Comparative Ethnobiology in Mesoamerica:
A Digital Portal for Collaborative Research and Public Dissemination

1. Introduction: Comparative study of the nomenclature and classification of biological species (flora and fauna) has been an important tool for studying the cultural history of language groups. Two major research topics have emerged. The first, exemplified by scholars such as Catherine Fowler (1972a, 1972b, 1983), Paul Friedrich (1970), Frank Siebert (1967), and K. W. Whistler (1977) (and most recently in the debate over the Aztec homeland [see References, section 2]), have reconstructed the lexicosemantics of proto-language terms for biotaxa and taken the reconstructed meanings as reflecting ancestral homeland ecosystems. In the depth of his study of proto–Indo-European, Friedrich broke a new ground in rigor by linking protosemantic reconstruction of flora to prehistoric ecosystems. Siebert, in turn, was the first to apply this methodology to American languages. Fowler continued this effort to use lexical evidence to document ecological clues to homelands but at the same time noted the limitations of this approach "given the *quality* and *quantity* of data presently available" (Fowler 1983:224; emphasis added).

A second direction of research regarding the lexicosemantics of biological nomenclature relates to contact phenomena and what has been called linguistic stratigraphy: "the systematic investigation of the layering of grammatical and lexical material in a language or dialect which reflects its historical development and past contacts between its speakers and bearers of other linguistic and cultural traditions" (Andersen 2003b:1). Within this area of research a small subset of studies have either focused exclusively on biological nomenclature (Bowern 2007; Bowern and McConvell 2011; Bostoen 2007; Meroz 2013) or relied heavily on terms from this semantic domain (Dakin 2003). The strategy of these studies differs, as Yoram Meroz notes, from most comparative lexical surveys in that to elucidate genetic relationships among languages a set of basic words most resistant to change is preferred (see Haspelmath and Tadmor, 2009, particularly chapters 1–3). The nomenclature of flora and fauna, however, is probably more sensitive than basic vocabulary to change through contact and thus is a particularly propitious semantic domain in which to study migration and contact.[1]

The loans that are relevant to such stratigraphic studies, in turn, may be either of form (i.e., loan words) or meaning (calques or loan translations). The former is common, for example, among Nahuat speakers of the Sierra Nororiental de Puebla who have borrowed many, some even quite basic, terms from Totonac (*xopepe* 'cockroach', *āltsimit* 'wasp', and *chokoy* 'puss caterpillar' [Megalopygidae]), probably reflecting Nahuat migration into the area. The second type of loans are calques, loans in which meaning is translated from one language to another but the term itself is not borrowed. Calques have been used to support the definition of Mesoamerica as a cultural area, though these loan translations are only one of several features, many morphosyntactic, that are regionally shared (Campbell, Kaufman, and Smith-Stark, 1986; Smith-Stark 1982, 1994; about 18 percent of the calques these authors reference denote to flora or fauna, such as 'mother of the leaf-cutter ant' for 'coral snake'). Amith in his ethnobiological research has discovered quite a few more (e.g. camel spiders and whip scorpions are both called "shame [animal]" among the Aztecs, modern Nahuatl speakers from central Guerrero, and the Coastal Mixtecs). The viability of using loan words and calques of biological nomenclature for studying historical contact is, however, also hindered by the same poor quality and quantity of data noted by Fowler.

The present project seeks to remedy this paucity of data for New World biosemantics through a web portal that will bring together fieldworkers in linguistics and anthropology, Western experts in the taxonomy of Neotropical flora and fauna, and Indigenous natural historians interested in documenting their local traditional ecological knowledge. Over the past two decades scores of projects have dedicated immense efforts to documenting endangered languages, often focusing on the Boasian core of "deliverables": corpus, lexicon, grammar. In the process, a significant amount of data on biosemantic lexicons is being gathered, but with no convenient way to share, discuss, and disseminate this information. Often the information gathered is partial (nomenclature of biotaxa associated with narrative descriptions of morphology, habitat, and use; photos of culturally salient plants or animals named in the local language) and incomplete (not part of a systematic attempt at ethnobiological research). The many letters of support from fieldworkers in

---

[1] Balée and Moore (1991) and Berlin et al. (1969, 1973) have looked at the factors that shape the rate of retention and loss in closely related languages among different types of ethnobiological nomenclanture.

