



Sz'al, nemtom 'So, I dunno'

- The formal and functional variation of the discourse marker **NEM TUDOM** 'I don't know' in Hungarian

CIFU XIII., University of Vienna, 26. 08. 2022.

Dér, Csilla Ilona (ELTE DiAGram Research Group of Functional Linguistics)





Main clause > discourse marker (DM)

- Typical grammaticalization/pragmaticalization path (also in Hungarian):
main clause (predicates)> sentence adverb/conjunction/particle/DM
(Haader 2001, Dér 2013)
- DMs: *I think, you know, I mean, I suppose, I don't know*
- H. *asszem* 'I think', *tudod* 'you know', *nemtom~nem tudom* 'I don't know'
- still exist as main clauses (bound vs. unbound forms), see pictures (complex sentences without and with the conjunction *hogy* 'that')

(Sources: <https://memegenerator.net/instance/73779050/philosoraptor-j-lenne-mert-nemtom-mit-vegyek-fel>, <https://memesgenerator.hu/meme/nem-tudom-hogy-ki-vagy-es-hogy-hol-vagy/>)

I DON'T KNOW

English: Tsui 1991, Pichler 2007, 2013 (others see: Pichler 2007)

German: Helmer & Deppermann 2017

Picher (2007): variational pragmatics (discourse variation analysis),
Berwick-upon-Tweed (near to the Scottish-English border)

I don't know~I divn't knaa~I dinnae ken~I dono~I dunno

I don't know (45%) > *I dunno*, *I divn't knaa* (22% each) > *I dono* (9%) > *I dinnae ken* (2%) **for referential uses**

I dunno (58%) > *I divn't knaa* (22%) > *I don't know* (11%) > *I dono* (8%)
> *I dinnae ken* (1%) **for all pragmatic uses**

(Pichler 2007: 182)

(Source: <https://memegenerator.net/instance/55514255/i-dunno-girl-well-ierah-i-dunno>)

WELL, I...ER.....AH....



memegenerator.net



Functions of unbound I DON'T KNOW

PREDICATE (MAIN CLAUSE): ?

- Referential use 'I don't have the knowledge' – **exist!**

DISCOURSE MARKER:

- **Textual (pragmatic) uses:**
 - repair
 - hesitation
 - turn-exchange
 - **Epistemic** (subjective/intersubjective) (**pragmatic**) uses:
 - doubt, tentativeness, uncertainty
 - mitigation (of disagreement)
- + mixed cases (e.g. subjective-textual)



(Pichler 2007)

Source: <https://hu.pinterest.com/pin/introspective-pug-is-introspective--217369119486071176/>



Example: mitigation

(1)

Luke: *For the kids that are on drugs I blame the parents me.* @

HP: *Why.*

Luke: *I divn't knaa I think they're just (?) they're no looking after their kids properly or they just (.) just letting them get away wi it.*

(Pichler 2007: 179)



Example: repair

(2)

Daniel: *No well (.) I s- I dunno. I keep saying if if it's so good where you came from, why don't you go back. That's what I say, you know.*

(Pichler 2007: 179)



Material, method

2 unbound forms: (én)/ *nem tudom* /(én) ~ (én) / *nemtom* / (én)

- **BEA** = Magyar Spontán Beszéd Adatbázis [Hungarian Database of Spontaneous Speech]. Subcorpus of 30 **spontaneous conversations** with 3 speakers
 - Size: ~8 hours (~37.000 words)
 - all speakers from **Budapest**
 - <http://www.nytud.hu/adatb/bea/index.html> [gender, age]
 - A (main speaker, 15 women & 15 men, 20–71 ys, 3 age groups: 20–35 ys /youngsters/, 36 – 50 ys /middle-aged/, 51–71 ys /older speakers/ - in each groups: half are male, half female)
 - T1 (experiment leader, mostly the same woman, 29 ys)
 - T2 (conversational partner, 2 men, 2 women in their 20s) (Gósy 2012; Horváth et al. 2019)
- **MNSz2** = Magyar Nemzeti Szövegtár 2. változat [Hungarian National Corpus 2.0.5]. Size: 1.04 billion tokens. [subcorpora by regions (Hungary, Slovakia, Subcarpathia, Transylvania, Vojvodina), by registers (spoken press language – news&reports, written press language, fiction, academic, political, private) <http://clara.nytud.hu/mnsz2-dev/>
 - 250–250 random sample hits (Oravecz et al. 2014)

Method: discourse variation analysis (Pichler 2013), 1 expert annotator

Main question: is the shorter form more pragmaticalized?



