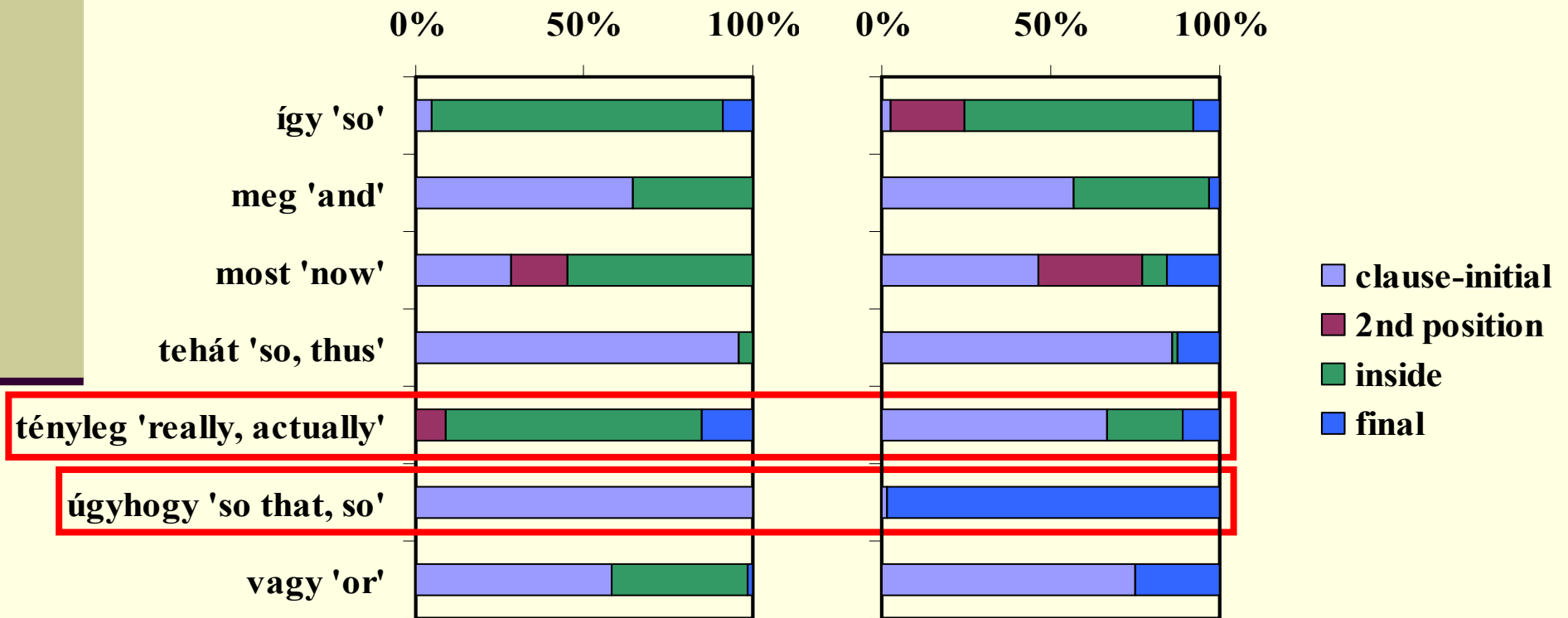
Results: Homophones



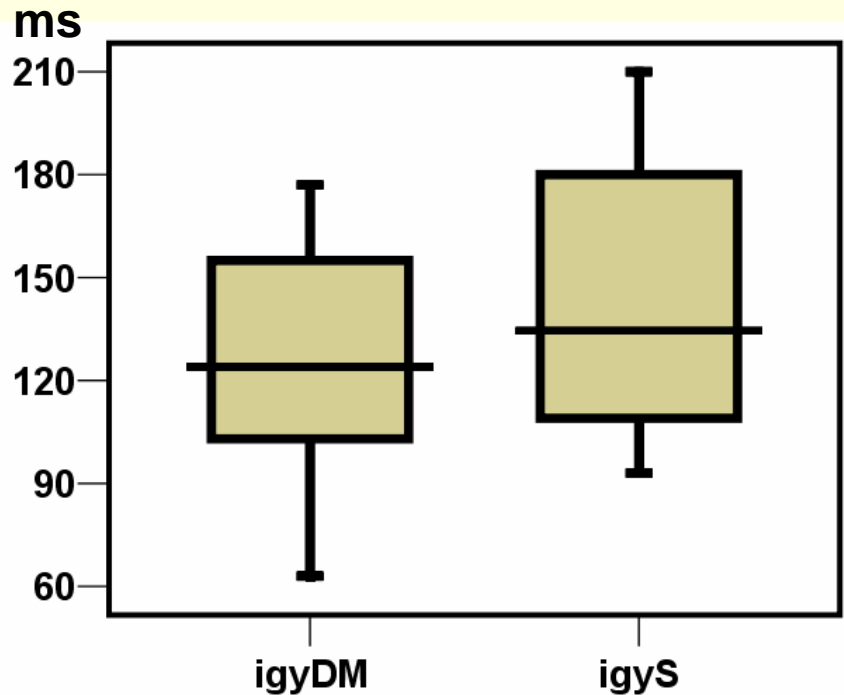
Results: Homophones – position in the clause