Mesoamerican languages attests to the enthusiastic interest that this project has generated. At the same time, personal experience combined with communication from other researchers (see particularly the letters of Cruz and Swanton) suggest that native speakers have a strong interest in documenting their cultural heritage within which traditional ecological knowledge is prominent. If this proposal is successful every effort (personal contacts of the collaborators on this project, listserves, social media) will be made to disseminate knowledge of the website among Indigenous academics and communities.

2. Enhancing the humanities through innovation: This project is innovative in content, structure, and impact. It will create a substantive resource for research in cultural history (origins of and contacts among Mesoamerican Indigenous populations) and in comparative ethnobiology (nomenclature, classification, and use of flora and fauna), and will contribute to the documentation of species diversity in botany and entomology. The information that will constitute the foundation of this resource will be unique and important, unavailable from any other source. It will forge an interdisciplinary community of collaborators and it will tap into the potential of Indigenous communities, and researchers who work in these communities, as well as published and unpublished materials on traditional ecological knowledge.

The digital structure of this resource will be based entirely on open source codes and technologies and informed consent procedures will be documented for all relevant content (e.g., digital recordings of native speaker experts; see discussion of Mukurtu below). The information architecture will combine robust basics (Drupal, MySQL, GPS mapping) with innovative processes for linking materials including faceted search interfaces, highly linked metadata interfaces, and relevant/similar content discovery. The site will also support community based commenting, content rating, and curation interfaces. Any newly developed and improved module/library code will be contributed back to the open source communities. Although an early instantiation of a fully developed portal, this pilot project will explore innovative means to make the website highly portable and interactive while leveraging existing material available online.

3. Environmental scan: To the degree that the present project seeks to involve native speakers and their communities in the documentation of nomenclature and classification of local fauna, it is superficially related to many citizen science projects that target biodiversity (e.g, National Geographic Great Nature Project and, in Mexico, NaturaLista, hosted by the Comisión Nacional para el Conocimiento y Uso de la Biodiversidad.). The mapping software that they use, in which data points are displayed in their geographic location, and the possibilities for participant comment and exchange will certainly be adapted in the present project. To the degree that the present project seeks to establish interdisciplinary communication among social scientists and biologists, however, the closest similarity might well lie with the many academic listserves and the discussions that they promote, though these exchanges generally lack value for added data. These lists, moreover, are invariably poorly presented, bereft of any but the most basic graphical interface. Thus in terms of topical focus (ethnobiology), targeted communities (social scientists, biologists, native speakers), objectives (data for research on cultural history), and presentation, this project creates a unique environment for collaborative research in the humanities.

Database structures and graphic interfaces both for uploading and searching through ecological knowledge are woefully deficient. The challenge this project faces is to simplify contributions; control for quality; and, most importantly for the humanities, facilitate discovery and organization of results to stimulate studies in the humanities (cultural history, linguistics, anthropology) as well as the natural sciences. To facilitate contributions and access by Indigenous people in Mexico, all user interfaces will offer presentation in Spanish as well as English. Finally, those responsible for this project will look at open source Mukurtu, a content management system that was built in collaboration with Indigenous communities to manage and share their digital heritage. After evaluating this software those aspects relevant to this project's goals, such as their traditional knowledge protocols for Indigenous creators and custodians, will be utilized.

