Voicing and reanalysis of Czech prefix and preposition *s*

Guy Tabachnick Markéta Ziková

Masarykova univerzita

Summary

- Czech orthography distinguishes between prefixes and prepositions *s(e)* and *z(e)*
- Voicing assimilation neutralizes the distinction between the two in some cases
- Work on Czech voicing assimilation typically assumes that orthography cues underlying form
- A pilot corpus study supports previous claims that some speakers are reanalyzing *s(e)* as /z(ɛ)/
- The shift is heavily lexicalized in the prefix, but only partially lexicalized in the preposition

Background

Two prepositions and prefixes *s(e)* and *z(e)* in Standard Czech, though the distinction between them is neutralized before obstruents:

(1) se dv-ěma hrad-y
sɛ dv-jɛma ɦrad-ɪ
with two-INS castle-INS.PL
'with two castles'

(2) ze dv-ou hrad-ů
zɛ dv-ow ɦrad-uː
from two-GEN castle-GEN.PL
'from two castles'

(3) s díl-y
z ɟiːl-ɪ
with part-INS.PL
'with parts'

(4) z díl-ů
z ɟiːl-uː
from part-GEN.PL
'from parts'

(5) směnit
smɲɛɲɪt
'exchange'

(6) změnit
zmɲɛɲɪt
'change'

(7) stěžovat si
scɛʒovat sɪ
'complain'

(8) ztěžovat si
scɛʒovat sɪ
'make difficult for oneself'

Eastern (Moravian and Silesian) dialects also voice obstruents at morpheme boundaries before sonorants, further neutralizing the distinction e.g. between (5) and (6)

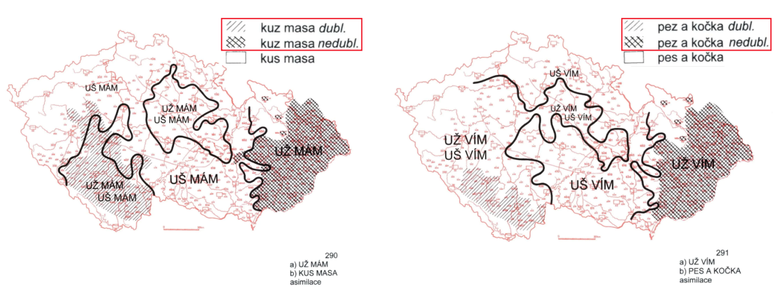


Figure 1. Geographic distribution of voicing at morpheme boundaries before sonorants and vowels (Balhar et al., 2016)

Historical developments

A second preposition *s(e)* 'off' has already merged with *z(e)* 'from'

Figure 2. Rate of *ze stolu* vs. *se stolu* 'off the table' in SyD (Cvrček & Vondřička, 2011)

A similar change in underlying voicing (without a merger) occurred much earlier from *ot(e)* to *od(e)* 'from'

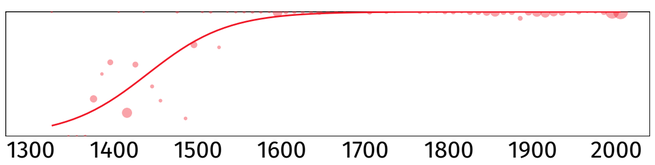


Figure 3. Rate of *ode mne/mě* vs. *ote mne/mě* 'from me' in SyD (Cvrček & Vondřička, 2011)

Comparison within West Slavic

Standardized versions of other West Slavic languages have greater use of prefix and preposition *z(e)*:

	'with me'	'with a sister'	'clench'
Czech	se mnou [sɛ mnow]	se sestrou [sɛ sɛstrow]	sevřít [sɛvr̝iːt]
Slovak	so mnou [so mnow]	so sestrou [zo sɛstrow]	zovrieť [zovrɟɛc]
Polish	ze mną [zɛ mnɔ̃w]	zε siostrą [zɛ ɕɔstrɔ̃w]	zewrzeć [zɛvʐɛtɕ]

Table 1. Cognates of Czech prefix and preposition *se* in West Slavic languages

The Slovak prescriptive codified norm (Kráľ, 1983/1996, pp. 117–8) distinguishes between pronouns and other nouns, capturing two intermediate stages of the shift.

Previous work

Balhar et al. (2016): Preposition and prefix [z(ɛ)] attested for standard *s(e)*, primarily in regions that also have pre-sonorant voicing

Preposition: *se sestrou* [zɛ sɛstrow] 'with a sister'

- Common throughout Moravia and Western Bohemia
- Most frequent with words beginning with [s] (possible voicing dissimilation)
- Prepositions *s(e)* and *z(e)* possibly merging in Moravia

Prefix: *sebrat* [zɛbrat] 'collect'

- Common in eastern regions only
- Prefixes *s(e)* and *z(e)* merging in Moravia due to similar meanings (cf. Šlosar, 1981, pp. 91–3)
- Possibly supported by voicing assimilation

Corpus study

- ORTOFON (Lukeš et al., 2021): ~3 million word corpus of spontaneous spoken Czech with phonetic transcriptions, lemmatization, demographic information, etc.
- Speakers grouped by the dialect region in which they have spent the most time
- Distribution of orthographic preposition and prefix *se*

