Lexical comparison - basis words as 'genetic markers'

The choice of the words used in the language comparison is not easy and is the result of many tests.

To be suited for lexical comparisons, the words should fulfil following conditions:

- They must have existed with the same meaning 5.000 to 10.000 years ago so that related languages had these words in their protolanguage (common ancestor).
- They must have kept a good semantic stability over the years – which is very rare, as words change their meaning overtime ("semantic shift").
- They must have not been subject to borrowing as such cases would lead to overestimate language proximity between not related languages.
- Their erosion must have been as limited as possible.

Try out a few lexical samples:

Avesta
Arabic
Submit

The 18 words used in this study have been chosen among words which are often in use in comparative linguistics studies. The words combination delivering the best results with this methodology when compared with results from other studies have been kept. For comparisons, only lexical morphemes are relevant - grammatical elements like nominative marks (eg. Latin, Hittite, Lithuanian, Gothic "s" desinences) or infinitive marks of verbs (Germanic "n", Slavic "t", Romance "r",...) are not taken into account and ignored in the **cognate scoring during lexical comparison**.

word	Comments
Eye	Stable word, with little exposure to semantic shift.
Ear	Pretty stable, semantically and also against erosion. Little probability to get borrowed from another language!
Nose	Very stable word, with little exposure to semantic shift and moderate daily use. Probably the best suited word for comparative linguistics!
Hand	As many other parts of the body, little exposure to borrowing and good semantic stability. However, in many languages, the meaning shift from "hand" to "arm" or the other way round.
Tongue	Very stable - similar to nose, although it is also being used for "language" in many languages and gives it an exposure to semantic shift or at least to confusion.
Tooth	Very stable - similar conditions as "nose". However, this word has been subject to semantic shift in parts of the Indo-European family - with a mix "Tooth"/"Tongue" ("-Z-B-" in Slavic/Indo-Iranan languages).
Death	As an abstract concept, the use of this word for comparing remote languages is somewhat hazardous. However, it is the one best linking the Indo-Europea and Semitic language families (Arabic الموت (mut)/ Hebrew מוות (mavet) -> French "Mort" / Slavic "Mertv"... In some languages, the root of the verb "to die" has been taken instead of the substantive "death" when it was not available ("to die" is an element of the Swadesh list "death" is not)
Water	Very interesting word, although it is in intensive use and as such subject to more erosion. Moreover, semantic shift exposure is higher than for body parts. Water is the word best linking the Indo-European and Finno-Ugric language families (Finnish "Vesi" / Hungarian "Vez" -> German "Wasser" / Slavic "Voda")
Sun	This word has a big exposure to semantic shift but delivers good results in comparative linguistics. Probably less suited for remote language relationships
Wind	As all nature related words, should have existed in early languages.
Night	Very classical example in Indo-European studies...
Two	Little exposure to semantic shift but intensive use in daily life ("erosion")
Three	Little exposure to semantic shift but intensive use in daily life ("erosion"). Sometimes exposure to borrowing like in Kabylian (see Kabylian to Arabic comparison)
Four	Little exposure to semantic shift but intensive use in daily life ("erosion"). Exposure to borrowing similar to "three".