4. History and duration of the project: Since the year 2000 Amith has conducted grant supported (NSF, NEH, Ford Foundation, Endangered Language Documentation Programme [ELDP]) research on three Indigenous languages of Mexico: Guerrero Nahuatl, Sierra Nororiental de Puebla Nahuat, and Pacific Coast of Guerrero Mixtec. Increasingly he has focused on documenting traditional ecological knowledge (nomenclature, classification, and economic and symbolic use) of local flora and fauna. He has done this with voucher specimens (collecting over 2,500 plants and 3,000 arthropods), photographs (over 5,000 images, see example photos in the appendix), and extensive digital recordings of native natural historians discussing local flora and fauna (over 1,000 recordings, all transcribed, in the three languages). To obtain scientific

determinations to species, Amith has forged a network of over 200 colleagues, expert taxonomists in plants and arthropods (see partial list in appendix, pp. 4–6), who have generously determined to scientific species the vouchers Amith has collected. Recently Amith has obtained significant support ($800,000) from NSF, NEH, and ELDP to use molecular data (DNA barcoding) to facilitate identification to species of plants collected in ethnobotanical research. That project will generate a significant amount of new material on nomenclature, classification, and use in seven Sierra Nororiental de Puebla communities (5 Nahuat and 2 Totonac).

Amith has reached out to other researchers on endangered languages and cultures in Mexico to gauge their interest in incorporating documentation on the nomenclature, classification, and use of flora and fauna into their efforts. The response has been extremely enthusiastic (as exemplified by letters of support from nine linguist and anthropologist colleagues). In sum, this present project offers a unique opportunity to create a virtual web of interdisciplinary and interethnic collaboration that will generate significant new material of prime importance to the cultural history of Indigenous Mesoamerica.

5. Staff: Amith (PI) heads a project team of highly qualified individuals combining skills in anthropology and linguistics (Amith), digital technology (Remy [Gettysburg College] and Ogilvie and Raley [Civic Actions], and botany (Lott and Steinmann). Collaborators working in Indigenous communities will Beta test functionality on uploading and searching, providing feedback as the website is developed.

6. Work plan: Phase One (May–Aug. 2015) Amith will work with Indigenous partners in Yoloxóchitl, Guerrero (Mixtec) and Cuetzalan, Puebla (Nahuat) who are collaborating on active research grants to develop a varied set of contemporary materials for inclusion in the prototype website. William Merrill's extensive database of over 8,000 terms for plant and animal nomenclature in Uto-Aztecan (see letter of commitment) will be converted to a form amenable to uploading to a MySQL database. Consultation between Amith, Remy, and Raley will determine the best database structure and tagging system for this material. **Phase Two (Sept.–Dec. 2015)**: Amith, Remy and Civic Actions will develop a graphical user interface for uploading material to the database. The interface will be designed to accept input from: (1) Contemporary data in multiple forms and types from fieldwork; (2) Contemporary published and unpublished information (e.g Merrill's database); (3) Information from historical sources (for the rich number of sources on Aztec flora and fauna, see References, section 3). Beta testing will be carried out by Amith's indigenous collaborators and the interface will be adjusted as needed. **Phase Three (Jan.–March 2015)**: Search engine capabilities will be developed to facilitate the generation of result sets that are beneficial to research agendas that potential users might have. A display template will be developed to display search results. **Phase Four (April–June 2015)**: Community based commenting, content rating, and curation interface functionality will be implemented. Playback capabilities will be created for digital recordings in Indigenous languages with presentation of transcriptions and translations; if transcriptions are time-coded playback will be aligned with text. **Phase Five (July–Sept. 2015)**: Collaborators (see letters of commitment) will be asked to Beta test upload, search, and display capacity of the Website. **Phase Six (October 2015)**: A bilingual manual for users will be written in English/Spanish and posted on the Website along with a metanarrative of the project history and goals. A register system will be developed to control for access by the general public. Finally, all substantive data will be exported from MySQL to XML to permit permanent archiving at the Archive of Indigenous Languages of Latin America, University of Texas, where Amith has already archived approximately 500 gigabytes of material although the principle results of this project will be txt files (Amith's photos, recordings, and transcriptions are already archived at AILLA).