Preposition

1089 tokens (606 from high-[zɛ] regions) of orthographic preposition *se* 'with' followed by a word with instrumental case (before sibilants and consonant clusters only)

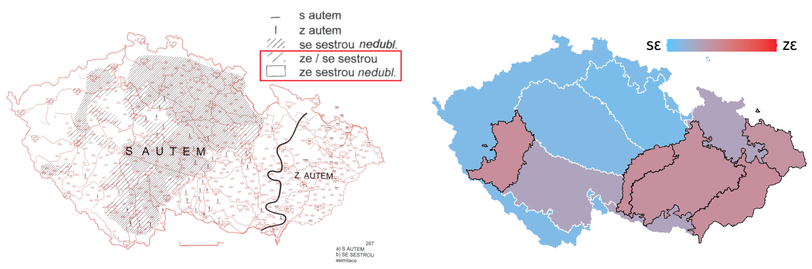


Figure 4. Realization of preposition *s(e)*: geographic comparison of Balhar et al. (2016) with corpus results (regions outlined in black used in analysis below)

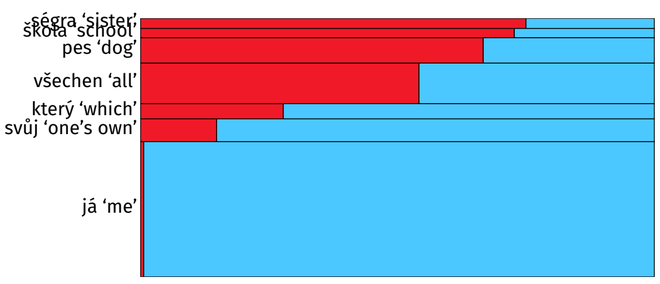


Figure 5. Rate of preposition [zɛ] vs. [sɛ] for frequent words (≥10 tokens)

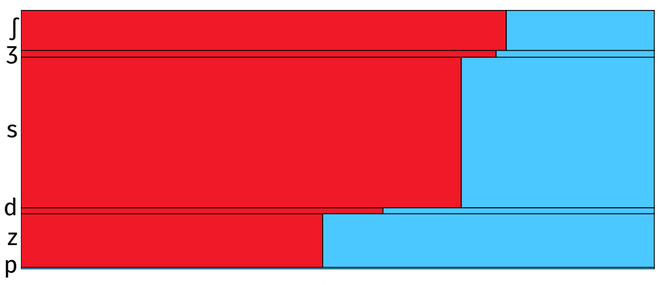


Figure 6. Rate of preposition [zɛ] vs. [sɛ] for less frequent words, by initial consonant

- **Lexicalization**: Very frequent pronouns have low rates of [zɛ], especially [sɛ mnow] 'with me', as in Slovak
- **Phonologization**: Non-pronouns have higher rates of [zɛ] that vary based on their initial consonant
- **Dissimilation**: As previously reported, [zɛ] is most common before voiceless sibilants and less common before voiced ones

Prefix

1157 tokens (623 from high-[zɛ] regions) of words (mostly verbs) with orthographic prefix *se* (typically before sibilants and consonant clusters)

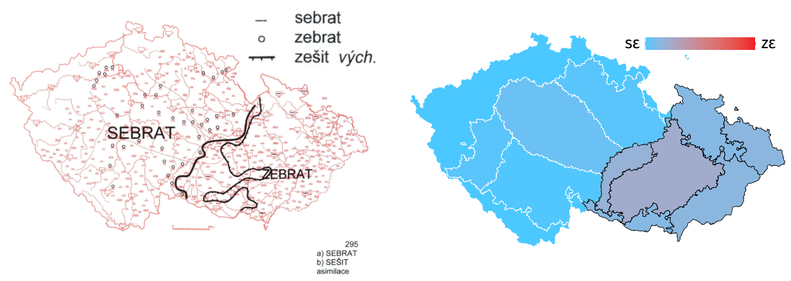


Figure 7. Realization of prefix *s(e)*: geographic comparison of Balhar et al. (2016) with corpus results (regions outlined in black used in analysis below)

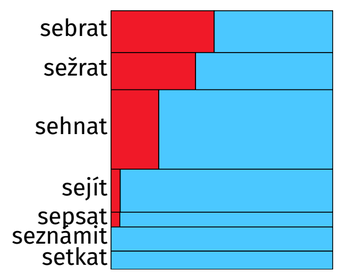


Figure 8. Rate of prefix [zɛ] vs. [sɛ] for frequent verbs (≥10 tokens)

	PERF	IMPF
'collect'	sebrat	[zb]írat
'eat up'	sežrat	[zʒ]írat
'get'	sehnat	[zɦ]ánět [sx]ánět
'meet up'	sejít	[sx]ázet
'write up'	sepsat	[s(ɛ)p]isovat
'acquaint'	seznámit	[sɛz]namovat
'meet'	setkat	[sɛt]kávat

