

On the prevalence of different types of lexical semantic changes: Preliminary findings

Ecesu Ürker^{*1} and Gemma Boleda^{1,2}

^{*}Corresponding Author: ecesu.urker@upf.edu

¹Department of Translation and Language Sciences, Universitat Pompeu Fabra

²Catalan Institution for Research and Advanced Studies (ICREA)

Change is intrinsic to language—any living language continues to change over time (Campbell, 2013). This study focuses on lexical semantic change, examining the **cross-linguistic prevalence of different types of semantic change**. Prior theoretical work has converged on specific types of lexical semantic change, which we adopt here (Geeraerts, 2020; Traugott, 2017, 2000, see below), yet their relative prevalence across languages remains largely unexplored. As argued by Jensen and McGillivray (2017), advancing historical linguistics requires quantitative frameworks for phenomena traditionally studied theoretically; investigating these cross-linguistic patterns is therefore a crucial step toward understanding the mechanisms underlying semantic change. Prior computational research has largely focused on detecting semantic change without distinguishing among types (Cassotti et al., 2024), while annotation efforts have typically targeted specific phenomena, such as metaphor and word usage in context, rather than lexical change (Ptiček & Dobša, 2023). The closest studies related to ours include synchronic polysemy (Carling et al., 2023), whereas we focus on completed changes; the only other quantitative study is methodologically problematic (Asri et al., 2024).

To address this gap, we analyze the **DatSemShift** database (Anna Zalazniak, 2000-2026), which contains 10,172 semantic shift patterns across 2,208 languages.¹ We focus on 234 shifts labeled *Semantic Evolution* or *Microevolution*. We manually annotate each shift with part-of-speech (PoS) tags and classify along three dimensions of semantic change: **extension (broadening)**: *moon*, Earth’s satellite → any planet’s satellite; **narrowing**: *deor*, animal → deer), **emotive value (amelioration)**: *nescius*, ignorant → nice; **pejoration**: *seely*, blessed/innocent → *silly*), and **semantic frame (metaphor)**: *before*, in front → earlier; **metonymy**: *board*, table → governing body). We also provide each an-

¹As of February 2026.

notation with a confidence score and a brief rationale.²

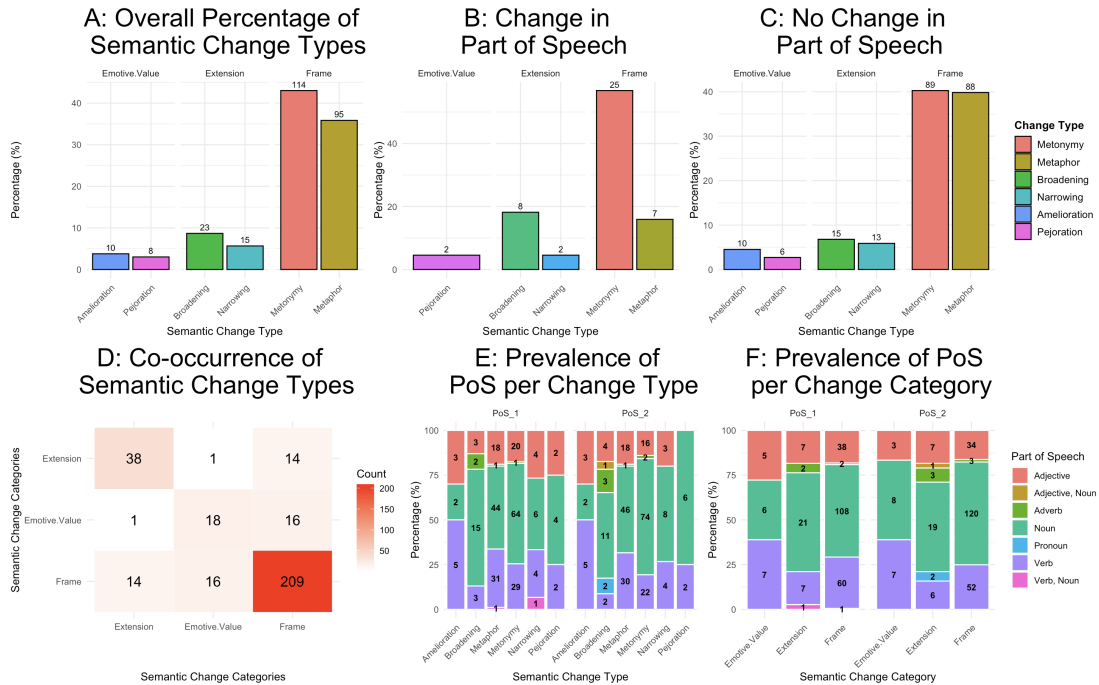


Figure 1. Overview of the study results. A) Overall frequency of semantic change types. B) Frequency of semantic change types in cases with PoS change. C) Frequency of semantic change types in cases without PoS change. D) Heatmap showing co-occurrence counts of semantic change categories. E) Percentage and counts of PoS tags per semantic change type. F) Percentage and counts of PoS tags per semantic change category.