Sentence function

DM function



Results: Homophones – word length

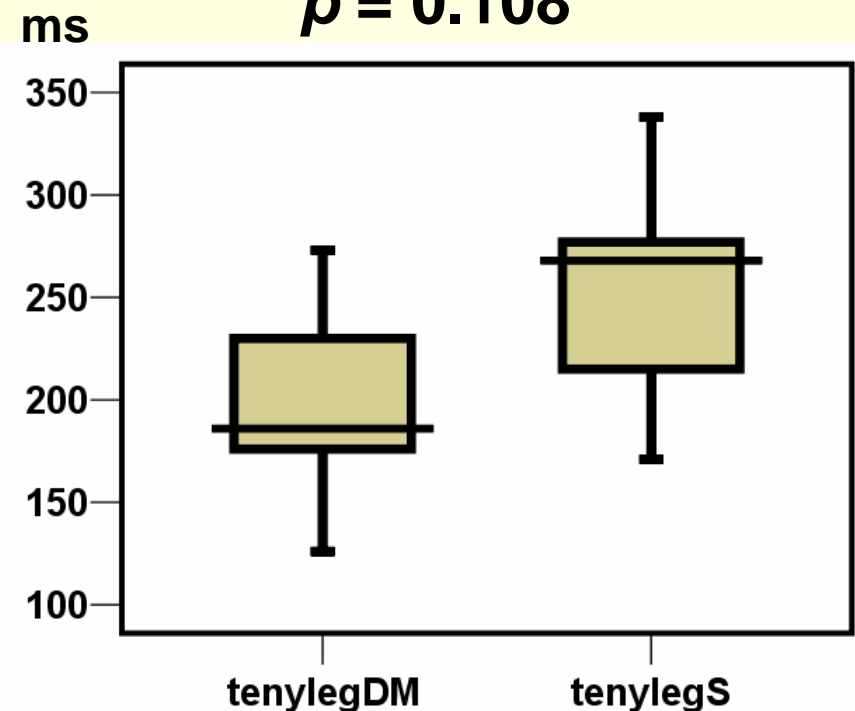


így 'so' [i:ʝ, i:ʝə, iʝ, i:c, ic]

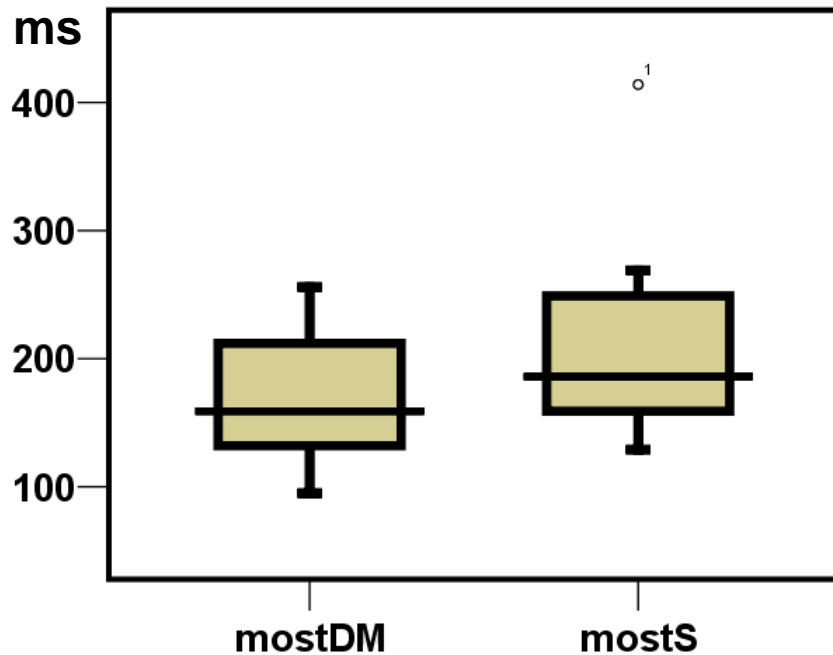
$p = 0.056$

tényleg 'really, actually'
[te:l:ɛg, te:l:ɛk, te:lɛg,
te:lɛk]

