



The
Word Order in
Western Asia
Corpus

metadata for the data set

Iranian — Balochi (Turkmen)

3 November 2022
(doc.v1.1)

Overview of the data set

data set	iran_balochi_turkmen
contributors	Geoffrey Haig
published	08 September 2021
texts	4
tokens	580 analyzed (628 total)
notes	XLS/TSV contain entire source text

Metadata for individual texts

Text A

source	Axenov (2006: 272–284) (WAV ✓ PDF ✓)
location	Mari region, Turkmenistan (36.5486° N, 62.5606° E)
recorded	c. 1995
text type	traditional narrative
speaker	01 — male, age >50
tokens	326 analyzed (370 total)

Text B

source	Axenov (2006: 284–287) (WAV ✓ PDF ✓)
location	Mari region, Turkmenistan (36.5486° N, 62.5606° E)
recorded	c. 1995
text type	traditional narrative
speaker	02 — male, age >50
tokens	88 analyzed (91 total)

Text C

source	Axenov (2006: 287–289) (WAV ✓ PDF ✓)
location	Mari region, Turkmenistan (36.5486° N, 62.5606° E)
recorded	c. 1995
text type	procedural description
speaker	03 — male, age unknown
tokens	68 analyzed (69 total)

Text D

source	Axenov (2006: 291–294) (WAV ✗ PDF ✓)
location	Mari region, Turkmenistan (36.5486° N, 62.5606° E)
recorded	c. 1995
text type	personal narrative
speaker	04 — male, age unknown
tokens	98 analyzed (98 total)

Background to the data set

The texts are taken from a documentation of the Balochi language of Turkmenistan (Axenov 2006). They were recorded, transcribed, and annotated by Serge Axenov, and comprise traditional narratives and other oral genres.

References

Axenov, Serge. 2006. *The Balochi Language of Turkmenistan*. Uppsala: Acta Universitatis Upsaliensis.