I. *nem tudom* in MNSz2

Total 331 hits, 250 hits random sample – valid hits: **only 27 unbound cases!** → **not typical as unbound form in today's Hungarian**

- all from **personal** subcorpus (social media)
- 77% on the **left periphery**, 11% on the right, 11% independent (Degand 2014)
- 3x in DM-combinations: *hát, úgyhogy, szóval* 'well, so, therefore'
- Main functions: **60% referential**, 19% uncertainty

Mixed: uncertainty&mitigation (of disagreement)

(3) 2013-09-23 leti0417 *Egyértelmű h a legszebb város Los Angeles és New York!!* 2013-09-23 Nina Costaninova *Oké, de nem ez volt a kérdés mind1. Isztambul?? (...)* 2013-09-23 rednail **nem tudom**, biztos Los Angeles?

'2013-09-23 leti0417 It is clear that the most beautiful cities are Los Angeles and New York!! 2013-09-23 Nina Costaninova Okay, but that was not the question nevermind. Istanbul?? (...) 2013-09-23 rednail **I don't know**, surely Los Angeles?'

(MNSz2, #1014978429, doc#2718, Hungary, pers-soc)



II: *nemtom* in MNSz2

- written in 1 word (cf. *nem tom*)

Total 2421 hits, valid (out of 250): **59 valid (unbound) hits**

- personal subcorpus: **internet forums** (53%), social media (39%)
- punctuation problems: which periphery? 8,5%
- 68% on the **left** periphery

referential:

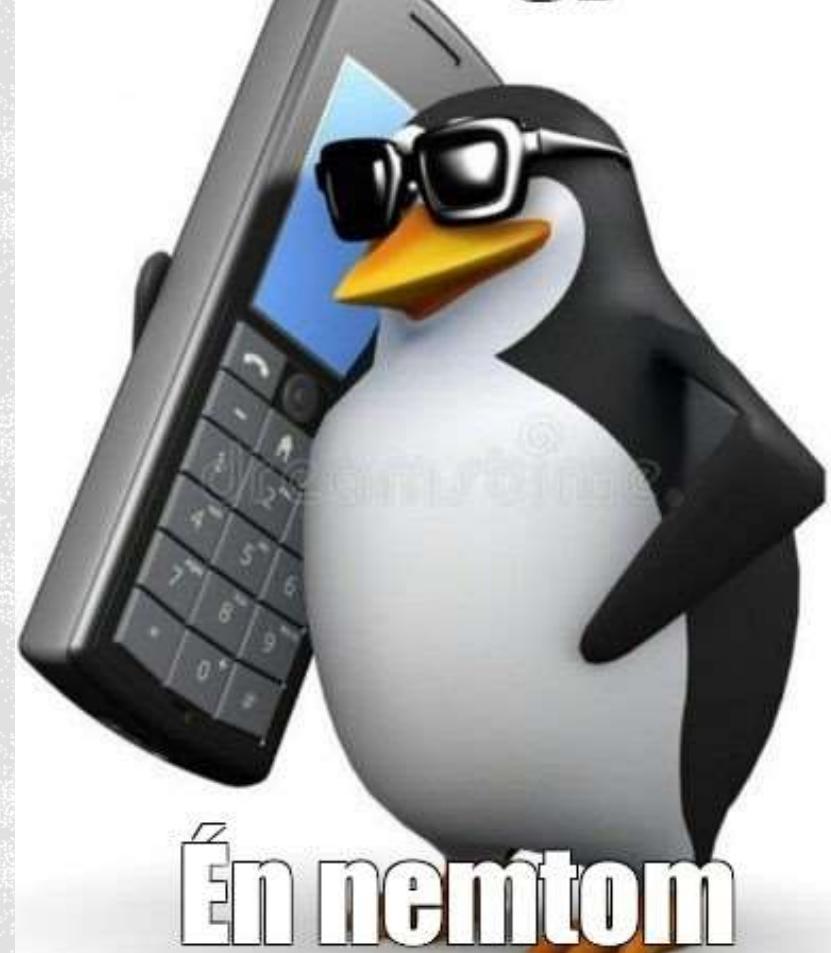
(4) 2013-09-26 xcbiker (...) A kihangosítás milyen ennél? 2013-09-26 HAL9001 **Nemtom**, van külön kihangosítóm.