$p = 0.108$



Results: Homophones – word length

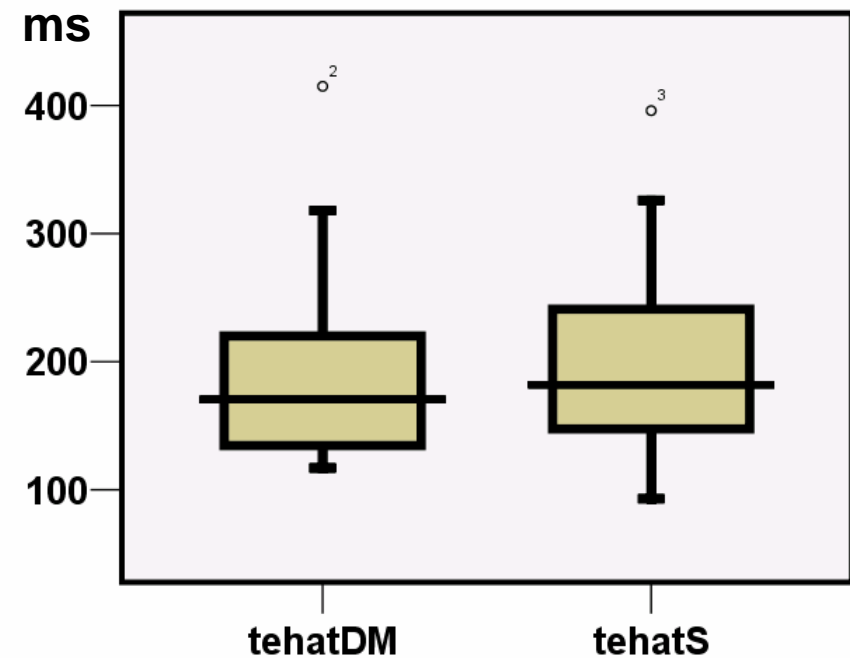


most 'now' [moʃt, moʃ, moʒ]

$p = 0.124$

tehát 'so, thus' [tɛha:t, tɛhat, tɛat, ta:t, tat, tət, ta:, ta]

