A feature, as the term is used here, is a structural property of a language. A conscious choice was made to define features binary. That means that a given language is classified as either displaying a particular feature (having the value=1) or that the feature is absent (having the value=0) as far as the present data set is concerned. In some cases, for instance when conclusive data is missing, it is marked as indeterminate.

Our features are not meant to reflect all possible language properties. Instead they have been chosen, a) to represent areally relevant features, and b) to cover different parts of the language system (hence distributed across five macro-categories: Clause structure, Grammatical categories, Lexicon, Phonology and Word order). The choice of features to include is also constrained by our present, and relatively limited, data set.

Each feature presentation contains: 1. A table showing the number of sample languages displaying the feature, the number of languages in which it is absent, and the number of languages for which the feature value is indeterminate. 2. A map showing the geographical distribution of feature values across the region. 3. A short prose description of the feature as it occurs in the region, accompanied by one or more illustrative examples drawn from the data set.

Data references are given in the following way: Sample language-Data component-Language consultant:Item number. The reference HNO-Val-RH:061 should for instance be read as Hindko language data, provided by a consultant given the identifier RH, related to item number 061 in the Valency Questionnaire. Most of the data used in the feature descriptions belong to one of the seven components listed (along with their abbreviations) in Project design and data collection.

The abbreviations in the interlinear glossing are to a large extent standard abbreviations according to the Leipzig Glossing Rules. Additional abbreviations are listed below.

|  |  |
| --- | --- |
| AN | animate |
| CONJ | conjunction |
| EMPH | emphatic |
| EQ | equative copula |
| ESS | essive |
| EX | existential copula |
| HF | human feminine (gender) |
| HM | human masculine (gender) |
| INAN | inanimate |
| MED | medial (distance) |
| R | recipient (argument) |
| T | theme (argument) |
| V | verb |
| X | x (non-human gender) |
| Y | y (non-human gender) |