'2013-09-26 xcbiker (...) How's the speakerphone on this one?
2013-09-26 HAL9001 I dunno, I have a separate speakerphone.'

(MNSz2, #1103493216, doc#2805, magyarországi, személyes-közösségi)

(Source: <https://www.meme-arsenal.com/en/create/meme/3136157>)

Hát figyi



II. *nemtom* in MNSz2

24% on the right periphery (2x as for *nem tudom*), mostly with the DM *vagy* 'or' – very similar to disjunct **general extenders** (Overstreet–Yule 2021), e. g. *or something; and stuff; and things (like that); and everything*

H. *meg minden* 'and everything', *meg ilyenek* 'and stuff/things like that'

'there are **other/relevant/obvious options** we can mention'

(5) 2013-12-29 *rolikaaa2000* @ Olh_Tndi jaaa...:D és nem találod fel magad semmivel?
esetleg alvás, rajzolás, filmnézés, ha van tesó akkor tesó akkor beszéljetek **vagy nemtom**

'2013-12-29 *rolikaaa2000* @ Olh_Tndi yaaah...:D and you don't come up with anything?
perhaps sleeping, drawing, watching movies, if there is a brother, then brother, then talk **or I dunno**'

(MNSz2, #1228756474, doc#2928, Hungary, pers-soc)

- **Functions:** 44% (of all valid hits) expresses **uncertainty**
- mixed with mitigation (25%), referential (20%)



I. *nemtom* in MNSz2

- Combinations with *hát* 'well, so', *vagy* 'or', *szóval* 'so', *amúgy meg* '(but) anyway' – rarely!

(6) [topic: weaning from trolling] 06/11/99 15:50:14 p.m.s.fej azé egy kissé durva beszólás volt de mint leszoktatás ... ***hát nemtom***

'06/11/99 15:50:14 p.m.s.fej that was a bit rude, but as a weaning ...**well, I dunno'**

(MNSz2, #108094955, doc#1002, Hungary, pers-forum)



I. *nem tudom* in BEA

- 118 hits: 34% by A (main speaker), 42% by T1 (female, 29 ys old), 24% by T2

Speaker A:

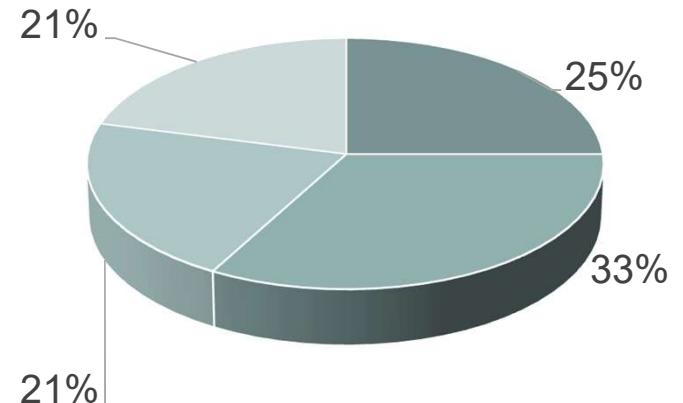
- 10% difference in favor of women
- 47,5% was produced by speakers in their 30s, 25% in their 20s, 30% in their 50–60s (av. 39,55 ys)
- positions: much more variable than in MNSz2

very mobile – why? pragmatalization: position can reflect DM-status (no correlation with functions)

(7) A: *lesz csak öö nem tudom valami más neve van ne-ne- nem tudom*

'A: there will be just uh I don't know it has another name I-I-I don't know'

(bea037, female, 20 ys)



■ left p. ■ right p. ■ inner ■ ???