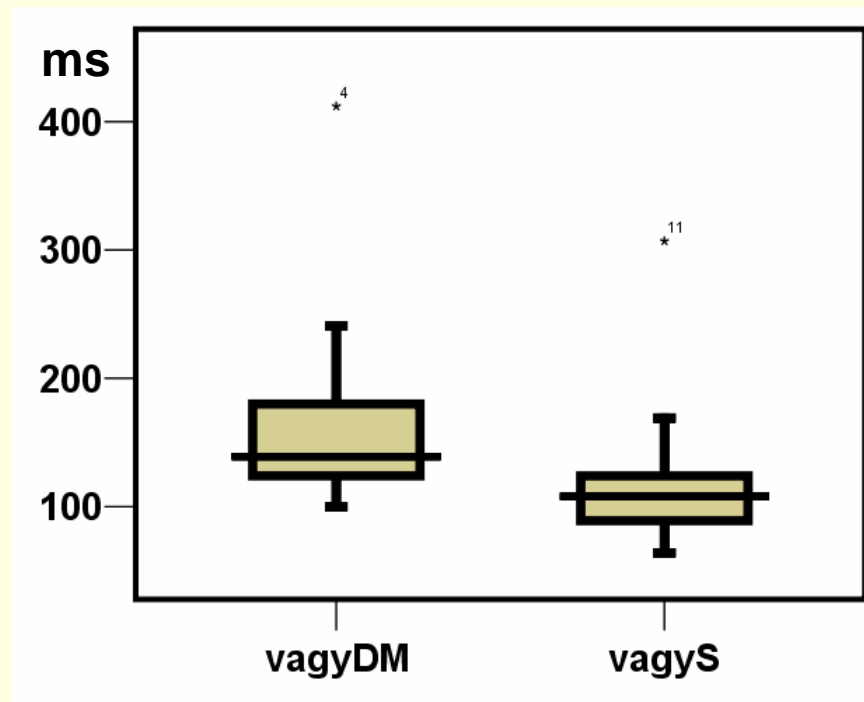
$p = 0.067$



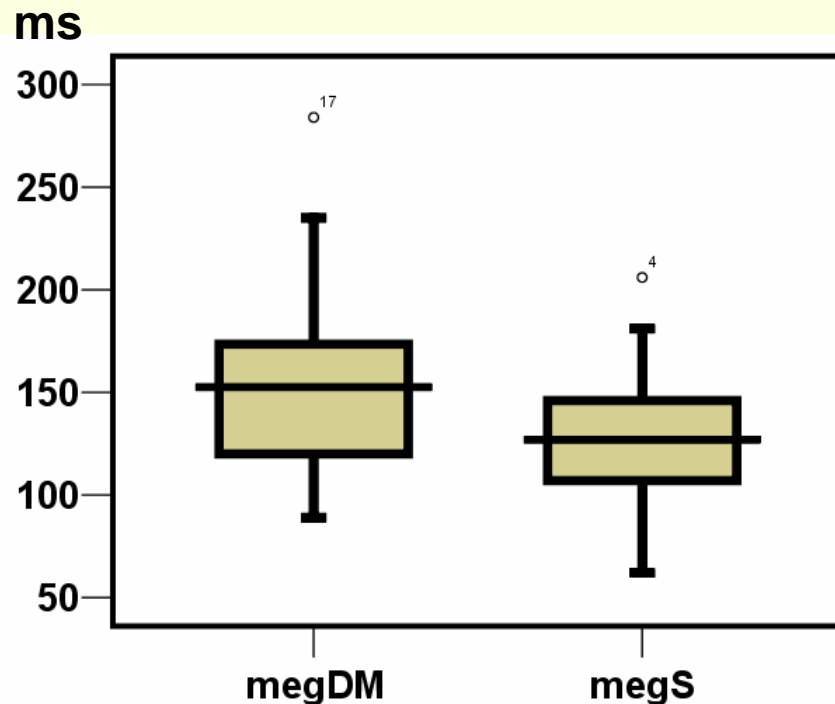
Results: Homophones – word length

vagy 'or' [vɔɟ, vɔɕ, vɔɟə]

$p = 0.056$



Results: Homophones – word length

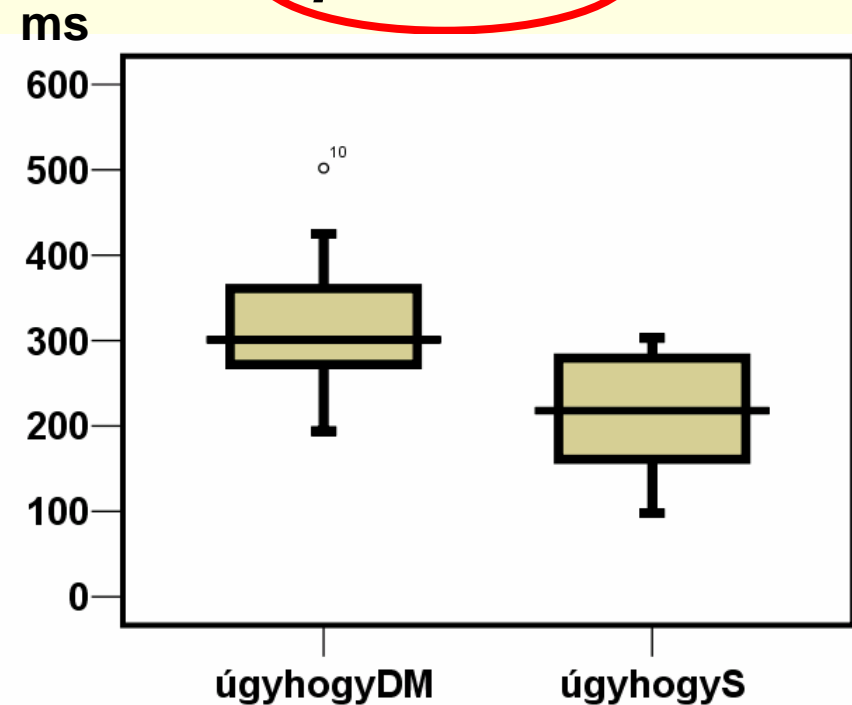


meg 'and' [mɛg, mɛk, mɛ]

$p = 0.045$

úgyhogy 'so that, so'
[u : choʃ, u : choc, uchoʃ, uchoc,
u : cho, ucho, choʃ, cho, ch]

$p < 0.001$



Results: Homophones – F₀ and stress

F₀: no regular difference between DM and S versions of homophones

Accent: remarkable difference in two cases

	DM	S
<i>így</i> ('so')	no stress	22.7% stressed
<i>tényleg</i> ('really, actually')	11.1% stressed	50.0% stressed

Conclusions

- i) **The data do not or only partly support** the statements of the literature about
 - the **turn-initial position** of DMs and
 - the **preceding and/or following pauses** in Hungarian.

- ii) **It is not true** that for the Hungarian **homophones** the investigated parameters are **unambiguous cues** (either alone or together) for the listeners to distinguish the DM and the sentence function.

Conclusions

Then how can the listeners distinguish DM and S – if they can at all?

- Methodological problems ➡ the decision is not easy.
- In some cases (*tényleg*, *úgyhogy*) the position in the clause can help.
- The average word duration of *meg* and *úgyhogy* is significantly longer in the DM function ➡ gaining time.
- The stress can help in the case of *így* and *tényleg*.

Conclusions

A new point arises: the speaker's usual manner of speaking.

