

LINGUIST Codes for Ancient and Constructed Languages

Ethnologue Codes

- Consistently apply an operational definition of language so that all entities for which an identifier is assigned are of a comparable nature
- Encompass all of the languages of the world,
- Clearly document the speech variety that each identifier denotes
- Maintain and update the system on an on-going basis
- Make the system freely and readily accessible to the public over the Internet

For every language description:

- The countries the language is spoken in
- The alternate names that refer to the language
- The number of speakers of the language
- The classification of the language

Mutual Unintelligibility

- Varieties of language are only assigned a code if they are mutually unintelligible with varieties of any language to which a code has already been assigned.

Current Use

- The Ethnologue system is intended to encompass only those languages of the world in current use. Thus the Ge'ez (Ethnologue code GEE) and Sanskrit (Ethnologue code SKT) languages both appear in Ethnologue
- Most ancient languages are thus absent

Shortcomings in Ethnologue

- Every language in Ethnologue is documented to a greater or lesser degree. But we usually do not have a clear idea of the evidence upon which it was decided to assign the language a unique code. Nor does the system allow for conflicting language classifications. For example, there is disagreement amongst scholars as to the classification of Low German dialects. This is not indicated in Ethnologue.

Criteria for Ancient and Constructed languages

- Conform as closely as is reasonable to the standards set by Ethnologue

But...

- The criterion of mutual intelligibility has to be abandoned

e.g. Anglo-Norman, which was an aberrant dialect of Old French. However, it evolved independently, and has a literature distinct from that of Old French. This scholars treat separately. Thus it must be assigned a distinct code so that work on it can be discriminated from work on Old French.

Mutual intelligibility breaks down in another way

- Ancient languages often have a diachronic dimension that can usually be ignored with modern languages

e.g. Old Latin gave rise to Classical Latin, which in turn gave rise to Late Latin, which in turn gave rise to Vulgar Latin or Proto-Romance...

- It is likely that no two adjacent stages of this complex process would have been mutually incomprehensible, had there been any speakers who could speak the two versions. How many codes do we assign here on the basis of mutual intelligibility?

Undeciphered Scripts

- Ancient languages in scripts which have as yet not been deciphered, e.g. Minoan

Conclusion

- Codes should be assigned to ancient languages which are treated distinctly by the scholarly community.
- The standard of mutual intelligibility should apply as far as possible. All apparently mutually intelligible ancient languages spoken at approximately the same period should be assigned one code, unless this conflicts with scholarly usage.
- In cases where the level of mutual intelligibility cannot be clearly ascertained, separate codes should be assigned.
- Codes should be assigned to undeciphered scripts, and to uninterpretable ancient languages in known scripts.

Conclusion (cont)...

- The system should be as complete as possible. Ancient languages should not be excluded simply because they are obscure.
- All alternate names of ancient should be listed, even those which are deprecated by scholars
- To integrate with Ethnologue codes, the primary geographic categorization of ancient languages should be by the modern countries in which they once existed
- All codes should have provenance information
- Committees of specialists will provide the provenance information

Constructed Languages

- Constructed languages cannot be treated by the criterion of mutual intelligibility, since they are almost never actually spoken, and are as much cultural objects as linguistic. In some cases there exist variants of originally identical constructed languages which have begun evolving independently. Esperanto (Ethnologue code ESP) and Ido (LINGUIST code CIDO) are instances of this phenomenon. These should be assigned distinct codes.
- No attempt should be made to assign constructed languages to geographical regions, since they do not exist in the real world.

The Canary Agreement

- All languages which require codes and which became extinct before 1950 should become the responsibility of LINGUIST. All languages after 1950 will be in the purview of Ethnologue.
- The two code sets will be unified into one three-letter code-set.