7. Final product and dissemination: The final product will be a robust and stable open source and technology website developed and hosted at an academic institution or herbarium. Access to search and display functionality will be open to the general public although uploading permission will be granted on a case-by-case basis while any patches to the website are needed. Possible future enhancements include: (1) active hyperlinks to on-line field guides and herbaria (e.g., Tropicos at the Missouri Botanical Gardens, Animal Diversity Web); (2) database of sounds (e.g., birds, cicadas) filtered for regional specification; (3) adaptation of download capabilities to hand-held devices; (4) TEI encoding of textual material when appropriate.

The completion of the pilot project, a low-cost effort with potential for high-impact results, will facilitate efforts to secure funding both in the United States (NSF, NEH), Mexico (CONABIO), Canada, and Europe, both in government and private foundations. All potential sources of funding will be pursued.

References

1. General bibliography of works cited

Andersen, Henning. 2003. "Introduction," in Henning Andersen, ed., *Language Contacts in Prehistory: Studies in Stratigraphy*. Philadelphia: John Benjamins, pp. 1–10.

Andersen, Henning, ed. 2003. Language Contacts in Prehistory: Studies in Stratigraphy (Papers from the Workshop on Linguistic Stratigraphy and Prehistory at the Fifteenth International Conference on Historical Linguistics, Melbourne, 17 August 2001. Current Issues in Linguistic Theory 239. Philadelphia, Penn.: John Benjamins.

Balée, William, and Denny Moore. 1991. "Similarity and variation in plant names in five Tupi-Guarani languages (Eastern Amazonia)." *Biological Sciences* 35(4):209–62.

Berlin, Brent, Dennis E. Breedlove, Robert M. Laughlin, and Peter H. Raven. 1973. "Cultural significance and lexical retention in Tzeltal-Tzotzil ethnobotany. In Munro S. Edmonson, ed., *Meaning in Mayan Languages*. The Hague: Mouton, pp. 143–64.

————. 1969. "Lexical retention and cultural significance in Tzeltal-Tzotzil comparative ethnobotany." Paper presented at the 68th Annual Meeting of the American Anthropological Association, 20–23 November, 1969.

Bostoen, Koen. 2007. Bantu plant names as indicators of linguistic stratigraphy in the western province of Zambia. Doris Payne and Jaime Peña, eds., *Selected Proceedings of the 37th Annual Conference on African Linguistics*, pp. 16–29.

Bowern, Claire. 2007. "On Eels, Dolphins, and Echidnas: Nyulnyulan Prehistory through the Reconstruction of Flora and Fauna.: In Alan Nussbaum, ed., *Verba Docenti: Studies in historical and Indo-European Linguistics, Presented to Jay H. Jasanoff by Students, Colleagues, and Friends*. Ann Arbor: Beechstave Press, pp. 39–53.

————. (work in progress with Patrick McConvell). 2011. "Reconstruction in ethnobiological systems of Australian languages." PowerPoint presentation. Australian Nacional University.

Campbell, Lyle, Terrence Kaufman, and Thomas Smith-Stark. 1986. "Meso-America as a linguistic area." *Language* 62:530–70.

Dakin, Karen. 2003. "Uto-Aztecan in the linguistic stratigraphy of Mesoamerican prehistory," in Henning Andersen, ed., *Language Contacts in Prehistory: Studies in Stratigraphy*. Philadelphia: John Benjamins, pp. 259–88.

Fowler, Catherine S. 1972a. "Comparative Numic ethnobiology." Ph.D. dissertation, University of Pittsburgh.

————. 1972b. "Some ecological clues to Proto-Numic homelands." In D. D. Fowler, ed., *Great Basin Cultural Ecology: A Symposium*. Reno, Nev.: Desert Research Institute Publications in the Social Sciences, no. 8, pp. 105–21.

————. 1983. Lexical clues to Uto-Aztecan prehistory. *International Journal of American Linguistics* 49:224–57.

