

Why *almost* and *almost* ended up being different: The diachrony of approximatives in Hungarian

SUMMARY

In our paper, we will explore the diachronic syntax and semantics and syntax of two *almost*-approximatives in Hungarian: *majdnem* ('almost-M') and *szinte* ('almost-S'). We will show that there is a neat division of labour between these two: *majdnem* encodes epistemic vagueness (possible worlds, cf. Sadock 1981, Nouwen 2006 a.o) whereas *szinte* encodes scalar vagueness (precision standards, cf. Penka 2005, Sauerland and Stateva 2007, Amaral and del Prete 2010, Greenberg and Ronen 2013 a.o.). We will show that this can be straightforwardly derived from the grammaticalization trajectories of the two approximatives: *majdnem* derives from the adverb *majd* 'soon' (and the pleonastic negator *nem* 'no') and *szinte* derives from an adverb originally meaning 'by outward appearance, superficially'.

DATA

Hungarian has two words which are usually rendered as *almost* in English: *majdnem* és *szinte*. While at first sight, these may appear to be just stylistic alternatives, the reality is more complicated. Consider:

- (1) a. *János és Máté majdnem egyidősek.*
John and Matthew almost-M same.age_{Adj:PL}
- b. *János és Máté szinte egyidősek.*
John and Matthew almost-S same.age_{Adj:PL}
'John and Matthew are almost the same age.'
- (2) a. *Majdnem első lettem a futóversenyen.*
almost-M first became:1SG the running.competition:on
- b. *#Szinte első lettem a futóversenyen.*
almost-S first became:1SG the running.competition:on
'I almost came first in the running competition.'
- (3) a. *#A mozaikok majdnem véletlenül kerültek Ravennába.*
the mosaics almost-M by.accident got Ravenna:into
- b. *A mozaikok szinte véletlenül kerültek Ravennába.*
the mosaics almost-S by.accident got Ravenna:into
'The mosaics ended up in Ravenna almost by accident.'

As (1) to (3) show, *szinte* and *majdnem* are felicitously interchangeable in some but not in all contexts. (2b), if uttered, would probably be met with a rejoinder along the lines: 'Well, you cannot be almost the first, you are either first or not first, there is no in-between.' (3a) can only receive a felicitous reading in the rather absurd scenario where there was a plan to get the mosaics to Ravenna in an accidental manner, but in the end, this plan failed, and the mosaics got to Ravenna in a non-accidental manner.

We argue that these distributional differences (and others to be discussed in our talk) show that *majdnem* is an epistemic indefinite (the proposition is false in w_0 but true in an alternative world w_1 close to w_0 , cf. Piñón 2008) whereas *szinte* is a scalar indefinite (under the current precision standard, the proposition is untrue, but there is a slightly laxer precision standard under which it is true that John is 30 years old):

EPISTEMIC (one proposition, alternative possible worlds, distance between possible worlds):

- (5) a. POLAR $\neg p_{w_0}$
b. PROXIMAL $\exists w_1 \in S_{ALT}(w). \text{close}_s(w_1, w_0) \wedge p_{w_1}$

SCALAR (alternative propositions, one possible world, distance between precision standards):

- (6) a. POLAR $\neg p_{preC, w_0}$
b. PROXIMAL $\exists pre' \in S_{ALT}(preC). \text{close}_s(pre', preC) \wedge p_{pre', w_0}$

DIACHRONIC ANALYSIS

Majdnem (and its more formal variant *majdbogynem*) is the result of the amalgamation of the approximator *majd* (which is obsolete today but is still available in some set expressions such as *Majd elfelejtettem*. ‘I almost forgot.’) and the pleonastic/expletive neg-word *nem* ‘not’ (Simonyi 1888). The approximator *majd* is first attested in the Guary-Codex (before 1508, 027):

- (7) *zertelen tezőn magaul, majd el vezti feiet*
 unrestrained do.3SG with.himself almost PRT lose.3SG his.head.ACC
 ‘He fails to restrain himself, he almost loses his head.’

We will argue that the approximator *majd* ‘almost’ is derived from the temporal adverb *majd* ‘soon’ (Historical Dictionary of Hungarian 2:819): if p will be true soon, this means that there is a w1 close to w0 (a continuation world of w0) where p is true. This explains why *majdnem* ended up encoding epistemic vagueness. The appearance of pleonastic negation with the approximator *majd* is a later development, it was first attested in 1793:

- (8) *Az ördög vigye el, majd hogy a kórság belém nem áll.*
 the devil carry.IMP VM almost that the disease VM not stand.
 ‘To the devil with him, the disease almost knocked me down!’

Following É. Kiss (2010), I analyze this as follows: [CP *majd* [C’ *hogy* [AspP *a kórság belém nem áll*]]]. In sentences without a topicalized constituent and without a verbal modifier; *majd*, the optional complementizer *hogy* and the pleonastic negator ended up being in an uninterrupted sequence. This led to the reanalysis of *majd (hogy) nem* as the adverb *majdbogynem / majdnem*:

- (9) [CP *majd* [C’ (*hogy*) [AspP *nem ...*]]] -> [AspP *majd(hogy)nem* [AspP ...]]

The appearance of the pleonastic negator in sentences such as (8) is due to mirativity. In Hungarian, pleonastic negation is widespread in mirative exclamatives:

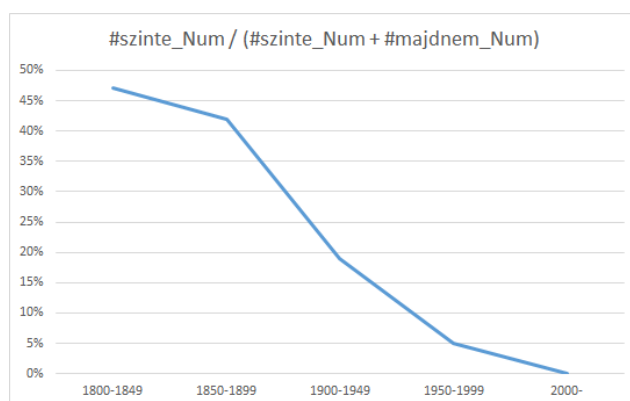
- (10) *Miket nem mondasz!*
 what not say.2SG

What extraordinary things you are saying!

Szinte derives from the word *szín* ‘outward appearance, surface’, and its original meaning was ‘by appearance, by outward similarity’ (Historical Dictionary of Hungarians 3:760). This adverb was easy to reinterpret as a scalar approximator indicating a more permissive precision standard, since identity by appearance clearly means identity by a less strict standard. Such a use is first attested in 1519. (Debrecen Codex, 177):

- (11) *olian kemeniön [...] mint szinte a száraz föld az o labai alatt*
 so hard as almost the dry land the he foot.3SG.PL under
 ‘Almost as hard as the dry land under his feet.’

Finally, we will discuss an ongoing development: in those environments where *majdnem* and *szinte* are both admissible and are, in a sense, in competition (such as before numerals), *majdnem* appears to be crowding *szinte* out:



We will argue that this is due to the fact that *majdnem* can freely combine with numerals suggestive of a precise measurement (such as *majdnem 190 cms* ‘almost 190cm’) while *szinte* cannot (*#szinte 190 cm* ‘almost 190 cm’). This means that *majdnem* has a slight edge over *szinte* in competitive environments, which can lead to the developments observed in the chart (cf. Niyogi 2002).