Figure 1A shows that semantic frame changes (metaphor and metonymy) are the most common types of semantic change significantly (χ^2 test: $p < 0.05$), which may explain their prominence in the literature and provide quantitative support for their proposed role in word formation (Basilio, 2006). Emotive value changes were the least frequent, possibly accounting for their absence in previous quantitative studies (Carling et al., 2023). Figures 1B–C indicate that metaphor and metonymy occur at similar rates when PoS remains stable, whereas metaphor is significantly less frequent when PoS changes (e.g., verb $\rightarrow$ noun; Fisher’s exact test: $p < 0.05$). In contrast, metonymy does not differ significantly between these two conditions ($p > 0.05$). This asymmetry may reflect the fact that metaphor typically relies on functional or perceptual similarity, whereas metonymy can involve a broader range of relations. Figure 1D shows that different change types can co-occur, a pattern not previously reported. Figures 1E–F show that nouns generally

²Data, annotation guides, and code are available here: https://github.com/ecesuurker/DatSemShift_Semantic_Change_Types

dominate across categories, with the possible exception of emotive value; however, this difference is not significant (χ^2 test: $p > 0.05$), likely due to data sparsity. Previous studies suggested that metaphors primarily involve verbs (Cameron, 2003; Jamrozik et al., 2013), but we do not find evidence for this pattern; future work should investigate the source of this discrepancy.

Overall, this study provides the first quantitative overview of semantic change types and offers a foundation for future work on the mechanisms driving lexical semantic change.

References

- Anna Zalizniak, D. G. I. G. V. K. V. K. T. M. T. M. M. O. M. R. R. E. R. A. S. M. B.-F. O. D. M. B. M. V., Maria Bulakh. (2000-2026). *Database of semantic shifts*. Moscow: Institute of Linguistics, Russian Academy of Sciences.
- Asri, W. K., Rhamadanty, W. A. U., Burhamzah, M., & Alamsyah, A. (2024). Analyzing semantic shifts in english and german by exploring historical influences and societal dynamics. *Studies in English Language and Education*, 11(2), 1085–1100.
- Basilio, M. (2006). Metaphor and metonymy in word formation. *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada*, 22, 67–80.
- Cameron, L. (2003). *Metaphor in educational discourse*. A&C Black.
- Campbell, L. (2013). *Historical linguistics*. Edinburgh University Press.
- Carling, G., Cronhamn, S., Lundgren, O., Bogren Svensson, V., & Frid, J. (2023). The evolution of lexical semantics dynamics, directionality, and drift. *Frontiers in Communication*, 8, 1126249.
- Cassotti, P., De Pascale, S., & Tahmasebi, N. (2024). Using synchronic definitions and semantic relations to classify semantic change types. *arXiv preprint arXiv:2406.03452*.
- Geeraerts, D. (2020). Semantic change: “what the smurf?”. *The Wiley Blackwell Companion to Semantics*, 1–24.
- Jamrozik, A., Sagi, E., Goldwater, M., & Gentner, D. (2013). Relational words have high metaphoric potential. *Metaphor in NLP*, 1, 21–26. (Originally published in the Proceedings of the 1st Workshop on Metaphor in NLP (ACL 2013))
- Jenset, G. B., & McGillivray, B. (2017). Methodological challenges in historical linguistics. In G. B. Jenset & B. McGillivray (Eds.), *Quantitative historical linguistics: A corpus framework* (pp. 1–35). Oxford, United Kingdom: Oxford University Press.
- Ptiček, M., & Dobša, J. (2023). Methods of annotating and identifying metaphors in the field of natural language processing. *Future Internet*, 15(6), 201.
- Traugott, E. C. (2000). Semantic change: An overview. In L. Cheng & R. Sybesma (Eds.), *The first glot international state-of-the-article book* (pp. 385–406). Berlin, Boston: De Gruyter Mouton.