Friedrich, Paul. 1970. *Proto-Indio-European Trees: The Arboreal System of a Prehistoric People*. Chicago: University of Chicago Press.

Haspelmath, Martin, and Uri Tadmor, eds., 2009. *Loanwords in the World's Languages: A Comparative Handbook*. The Hague: De Gruyter Mouton (particularly chap. 1: Martin Haspelmath and Uri Tadmor, "The loanword typology project and the World Loanword Database" (pp. 1–34), Martin Haspelmath, "Lexical borrowing: Concepts and issues" (pp. 35–54), Uri Tadmor, "Loanwords in the world's languages: Findings and results" (pp. 55–75).

Meroz, Yoram. 2013. "Large-scale vocabulary surveys as a tool for linguistic stratigraphy: A California case study." Paper presented at the 39th Berkeley Linguistic Society Meetings, Feb. 16–17, 2013.

Siebert, Frank T. 1967. "The original home of the proto-Algonquian people." *Contributions to Anthropology: Linguistics I. National Museum of Canada Bulletin* 214:13–47.

Smith-Stark, Thomas C. 1982. "Mesoamerican calques." Paper presented at the symposium on Mesoamerican dialectology and language history, 44th International Congress of Americanists, Manchester, England.

————. 1994. Mesoamerican calques. I. In Carolyn J. MacKay and Verónica Vázquez, eds., *Investigaciones lingüísticas en Mesoamérica.* Mexico City: Universidad Nacional Autónoma de México, pp. 15–50.

Whistler, Kenneth W. 1977. "Wintun prehistory: An interpretation based on linguistic reconstruction of plant and animal nomenclature." *Proceedings of the Berkeley Linguistic Society* 3:157–74.

2. The debate on the Uto-Aztecan homeland

Campbell, Lyle. 2003. What drives linguistic diversity and language spread?, in Peter Bellwood and Colin Renfrew, eds. *Examining the Farming/Language Dispersal Hypothesis.* Cambridge, UK: McDonald Institute for Archaeological Research, pp. 49–64.

Hill, Jane H. 2001. "Proto–Uto-Aztecan: A community of cultivators in central Mexico." *American Anthropologist* 103:913–34.

————. 2002. Toward a linguistic prehistory of the southwest: 'Azteco-Tanoan' and the arrival of maize cultivation. *Journal of Anthropological Research* 58:457–75.

————. 2003. "Proto–Uto-Aztecan and the northern devolution." In Peter Bellwood and Colin Renfrew, eds., *Examining the Farming/Language Dispersal Hypothesis.* Cambridge: McDonald Institute for Archaeological Research, pp. 331–40.

————. 2008a. "Northern Uto-Aztecan and Kiowa-Tanoan: Evidence for contact between the proto-languages?" *International Journal of American Linguistics* 74: 155–88.

————. 2008b. "Otomanguean loan words in Proto–Uto-Aztecan maize vocabulary?". In John D.Bengtson, ed., *In Hot Pursuit of Language in Prehistory.* Philadelphia: John Benjamins. pp. 309–20.

————. 2010. New evidence for a Mesoamerican homeland for Proto-Uto-Aztecan. Proceedings of the National Academy of Sciences 107(11):E33. doi:10.1073/pnas.0914473107

Merrill, William L. n.d. "The historical linguistics of Uto-Aztecan agriculture." ms.

Merrill, William L. et al. 2009. "The diffusion of maize to the southwestern United States and its impact. *Proceedings of the National Academy of Sciences* 106:21019–26.

————. 2010. "Reply to Hill and Brown: Maize and Uto-Aztecan cultural history. *Proceedings of the National Academy of Sciences* 107(11):E35–36. doi:10.1073/pnas.10000923107.

3.Historical sources on Nahuatl (Aztec) ethnobiology

Bravo Hollis, Helia. 1934. "Las cáctaceas entre los antiguos mexicanos." *Anales del Instituto de Biología Mexicana* 5:139–48.

