



The
Word Order in
Western Asia
Corpus

metadata for the data set

Semitic — Arabic (Christian, Ka'biye)

23 August 2024
(doc.v1.0)

Overview of the data set

data set	semi_arabic_christian+kabiye
contributors	Paul Noorlander
published	23 August 2024
texts	4
tokens	643 analyzed (788 total)
notes	XLS/TSV contain entire source text

Metadata for individual texts

Text A

source	Jastrow (2022: Text II) (WAV ✓ PDF ✓)
location	Diyarbakır, Turkey (37.8990° N, 40.2950° E)
recorded	unknown
text type	traditional narrative
speaker	01 — female, age unknown
tokens	160 analyzed (186 total)

Text B

source	Jastrow (2022: Text IX) (WAV ✓ PDF ✓)
location	Diyarbakır, Turkey (37.8990° N, 40.2950° E)
recorded	unknown
text type	traditional narrative
speaker	02 — male, age unknown
tokens	170 analyzed (207 total)

Text C

source	Jastrow (2022: Text XIII) (WAV ✓ PDF ✓)
location	Diyarbakır, Turkey (37.8990° N, 40.2950° E)
recorded	unknown
text type	traditional narrative
speaker	03 — female, age unknown
tokens	120 analyzed (173 total)

Text D

source	Jastrow (2022: Text XIV) (WAV ✓ PDF ✓)
location	Diyarbakır, Turkey (37.8990° N, 40.2950° E)
recorded	unknown
text type	traditional narrative
speaker	03 — female, age unknown
tokens	193 analyzed (222 total)

Background to the data set

The texts in this data set were taken from a monograph on the now extinct Ka'biye variety of Arabic (Jastrow 2022), whose speakers fell victim to ethnic cleansings during the 1915 Armenian and Assyrian genocide. The texts were recorded by Jastrow in 1968 with surviving speakers in Diyarbakır (A, B) and Beirut (C, D).

The texts were annotated for WOWA by Paul Noorlander, with additional data handling work performed by Iuliia Berg.

References

Jastrow, Otto. 2022. *Der arabische Dialekt der Christen von Ka'biye (Diyarbakir)*. Wiesbaden: Harrassowitz.