I. nem tudom in BEA



- Functions: 10% mitigation, 25% referential, **32,5% uncertainty** → *nem tudom* (ÉN) gives „good enough” example (cf. *mit tudom én* ‘how should I know’)

(8) A: (...) fontosabb tehát az hogy **nem tudom én** Pistike szívét átültették jó ez is egy nagyon fantasztikus dolog de akkor inkább erről szól vagy valami botrányról

’A: (...) so the more important thing is that **I don't know** Pistike's heart was transplanted, it's also a very fantastic thing, but then it's more about this or about some scandal ’

(bea044, female, 31 ys)

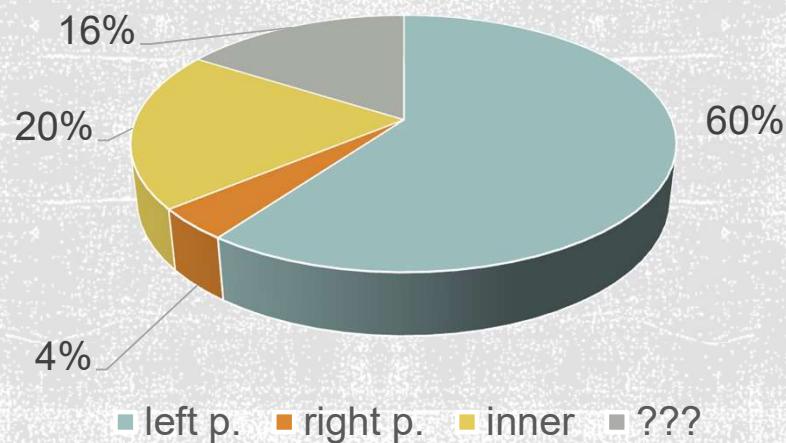
- DM-combinations: 47% (*vagy* 'or', *hát* 'well, so', *meg* 'and, but', *tényleg* 'really, indeed, by the way')

(Source: <https://hu.pinterest.com/pin/314337248966840446/>)



I. *nemtom* in BEA

- Total 62 hits (half as much as *nem tudom*)
- 40% by A (main speaker), 48% by T1 (female, 29 ys old), 12% by T2
- Only by speaker A:
 - 60% produced by males
 - Av. 29,24 ys



(Source: https://sargacsekk.blogspot.com/2014/02/csak-ugy-felrobbant_8576.html)



II. *nemtom* in BEA

Functions: **uncertainty** (40%), referential (8%), **hesitation** (25% - produced by the same young male speaker), mixed (27%, turn-taking&mitigation, turn-taking&uncertainty)

(9) A: *mi ööö azt csináljuk hogy hogy a termosztátot nemtomén tavaly huszonkét fokra volt beállítva idén huszonegy fokra van beállítva*

'A: we uh what we are doing is that that the thermostat **dunno** I was set to twenty-two degrees last year but this year it is set to twenty-one degrees'

(bea046, female, 57 ys)

(10) A: *na én most jártam arra és ööö nemtom tát ott elmegyünk egy éttere- ott ööö sz- Donovalyba egyetlenegy normális étterem van*

'A: well, I've been there and **uh I dunno so** we're going to a restaurant there uhm there's only one normal restaurant in Donovaly'

(bea031, male, 21 ys)

- Combinations: 52% (*hát* 'well, so', *és* 'and', *tehát* 'therefore', *de* 'but')



Statistics: no differences!

Speaker A	<i>nem tudom/nr of words</i>	<i>nemtom/nr of words</i>
female (15 ps)	0,001101377	0,000500626
male (15 ps)	0,001049318	0,000874432

Speaker A	<i>nem tudom/nr of words</i>	<i>nemtom/nr of words</i>
21–35 ys (12 ps)	0,001478054	0,0007068954
36–50 ys (6 ps)	0,000302709	0,00
51–70 ys (11 ps)	0,000906893	0,0002267231

Mann-Whitney test: no significant difference regarding gender (*nem tudom*: $U = 98,000, p > 0,05$; *nemtom*: $U = 109,000, p > 0,05$)

