



The
Word Order in
Western Asia
Corpus

metadata for the data set

Turkic — Oghuz (Erzurum)

09 September 2021
(doc.v1.0)

Overview of the data set

data set	turk_oghuz_erzurum
contributors	Mahîr Can Doğan
published	09 September 2021
texts	3
tokens	636 analyzed (834 total)
notes	XLS/TSV contain entire source text

Metadata for individual texts

Text A

source	Gemalmaz (1995: 14–16) (WAV ✕ PDF ✓)
location	Erzurum, Turkey (39.9167° N, 41.2833° E)
recorded	c. 1995
text type	traditional narrative
speaker	01 — male, age 52
tokens	193 analyzed (216 total)

Text B

source	Gemalmaz (1995: 17–21) (WAV ✕ PDF ✓)
location	Taşpınar village, Erzurum Province, Turkey (40.1826° N, 40.7246° E)
recorded	c. 1995
text type	personal narrative
speaker	02 — male, age 60
tokens	288 analyzed (413 total)

Text C

source	Gemalmaz (1995: 161–163) (WAV ✕ PDF ✓)
location	Yeniköy village, Erzurum Province, Turkey (39.9774° N, 41.6678° E)
recorded	c. 1995
text type	traditional narrative
speaker	03 — female, age 33
tokens	155 analyzed (205 total)

Background to the data set

The texts in this data set are part of a three-volume documentation of the Turkish dialect spoken in Erzurum province of Turkey (Gemalmaz 1995), compiled as part of a nation-wide survey of dialectal Turkish conducted by the Turkish Language Society (Türk Dil Kurumu). The texts recorded were published with a relatively narrow transcription in order to reflect dialect features of phonetics and morphology as closely as possible. Sound files were not available to us. We have imported the transcriptions via OCR-recognition from scans of the book, though some diacritics were lost in the process. We also added an English translation. The texts are a mix of traditional narratives and oral history.

References

Gemalmaz, Efrasiyap. 1995. *Erzurum İli Ağızları: II cilt (metinler)*. Ankara: TDK.