



The
Word Order in
Western Asia
Corpus

metadata for the data set

Turkic — Oghuz (Ankara)

24 August 2021
(doc v1.0)

Overview of the data set

data set	turk_oghuz_ankara
contributors	Kateryna lefremenko
published	18 July 2021
texts	28
tokens	587 analyzed (799 total)
notes	XLS/TSV contain entire source text

Metadata for individual texts

Text A

source	lefremenko (2021: TUBi01MT_isT) (WAV ✓ PDF ✓)
location	Ankara, Turkey (39.9300° N, 32.8500° E)
recorded	2018
text type	stimulus-based narrative
speaker	01 — male, age 28
tokens	17 analyzed (26 total)

Text B

source	lefremenko (2021: TUBi02FT_isT) (WAV ✓ PDF ✓)
location	Ankara, Turkey (39.9300° N, 32.8500° E)
recorded	2018
text type	stimulus-based narrative
speaker	02 — female, age 27
tokens	10 analyzed (11 total)

Text C

source	lefremenko (2021: TUBi03FT_isT) (WAV ✓ PDF ✓)
location	Ankara, Turkey (39.9300° N, 32.8500° E)
recorded	2018
text type	stimulus-based narrative
speaker	03 — female, age 26
tokens	12 analyzed (21 total)

Text D

source lefremenko (2021: TUBi04MT_isT) (WAV ✓ PDF ✓)
location Ankara, Turkey (39.9300° N, 32.8500° E)
recorded 2018
text type stimulus-based narrative
speaker 04 — male, age 27
tokens 7 analyzed (9 total)

Text E

source lefremenko (2021: TUBi05MT_isT) (WAV ✓ PDF ✓)
location Ankara, Turkey (39.9300° N, 32.8500° E)
recorded 2018
text type stimulus-based narrative
speaker 05 — male, age 27
tokens 22 analyzed (24 total)

Text F

source lefremenko (2021: TUBi06MT_isT) (WAV ✓ PDF ✓)
location Ankara, Turkey (39.9300° N, 32.8500° E)
recorded 2018
text type stimulus-based narrative
speaker 06 — male, age 26
tokens 14 analyzed (16 total)

Text G

source lefremenko (2021: TUBi07FT_isT) (WAV ✓ PDF ✓)
location Ankara, Turkey (39.9300° N, 32.8500° E)
recorded 2018
text type stimulus-based narrative
speaker 07 — male, age 40
tokens 26 analyzed (31 total)

Text H

source lefremenko (2021: TUBi08MT_isT) (WAV ✓ PDF ✓)
location Ankara, Turkey (39.9300° N, 32.8500° E)
recorded 2018
text type stimulus-based narrative
speaker 08 — female, age 25
tokens 14 analyzed (20 total)

Text I

source lefremenko (2021: TUBi09MT_isT) (WAV ✓ PDF ✓)
location Ankara, Turkey (39.9300° N, 32.8500° E)
recorded 2018
text type stimulus-based narrative
speaker 09 — male, age 27
tokens 13 analyzed (23 total)

Text J

source lefremenko (2021: TUBi10MT_isT) (WAV ✓ PDF ✓)
location Ankara, Turkey (39.9300° N, 32.8500° E)
recorded 2018
text type stimulus-based narrative
speaker 10 — male, age 25
tokens 9 analyzed (19 total)

Text K

source lefremenko (2021: TUBi11FT_isT) (WAV ✓ PDF ✓)
location Ankara, Turkey (39.9300° N, 32.8500° E)
recorded 2018
text type stimulus-based narrative
speaker 11 — female, age 32
tokens 42 analyzed (55 total)

Text L

source lefremenko (2021: TUBi13FT_isT) (WAV ✓ PDF ✓)
location Ankara, Turkey (39.9300° N, 32.8500° E)
recorded 2018
text type stimulus-based narrative
speaker 12 — female, age 33
tokens 27 analyzed (33 total)

Text M

source lefremenko (2021: TUBi14MT_isT) (WAV ✓ PDF ✓)
location Ankara, Turkey (39.9300° N, 32.8500° E)
recorded 2018
text type stimulus-based narrative
speaker 13 — male, age 26
tokens 7 analyzed (18 total)

Text N

source lefremenko (2021: TUBi17FT_isT) (WAV ✓ PDF ✓)
location Ankara, Turkey (39.9300° N, 32.8500° E)
recorded 2018
text type stimulus-based narrative
speaker 14 — female, age 31
tokens 9 analyzed (14 total)

Text O

source lefremenko (2021: TUBi18MT_isT) (WAV ✓ PDF ✓)
location Ankara, Turkey (39.9300° N, 32.8500° E)
recorded 2018
text type stimulus-based narrative
speaker 15 — male, age 24
tokens 47 analyzed (63 total)

Text P

source lefremenko (2021: TUBi20FT_isT) (WAV ✓ PDF ✓)
location Ankara, Turkey (39.9300° N, 32.8500° E)
recorded 2018
text type stimulus-based narrative
speaker 16 — female, age 25
tokens 15 analyzed (18 total)

