

The Politics of Code-Switching: Attitudes and Patterns in Boğaziçi Turkish

Furkan Akkurt^α, Nursena Teker^β, Helin Binici^β, Ahmet Demir^β, Konstantinos Sampanis^γ

Department of Computer Engineering, Boğaziçi University, Istanbul, Turkey^α

Department of Linguistics, Boğaziçi University, Istanbul, Turkey^β

Department of Byzantine and Modern Greek Studies, University of Vienna, Vienna, Austria^γ

^αCorresponding Author E-mail: furkanakkurt7242@icloud.com

ABSTRACT

The term *Boğaziçi Turkish* refers to the variety of Turkish influenced by English, commonly spoken by Boğaziçi University students and characterized by code-switching. This phenomenon is widely recognized in informal settings and is sometimes referred to as *Boğaziçi Tarzancası* (Boğaziçi Tarzan-speak). However, its systematic nature remains largely unexamined in the linguistic literature.

This study presents a two-fold contribution: (1) a comprehensive sociolinguistic analysis of attitudes toward Turkish-English code-switching among Boğaziçi University students, and (2) the initial release of a Universal Dependencies (UD) treebank for Turkish-English code-switched sentences. Using established sociolinguistic and computational methodologies, we provide a standardized resource for analyzing syntactic patterns in Turkish-English code-switching, facilitating further research in computational linguistics.

We conducted an online questionnaire ($n = 86$) to examine attitudes toward and patterns of Turkish-English code-switching. The survey collected demographic information, self-reported English proficiency, political orientation, and familiarity with “Boğaziçi Turkish.” Participants rated the acceptability of authentic code-switched sentences on a 6-point Likert scale and provided examples from their daily communications.

We employed non-parametric statistical methods suited for Likert scale data, including Kruskal-Wallis H-tests for categorical variables and Spearman correlations for continuous variables. The analysis included FDR correction for multiple comparisons, calculation of effect sizes, and 95% confidence intervals using Fisher’s z-transformation.

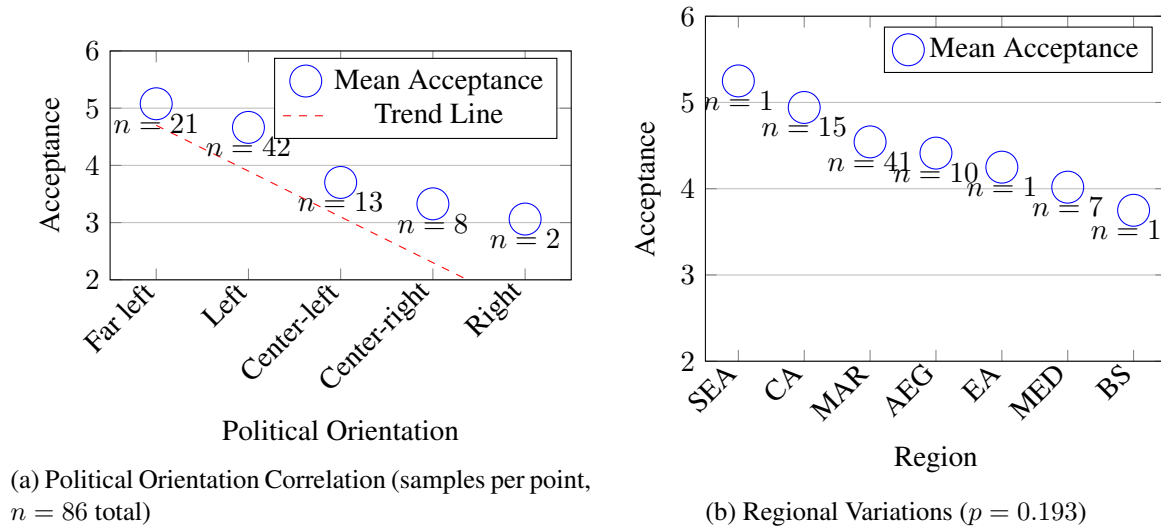


Figure 1: Factors Influencing Code-Switching Acceptance

Our statistical analysis revealed four key factors correlated with code-switching acceptance. Political orientation emerged as the strongest predictor, showing strong negative correlations on both left-right ($r = -0.480$, $p < 0.001$) and liberal-conservative ($r = -0.463$, $p < 0.001$) scales, with notably higher

acceptance rates among politically liberal/left-leaning students (Figure 1a). Mean acceptance declined steadily from far-left ($\mu = 5.077$, $n = 21$) to right ($\mu = 3.062$, $n = 2$) on the political spectrum. Similarly, concern about English influence on Turkish showed a significant negative correlation with code-switching acceptance ($r = -0.473$, $p < 0.001$).