Kruskal-Wallis test: no significant difference regarding age (*nem tudom*: $H = 3,269, p > 0,05$, *nemtom*: $H = 4,012, p > 0,05$)



Summary

MNSz2

MNSz2	<i>nem tudom</i>	<i>nemtom</i>
<i>register/genre</i>	social media	social media, forums
<i>preferred periphery</i>	left	left
<i>dominant functions</i>	referential (60%) > uncertainty (19%)	uncertainty (44%) > referential (20%)
<i>combination with DMs</i>	rare (under 15%)	above 50%



BEA	<i>nem tudom</i>	<i>nemtom</i>
<i>age group/gender</i>	no connection	no connection
<i>preferred periphery</i>	???	left
<i>dominant functions</i>	uncertainty (32,5%) > referential (25%)	uncertainty (40%) > hesitation (25%)
<i>combination with DMs</i>	above 45%	above 50%



Conclusions

nem tudom/nemtom én 'I, me' is peripheric, probably idiolectal (3 hits in BEA, 1 in MNSz2)

Functions: no repair; hesitation: only with other functions together

- Is *nemtom* more pragmatalized than *nem tudom*?

MNSz: as expected

But: **contradictory results in BEA:** in the case of *nemtom*

- **stronger mobility as DM:** not true (on the contrary!)
- **like to combine with DMs:** nearly the same percent
- **dominant functions are pragmatic:** no big difference, but *nemtom* has more pragmatic functions, *nem tudom* was more typical in referential function (25% vs. 8%)

Why acts the longer form so strangely in conversations?

Next aims: collecting more data from spontaneous conversations, other annotators



Acknowledgements

This research was supported by the project NKFI K 129040
"Verbal constructions in Hungarian. Research in usage-based
construction grammar".



THANK YOU VERY MUCH FOR YOUR ATTENTION!



References

- Dér, Csilla 2013. Grammaticalization: A specific type of semantic, categorical, and prosodic change. *Berliner Beiträge zur Hungarologie. Schriftenreihe des Fachgebiets für Ungarische Literatur und Kultur an der Humboldt-Universität zu Berlin* 18. Berlin–Budapest. 160–179.
- Degand, Lisbeth 2014. *So very fast very fast then*. Discourse markers at left and right periphery in spoken French. In: Beeching, Kate – Detges, Ulrich (eds): *Discourse functions at the left and right periphery. Crosslinguistic investigation of language use and language change*. Leiden–Boston: Brill. 151–178.
- Gósy, Mária 2012. BEA: Beszélt nyelvi adatbázis. In: Gósy, Mária (ed.): *Beszéd, adatbázis, kutatások*. Budapest: Akadémiai Kiadó.
- Haader Lea 2001. Mikrodiakrónia és változásvizsgálat (az összetett mondatokban). *Magyar Nyelvőr* 125: 354–371.
- Helmer, Henrike – Deppermann, Arnulf 2017. ICH WEIß NICHT zwischen Assertion und Diskursmarker: Verwendungsspektren eines Ausdrucks und Überlegungen zu Kriterien für Diskursmarker. In: Blühdorn, Hardarik – Deppermann, Arnulf – Helmer, Henrike – Spranz-Fogasy, Thomas (Hg.): *Diskursmarker im Deutschen. Reflexionen und Analysen*. Göttingen: Verlag für Gesprächsforschung. 131–156.
- Horváth, Viktória et al. 2019. Háromfős társalgások annotálása a BEA-adatbázisban: elvek és kihívások. *Nyelvtudományi Közlemények* 115: 255–274.
- Oravecz et al. 2014. The Hungarian Gigaword Corpus. In: *Proceedings of LREC 2014*. 1719–1723.
- Overstreet, Maryann – Yule, George 2021. *General extenders. The forms and function of a new linguistic category*. Cambridge: Cambridge University Press.
- Pichler, Heike 2007. Form-function relations in discourse: The case of *I don't know*. *Newcastle Working Papers in Linguistics* 13:174–187.
- Pichler, Heike 2013. *The structure of discourse-pragmatic variation*. Amsterdam – Philadelphia: John Benjamins.
- Tsui, Amy B. M. 1991. The pragmatic functions of *I don't know*. *Text* 11: 607–622.