Text Q

source lefremenko (2021: TUBi21MT_isT) (WAV ✓ PDF ✓)
location Ankara, Turkey (39.9300° N, 32.8500° E)
recorded 2018
text type stimulus-based narrative
speaker 17 — male, age 27
tokens 30 analyzed (36 total)

Text R

source lefremenko (2021: TUBi22MT_isT) (WAV ✓ PDF ✓)
location Ankara, Turkey (39.9300° N, 32.8500° E)
recorded 2018
text type stimulus-based narrative
speaker 18 — male, age 38
tokens 85 analyzed (110 total)

Text S

source lefremenko (2021: TUBi30FT_isT) (WAV ✓ PDF ✓)
location Ankara, Turkey (39.9300° N, 32.8500° E)
recorded 2018
text type stimulus-based narrative
speaker 19 — female, age 24
tokens 16 analyzed (23 total)

Text T

source lefremenko (2021: TUBi12MT_isT) (WAV ✓ PDF ✓)
location Ankara, Turkey (39.9300° N, 32.8500° E)
recorded 2018
text type stimulus-based narrative
speaker 20 — male, age 24
tokens 19 analyzed (29 total)

Text U

source lefremenko (2021: TUBi15MT_isT) (WAV ✓ PDF ✓)
location Ankara, Turkey (39.9300° N, 32.8500° E)
recorded 2018
text type stimulus-based narrative
speaker 21 — male, age 32
tokens 26 analyzed (40 total)

Text V

source lefremenko (2021: TUBi25MT_isT) (WAV ✓ PDF ✓)
location Ankara, Turkey (39.9300° N, 32.8500° E)
recorded 2018
text type stimulus-based narrative
speaker 22 — male, age 24
tokens 14 analyzed (19 total)

Text W

source lefremenko (2021: TUBi26MT_isT) (WAV ✓ PDF ✓)
location Ankara, Turkey (39.9300° N, 32.8500° E)
recorded 2018
text type stimulus-based narrative
speaker 23 — male, age 25
tokens 9 analyzed (11 total)

Text X

source lefremenko (2021: TUBi23MT_isT) (WAV ✓ PDF ✓)
location Ankara, Turkey (39.9300° N, 32.8500° E)
recorded 2018
text type stimulus-based narrative
speaker 24 — male, age 25
tokens 10 analyzed (13 total)

Text Y

source lefremenko (2021: TUBi24MT_isT) (WAV ✓ PDF ✓)
location Ankara, Turkey (39.9300° N, 32.8500° E)
recorded 2018
text type stimulus-based narrative
speaker 25 — male, age 24
tokens 19 analyzed (27 total)

Text Z

source lefremenko (2021: TUBi27MT_isT) (WAV ✓ PDF ✓)
location Ankara, Turkey (39.9300° N, 32.8500° E)
recorded 2018
text type stimulus-based narrative
speaker 26 — male, age 37
tokens 25 analyzed (38 total)

Text ZA

source lefremenko (2021: TUBi28MT_isT) (WAV ✓ PDF ✓)
location Ankara, Turkey (39.9300° N, 32.8500° E)
recorded 2018
text type stimulus-based narrative
speaker 27 — male, age 31
tokens 23 analyzed (27 total)

Text ZB

source	lefremenko (2021: TUBi29MT_isT) (WAV ✓ PDF ✓)
location	Ankara, Turkey (39.9300° N, 32.8500° E)
recorded	2018
text type	stimulus-based narrative
speaker	28 — male, age 29
tokens	20 analyzed (25 total)

Background to the data set

The data were gathered in Ankara in October 2018. The corpus consists of texts produced by 30 Kurmanji-Turkish speakers. The age of the participants varies between 25 and 37 years old.

The data was collected using the method “Language Situations” (Wiese 2020) within the framework of the project RUEG.¹ The method combines controlled elicitation with spontaneous data, and thus is suitable for systematic comparisons across contact-linguistic constellations as well as different languages. It captures quasi-naturalistic productions across different communicative situations, including informal versus formal, and written versus spoken settings. The elicitation comprised two sessions (one in Kurmanji as a family language and one in Turkish as the majority language of the society; the former forms the basis of the *iran_kurdish_northern+ankara* data set in WOWA), with at least three days between the sessions. Participants were shown a short video of a car accident and were asked to imagine themselves having witnessed it. After that they were asked to recount the incident in four different imagined situations: to a friend via a *WhatsApp* voice message (informal spoken), to a friend via a *WhatsApp* text message (informal written), to the police via a voice mail (formal spoken), to the police in a form of a written witness report (formal written). In order to exclude a possible effect of priming, we balanced the order of the situations and the order of the languages. Only informal spoken productions were included in this analysis.

References

- lefremenko, Kateryna. 2021. *RUEG-KMR-TUR 0.1 at ANNIS3*. (%7Bhttps://korpling.org/annis3/#_c=U1VFRy1LTVIitVFVSXzAuMQ%7D).
- Wiese, Heike. 2020. Language Situations: A method for capturing variation within speakers’ repertoires. In Asaki, Yoshiyuki (ed.), *Methods in Dialectology XVI*. Bern: Peter Lang.

multicast.aspra.uni-bamberg.de/resources/wowa/

¹ <https://www.linguistik.hu-berlin.de/de/institut/professuren/rueg/index.html>