Two demographic factors also showed notable correlations. Although regional variations were observed ($p = 0.193$), they were not statistically significant. Central Anatolia showed the highest acceptance rates ($\mu = 4.942$, $n = 15$), whereas the Black Sea region had the lowest ($\mu = 3.750$, $n = 11$), as visualized in Figure 1b. Gender differences were near significance ($p = 0.058$), with the ‘Other’ gender category showing the highest acceptance ($\mu = 5.444$, $n = 9$) compared to ‘Female’ ($\mu = 4.302$, $n = 50$) and ‘Male’ ($\mu = 4.417$, $n = 27$).

As the second outcome, we’ve annotated 10 representative sentences using a semi-automated pipeline in the Universal Dependencies annotation framework (Nivre et al., 2020). The process began with preliminary annotation using the language model Claude 3.5 Sonnet (Anthropic, 2024), followed by manual verification and correction by four annotators using ArboratorGrew (Guillaume, 2021). Our annotation scheme aligns with existing Turkish UD treebanks while incorporating adjustments for code-switching, particularly in head assignment within mixed-language constructions.

Qualitative analysis revealed distinctive code-switching patterns, including English verbs with Turkish auxiliaries, academic terminology, and pragmatic expressions. A notable pattern is the morphological integration of English verbs into Turkish syntax, exemplified by *drop-by-layacađım* “I will drop by,” where the English phrasal verb receives Turkish morphological markers. This is illustrated in Figure 2.

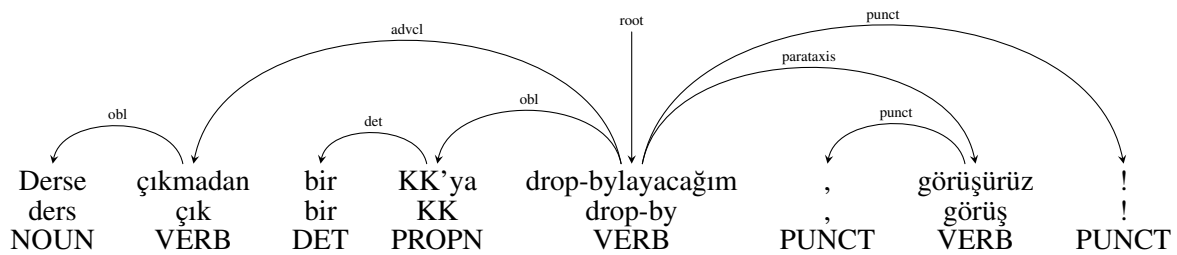


Figure 2: Dependency parse of a code-switched sentence in Boğaziçi Turkish

The Universal Dependencies analysis demonstrates three key syntactic patterns in Boğaziçi Turkish: preserving Turkish syntactic structure with English lexical insertions, morphological adaptation of English verbs, and code-switching at specific syntactic boundaries.

This study provides the first quantitative evidence of sociolinguistic factors influencing Boğaziçi Turkish usage, highlighting the strong link between code-switching attitudes and socio-political identities while laying the groundwork for computational linguistic research. The current annotation of 10 representative sentences will be significantly expanded, and the complete treebank resource will be publicly released in the next Universal Dependencies release.

Keywords: Sociolinguistics; Code-switching; Universal Dependencies; Corpus Linguistics.

References

- Anthropic (2024). Claude 3.5 Sonnet [Large language model]. <https://www.anthropic.com/claude>. Accessed: February 2025.
- Guillaume, B. (2021). ArboratorGrew: An online syntax annotation tool and dependency treebank building ecosystem. In *International Conference on Language Resources and Evaluation*. European Language Resources Association.

Nivre, J., de Marneffe, M.-C., Ginter, F., Hajič, J., Manning, C. D., Pyysalo, S., Schuster, S., Tyers, F., and Zeman, D. (2020). Universal Dependencies v2: An evergrowing multilingual treebank collection. In *Proceedings of the 12th Language Resources and Evaluation Conference*, pages 4034–4043. European Language Resources Association.