

The database of Hittite clauses involves clauses which contain *wh*-clitics such as relative pronouns, indefinite pronouns, subordinators *kuit* ‘as’ and *kuwapi* ‘when, where’. The database was elaborated and filled by E. Lyutikova and A. Sideltsev.

The database of Hittite *wh*-elements contains 2 257 clauses originating from the corpus of 516 texts; 1070 clauses feature relative pronouns. The total number of clauses in the corpus is 32 628. Every reasonably well-preserved clause of the corpus containing at least one *wh*-clitic was included into the database. Badly damaged clauses containing *wh*-clitics and those lacking crucial syntactic information were not included into the database. For the full composition of the Corpus see Appendix 1.

The relational database has been created in MS Access. The entry is a clause containing one or more of the following elements: a relative pronoun, an indefinite pronoun or adverb, subordinator *kuit* ‘while, as, because, that’, subordinator *kuwapi* ‘when, where’. The 47 fields accumulate the following information for each clause:

- Clause
- Translation
- Source
- Period
- Linear position of the *wh*-clitic in the clause
- Type of *wh*-clitics: relative, indefinite, *kuit* ‘as’, *kuwapi* ‘when, where’
- Number of *wh*-clitics: multiple relatives or multiple indefinites in one clause
- Additional information about *wh*-clitics: whether indefinites are marked by *-ki/ka*, whether subordinators split a phrase
- Clause properties: Wackernagel clitics, *-(y)a/-(m)a* conjunctions, clause connectives, *mān* ‘if’ or irrealis *man* or other delaying elements (see below), vocatives, negation markers, preverbs, additional subordinators in the same clause (the ‘freak’ type)
- If the *wh*-clitic is relative, the following information is provided: phrasal or proform. If proform, does it split another phrase? If phrasal, are the parts of the relative phrase adjacent or discontinuous? If they are adjacent, is the word order inverse (relative pronoun second) or direct (relative pronoun first)? If it is inverse, is the pronoun second after a word, a phrase or the status of the first element is unclear? If the relative phrase is discontinuous, what is fronted – a relative pronoun or a noun phrase? If the relative phrase is discontinuous, does the relative pronoun break another phrase?
- If the *wh*-clitic is indefinite, the following information is provided: phrasal or proform. If proform, does it split another phrase? If phrasal, are the parts of the indefinite phrase adjacent or discontinuous? If they are adjacent, is the word order inverse (indefinite pronoun second) or direct (indefinite pronoun first)? If it is inverse, is the pronoun second after a word, a phrase or is the status of the first element unclear? If the indefinite phrase is discontinuous, what is fronted – an indefinite pronoun or a noun phrase? If the indefinite phrase is discontinuous, does the indefinite pronoun break another phrase?
- Additional data: whether *wh*-clitic is clause-second, clause-second and preverbal, preverbal, postverbal. If the *wh*-clitic is clause-second, what is it second after – a word or a phrase? Is there a focus marker (*imma*) within the *wh*-clitic phrase, *parā* or *namma*, a serial verb?

The database design allows to easily access the complex combinatorial information about *wh*-clitics distribution. In this paper, we present our findings concerning the distribution of relative pronouns.

[Here we provide the list of clauses with relative phrases and some basic information concerning their marking in the database.](#)