The Herodotos Project: Determining the Rate of Language Loss in the Past

Julia Papke (et al., see below)
The Ohio State University

ABSTRACT

A common claim nowadays is that the rate of language loss is greater now than ever before. Without minimizing the real urgency such statements are meant to convey, one must realize that they are made without a clear empirical basis, in the form of a calculation of the rate of language loss in pre-modern times against which to measure the current rate. One estimate, by evolutionary biologist Mark Pagel, counts as many as 400,000 languages spoken since the beginning of human language (c. 75,000 B.P.); when put up against the current number of some 7,000 languages, it means that there has been a huge amount of language loss over many millennia.

Still, some empirical data on this question is certainly a desideratum. To that end, we have been establishing a point of comparison using information from historical sources that give a basis for saying approximately how many languages have died within pre-modern recorded history. Specifically, using Herodotos’s Histories, we are determining how many languages are known from the past that have not survived in any form; we then use that establishable standard from the past to calculate the rate of language loss was in ancient times, so that the modern rate of loss can be compared meaningfully.

In this paper, we report on this project (the “Herodotos Project”), which we see as a small step towards answering what some might consider a politically incorrect question, namely: is language loss just part of the natural “evolution” of languages or are globalization and other modern forces really speeding up the rate of loss? We discuss as well the project’s further implications, e.g. determining rate of “replacement” of languages from dialect split and creolization, and tying the results of historical language loss to other key events such as conquests and colonialization.

PRESENTATION HANDOUT:

1. Introducing the project:

   • This presentation describes an ongoing project, and will include a sampling of results
   • The project is based at Ohio State with connections to other institutions
   • Team members:
     - Ohio State: Christopher Brown (Dept. of Greek & Latin)
Primary Goal: **work toward an understanding of the rate of language loss in ancient times by surveying Herodotos and his mention of different peoples**

Secondary Goal: **creation of a resource for the study of language groups in the ancient world (including mapping their location over time (via LL-MAP project through LinguistList))**

2. Rationale:

A common claim nowadays is that the rate of language loss is greater now than ever before. There is a real urgency that such statements are meant to convey.

But, one must realize that they are made without a clear empirical basis, in the form of a calculation of the rate of language loss in **pre-modern** times against which to measure the current rate.

That is, in order to say that a rate is increasing (or decreasing) there must be a point of comparison.

Other relevant numericals: one estimate, by evolutionary biologist Mark Pagel (1995), counts as many as 400,000 languages spoken since the beginning of human language (c. 75,000 B.P.); when put up against the current number of some 7,000 languages, it means that there has been a huge amount of language loss over many millennia.

In this paper, we report on this project (the “Herodotos Project”), **still in its early stages**, which we see as a small step towards answering what some might consider a politically incorrect question, namely: is language loss just part of the natural “evolution” of languages or are globalization and other modern forces really speeding up the rate of loss? We discuss as well the project’s further implications, e.g. determining rate of “replacement” of languages from dialect split and creolization, and tying the results of historical language loss to other key events such as conquests and colonialization.

3. Methods:

• Starting with Herodotos’s Histories, we are creating a database of possible
ethnonyms. This includes as much information about the customs, appearance, religion and (particularly) language as is available in the text. It also includes any geographical information about a particular group, and the citations in Herodotos that provide useful information.

- This database includes other appearances of the ethnonyms in ancient sources (e.g. Stephanus of Byzantium), literary and epigraphical, and discussion in modern scholarship (e.g. How & Wells, Asheri.)

4. Caveats about working with Herodotos:

- Is Herodotos reliable as a source of information in general and in regard to language? (cf. Munson 2005)
- Language observations in the Histories are scattered and impressionistic.
  - Most ethnonyms have little or no mention of language at all
- Some of the groups he mentions are clearly unreal/mythological, e.g. the one-eyed people (such as the Ἀριμασποί/Arimaspians)
- Even if Herodotos is useable, there other issues with the data:
  - Questions about the difference between language and dialect, in general and in Herodotos’s conception (cf. Morpurgo-Davies 1987, Harrison 1998)
  - The challenge of determining what became of the group -- both how long they survived as a people, and how long they maintained a separate language.
  - determining rate of “replacement” of languages from dialect split and creolization, and tying the results of historical language loss to other key events such as conquests and colonialization.

5. Why use Herodotos?

- Unlike archaeological artifacts, what we get from Herodotos are actual language observations, with some degree of sophistication
  - his dialect observations (pace caveats above in (4))
  - his comments about Gelonians, and contact language
- Even if the data are imperfect, they are still the data that we have at hand.
  - Furthermore, a detailed compilation of information on the languages of the peoples in Herodotos is useful for many reasons, going beyond listing in Harrison 1988 (appendix 1)

6. Sample results:
There are, as of now, 128 ethnonyms in the database for books 1-4

- Of these,
  - There are some where he says something about culture/peoplehood
  - There are some where he says something about language

- Language death—Pelasgian, the language of a pre-Greek people who transmitted the names of the gods from the Egyptians to Greece—while conceding autochthony to the Athenians, Herodotus makes them descendants of the non-Greek speaking Pelasgians who were Hellenized

- Contact language—the Gelonian language resulting from contact of the originally Greek-speaking Gelonians with their neighbors the Budini; farmers apparently living on the Don (Asheri 595) they maintained Greek religion and customs while altering their language.

- Language experts—the Ichthyophagi, a tribe of Elephantine on the Red Sea, served as interpreters for Cambyses to the Ethiopians (3.19)

- Language as a marker of uniqueness—the androphagoi are unique (4.18, 4.106)—living to the north of the Scythsians and dressing like Scythsians but ἀγριώτατα πάντων ἀνθρώπων ἔχουσι ἡθέα, οὔτε δίκην νομίζοντες οὔτε νόμῳ οὐδενὶ χρεώμενοι

- irrelevance of language: the Caunian language is similar to that of the Carians, but the Caunians differ from them and all mankind in customs & religion (1.172); the Carians consider selves related to the Lydians and Mysians and admit them to temple of Zeus at Mylasa, despite their difference in language, while excluding speakers of their own language of a different ethnos (1.171) (add Munson ref.)

- Scythian languages—where Greeks like Persians referred to steppe-dwellers by a single word (Scythsians, Saka), Herodotus uses the word Scythian in a stricter sense (Asheri 657)—the Sauromatae are a nomadic culture notable for the high status of women, living east of the Don (Asheri 658 f), Scythsians who mixed with Amazons and thus speak Scythian ungrammatically (4.57)

7. Ultimate goal:

Use estimates about language occurrence in ancient times along with occurrence—or not—of (ostensibly) same languages at a later time, c. 500 years later, to see what percentage of existing languages at time X in the past were still in existence at time X + 500. That will give an estimate of the rate of language loss in ancient times against which the modern (undoubtedly correct) claims about a higher rate of language loss now than ever before can be calibrated.