Table 2. Imperfective counterparts of frequent verbs (≥10 tokens)

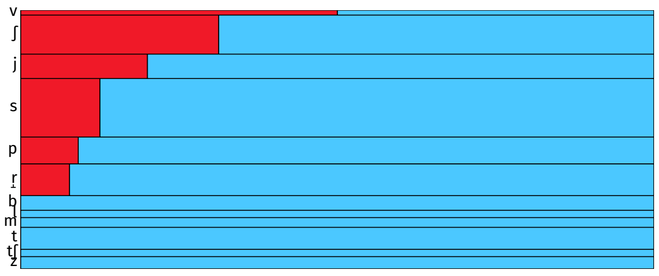


Figure 9. Rate of prefix [zɛ] vs. [sɛ] for less frequent words, by initial consonant

- **Lexicalization**: More frequent perfective verbs show moderate use of [zɛ] – if their imperfectives obligatorily have [z]
- **No phonologization**: Phonology creates conditions for reanalysis, but a direct effect of initial consonant is not so clear
- **Dissimilation?**: The phonological effect of dissimilation may play a small role in driving the lexicalized shift to [zɛ] before voiceless sibilants

Discussion

- Confirmation of claims in Balhar et al. (2016) that the prefix and preposition *s(e)* is merging with *z(e)*, especially in the east
- The prefix is shifting item-by-item
- The preposition is shifting in a more unified way, leaving scattered frequent exceptions
- Prefixes and prepositions remain a good way to test the application of voicing assimilation across phonological domains (e.g. Palková, 2013), adjacent sounds (e.g. Machač & Zíková, 2020; Ptak, 2020; Volín & Skarnitzl, 2006), and dialects (e.g. Šimáčková et al., 2012)
- But we must be careful to test the underlying form for each speaker and even lexical item – we cannot assume that it's faithful to orthography as is common practice (e.g. Hála, 1955, pp. 43–65; Palková, 1994, pp. 328–331; Hůrková, 1995, pp. 27–30)

References

Balhar, J., Bachmannová, J., Balátová, E., Čižmárová, L., Fic, K., Hladká, Z., Hlavsová, J., Ireinová, M., Kloferová, S., Komárková, Z., Konečná, H., Landsmann, R., Lamprechtová-Michálková, V., Majerčáková/Hlubinková, Z., Omelková, M., Pleskalová, J., Povolná, R., Přadková, P., Spinková, S., … Vojtová, J. (2016). *Český jazykový atlas* (2., elektronické, opravené a doplněné vydání, Vol. 5). Dialektologické oddělení Ústavu pro jazyk český AV ČR, v. v. i.

Cvrček, V., & Vondřička, P. (2011). SyD – korpusový průzkum variant. *FF UK*. syd.korpus.cz

Hála, B. (1955). *Výslovnost spisovné češtiny: Její zásady a pravidla: Vol. 1. Výslovnost slov českých*. Nakladatelství Československé akademie věd.

Hůrková, J. (1995). *Česká výslovnostní norma*. Scientia.

Kráľ, Á. (1996). *Pravidlá slovenskej výslovnosti* (3rd ed.). Slovenské pedagogické nakladatelstvo. (Original work published 1983)

Lukeš, D., Kopřivová, M., Laubeová, Z., Poukarová, P., Horký, V., Jelínek, T., Křivan, J., Waclawičová, M., Benešová, L., & Škarpová, M. (2021). ORTOFON v3: Korpus neformální mluvené češtiny s víceúrovňovým přepisem. *Ústav Českého národního korpusu FF UK*. www.korpus.cz

Machač, P., & Zíková, M. (2020). Distantní asimilace znělosti a signalizace hranice slov v češtině – případ pobočných slabik. *Naše řeč*, *103*(3), 189–211.

Palková, Z. (1994). *Fonetika a fonologie češtiny: S obecným úvodem do problematiky oboru*. Karolinum.

Palková, Z. (2013). Sandhiové modifikace v současné češtině. In A. Cychnerska (Ed.), *Sandhi w językach słowiańskich* (pp. 85–102). Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika.

Ptak, L. (2020). Sandhi na granicy międzywyrazowej: Badanie sondażowe w czeskim korpusie języka mówionego Ortofon. *Bohemistyka*, *20*(2), 153–165.

Šimáčková, Š., Podlipský, V. J., & Chládková, K. (2012). Czech spoken in Bohemia and Moravia. *Journal of the International Phonetic Association*, *42*(2), 225–232.

Šlosar, D. (1981). *Slovotvorný vývoj českého slovesa*. Univerzita J. E. Purkyně.

Volín, J., & Skarnitzl, R. (2006). K nevyhraněnosti vlastností české labiodentální frikativy. In P. Pořízka & V. P. Polách (Eds.), *Tzv. základní výzkum v lingvistice – desideratum, nebo realis? Sborník příspěvků z 5. mezinárodní konference Setkání mladých lingvistů, konané na Filozofické fakultě Univerzity Palackého v Olomouci ve dnech 17.–19. května 2004* (pp. 308–314). Univerzita Palackého v Olomouci.