Clayton, Martin, Luigi Guerrini, and Alejandro de Ávila. 2009. *The Paper Museum of Cassiano dal Pozzo. Serie B, Part Eight. Flora: The Aztec Herbal.* London: Royal Collection Enterprises.

Comisión Editora de las Obras de Francisco Hernández. 1984. *Comentarios a la obra de Francisco Hernández.* Mexico City: Universidad Nacional Autónoma de México. [see also Hernández, Francisco].

de la Cruz, Martín, and Juan Badiano. 1940 [1552]. *The Badianus manuscript, Codex Barberini, Latin 241, Vatican Library: An Aztec herbal of 1552.* Edited by Emily Walcott Emmart. Baltimore: Johns Hopkins University Press.

de la Cruz, Martín. 1991 [1552]. *Libellus de medicinalibus indorum herbis. Manuscrito azteca de 1552 según traducción latina de Juan Badiano. Versión española con estudios y comentarios por diversos autores.* 2 vols. Mexico City: Fondo de Cultura Económica; Instituto Mexicano del Seguro Social.

Dressler, Robert L. 1953. "The Pre-Columbian Cultivated Plants of Mexico." *Botanical Museum Leaflets* (Harvard University), Vol. 16, num. 6

Hernández, Francisco. 1959 [1571–76]. *Historia natural de Nueva España.* Vol. 1: *Historia de las plantas de Nueva España.* Vol. 2: *Historia de las plantas de Nueva España (conclusion). Historia de los animales de Nueva España. Historia de los minerales de Nueva España.* In *Obras completas*, vols. 2–3. Mexico City:

Universidad Nacional Autónoma de México. [see also Comisión Editora de las Obras de Francisco Hernández].

Kumate, Jesús, et al. 1992. *Estudios actuales sobre el* Libellus de medicinalibus indorum herbis. Mexico City: Secretaría de Salud.

Martín del Campo, Rafael. 1938. "Ensayo de interpretación del libro undécimo de la *Historia* de Sahagún [I: Los reptiles]." *Anales del Instituto de Biología* 9:379–91.

————. 1940. "Ensayo de interpretación del libro undécimo de la *Historia general de las cosas de la* Nueva España de Bernardino de Sahagún II: Las aves." *Anales del Instituto de Biología* 11:385–408.

————. 1941. "Ensayo de interpretación del libro undécimo de la Historia de Sahagún. III: Los mamíferos." *Anales del Instituto de Biología* 12:489–506.

————. 1957. "Las cactáceas entre los mexica." *Cactaceas y Succulentas Mexicanas* 2(2):27–38.

Navarro, Fray Juan. 1992 [1801]. *Historia natural, o, Jardín americano: Manuscrito de 1801*. Mexico City: Universidad Nacional Autónoma de México, Instituto Mexicano del Seguro Social, Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado.

Paso y Troncoso, Francisco del. 1886. "Estudios sobre la historia de la medicina en México. I: La botánica entre los nahuas. *Anales del Museo Nacional de México* 3:137–235. Reprinted in *La botánica entre los nahuas y otros estudios*. Introducción, selección y notas por Pilar Máynez. Mexico City: Secretaría de Educación Pública, pp. 27–188.

Reko, Blas Pablo. 1947. "Nombres botánicos del manuscrito Badiano." *Boletin de la Sociedad de Botánica Mexicana* 5:23–43.

Rivera M., Irene. 1941. "Ensayo de interpretación botánica del libro X de la Historia de Sahagún." Anales del Instituto de Biología 12(1):439–88.

Sahagún, Bernardino de. 1963. *Florentine Codex: General History of the Things of New Spain. Part 12: Book 11: Earthly Things.* Translated and edited by Charles E. Dibble and Arthur J. O. Anderson. Albuquerque: School of American Research; Salt Lake City: University of Utah Press.

Urbina, M. 1903. "Plantas comestibles de los antiguos mexicanos." *Anales del Museo Nacional de México.* II, 1:503–91.