



The
Word Order in
Western Asia
Corpus

metadata for the data set

Iranian — Bashkardi (Southern)

09 May 2022
(doc.v1.0)

Overview of the data set

data set	iran_bashkardi_southern
contributors	Agnes Korn, Ilya Gershevitch
published	9 May 2022
texts	5
tokens	234 analyzed (316 total)
notes	XLS/TSV contain entire source text

Metadata for individual texts

Text A

source	Gershevitch (1956: IG A3t1.1) (WAV ✓ PDF ✓)
location	Jakdân, Hormozgan Province, Iran (26.4431° N, 57.7517° E)
recorded	1956
text type	dialogue
speaker	01 — male, age unknown
tokens	57 analyzed (77 total)

Text B

source	Gershevitch (1956: IG A3t1.2) (WAV ✓ PDF ✓)
location	Jakdân, Hormozgan Province, Iran (26.4431° N, 57.7517° E)
recorded	1956
text type	personal narrative
speaker	02 — male, age unknown
tokens	41 analyzed (56 total)

Text C

source	Gershevitch (1956: IG A3t1.3) (WAV ✓ PDF ✓)
location	Jakdân, Hormozgan Province, Iran (26.4431° N, 57.7517° E)
recorded	1956
text type	dialogue
speaker	03 — male, age unknown
tokens	47 analyzed (74 total)

Text D

source	Gershevitch (1956: IG B2t35) (WAV ✓ PDF ✓)
location	Parmont, Hormozgan Province, Iran (26.3169° N, 58.0818° E)
recorded	1956
text type	personal narrative
speaker	04 — male, age unknown
tokens	28 analyzed (37 total)

Text E

source	Gershevitch (1956: IG Sp2Gour) (WAV ✗ PDF ✓)
location	Gurichi, Hormozgan Province, Iran (26.4541° N, 57.7531° E)
recorded	1956
text type	personal narrative
speaker	05 — male, age unknown
tokens	61 analyzed (72 total)

Background to the data set

All texts are based on material stemming from field work by Ilya Gershevitch in Iran in 1956 and belonging to the AIIT IG Bashkardi archive (cf. the description in Korn 2015). This archive contains the recordings made by Ilya Gershevitch, his field notes and material (mostly handwritten) of his subsequent work on the data.

For texts A, B, and C (IG At3.1–3) there is a handwritten transcription with some notes by Ilya Gershevitch in the archive (bundle “Tape I”). For texts A, B, C, and D there are sound files available in the archive. Texts A–C reoccur on another tape (B6t1), where the beginning of text A is missing. Not using Gershevitch’s notes, Bakhtiar Seddiqi Nejad in cooperation with Behrooz Barjasteh Delforoos produced preliminary transcriptions of the sound files and translations into Persian of A and D (as sound files). Based on these materials, Agnes Korn elaborated a considerably revised transcription (checked several times against the soundfiles) to reflect what she hears on the recordings, and a morphosyntactic analysis (using Flex software). For text E, there are two draft versions (one typed, one handwritten) by Ilya Gershevitch, containing his transcription and English translation. The transcription here is adjusted to Korn’s system, for instance attaching pronominal clitics to the preceding words by a hyphen and similar conventions. All translations in the present data set use the materials just mentioned, but differ from these to reflect Korn’s own understanding of the texts, their morphosyntax, and textual structure (Korn’s own token numberings are likewise included). Information on the accents noted by Ilya Gershevitch is not included in this project. The texts are all previously unpublished.¹

Texts A, B, and C (IG At1.1–3) seem to have been recorded in the same setting, perhaps even on the same occasion. Apparently Ilya Gershevitch is with the Kaxodâ of Jakdân, who is trying to motivate villagers to talk to Gershevitch, that is to the tape recorder (the tape recorder is probably referred to in A 41–44). The speaker of text D (IG B2t35) describes himself as being from Parmont, but he mentions having come to Bandar Abbâs to lodge a complaint. It is thus possible that the recording was made in Bandar Abbâs, which could explain some dialect mixture. There are no metadata available for text E (IG Sp2Gour) other than Ilya Gershevitch’s title “Gourečî specimen” on the drafts (see above).

¹ In Korn (2017: 97), text A is called ‘The Journey’, text B ‘Loading Camels’, and text C ‘Dialogue’.

The coordinates for the location of Parmont (text D) are somewhat pseudo-exact, as no place of a fitting name is to be found on current maps; they are those of the (on an officially looking map in Gershevitch's material) place Yamasûr, on the location of which there are houses to be seen on recent satellite images. There is a Parmân on the map, but at this location there are no houses on the satellite image. In Gershevitch (1959), there is a map that has Parmônt as a region; this region includes the coordinates given here.

The original data coding undertaken by Agnes Korn includes some additional information and colour-coding, which were not included in the version used for cross-corpus comparison. However, the original spreadsheet containing this original information is available on the WOWA website.² In Korn's numbering system, the "token" count does not include the data lines with extra-textual information (such as "interruption"), but does include everything which is text (also text in Persian, other speakers saying "yes" etc., all of which are then counted as non-analysed tokens).

References

- Gershevitch, Ilya. 1956. *AIIT IG Bashkardi archive*. Cambridge: Ancient India and Iran Trust (AIIT).
- Gershevitch, Ilya. 1959. Travels in Bashkardia. *Journal of the Royal Central Asian Society* 46(3-4). 213-225. (<https://doi.org/10.1080/03068375908731669>).
- Korn, Agnes. 2015. Voices from far away: Bashkardi material at the Ancient India and Iran Trust. *INDIRAN: Newsletter of the Ancient India and Iran Trust* 10. 4-5. (<https://hal.archives-ouvertes.fr/hal-01352800/document>).
- Korn, Agnes. 2017. Notes on the nominal system of Bashkardi. *Transactions of the Philological Society* 115. 79-97. (<https://doi.org/10.1111/1467-968X.12087>).

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

² https://multicast.aspra.uni-bamberg.de/resources/wowa/data/iranian/bashkardi_southern/southern-bashkardi__original-coding.xlsx