Kathryn Hudson

THE BENEFITS OF BIO(LEXICOGRAPHY)
A Topical Approach to Lexicographic Practice

Abstract This paper combines insights drawn from the author's ongoing linguistic and ethnographic work in the Americas and Micronesia with lessons from Lichtenberk's (2003) work on creating dictionaries for languages in transition to explore the utility of biolexicography and other topical lexicographic approaches. Particular attention is given to key features of biolexicography – including biolexica, the role of communities of practice, the use of ethnographic methodologies, considerations of identity, and linguistically-mediated ecological engagement – and to the relevance of this and other topical lexicographic approaches for languages in transition and their speakers. The ways in which these approaches. are shaped by the lexicographic inheritance and by engagements with languages experiencing attrition or change are also considered. This discussion illustrates that the utility of biolexicography is rooted in the linguistic and sociocultural significances of biolexica in ways that integrate lexicographic products into broader cultural and linguistic systems and center speaker communities in the lexicographic process.

Keywords ecology; ethnography; language change; lexicography

1. Introductory Remarks

Topical or thematic lexicographic materials focus, exclusively or primarily, on a particular subject matter recognized in a particular sociocultural context or on a cluster of topics closely related within a particular cultural tradition. They are often conceptualized as extensions of – or supplements to – more generalized lexicographic productions rather than as a potentially primary methodology; however, their possibly utility is far greater than these classifications imply. This is particularly true for languages experiencing attrition or change and for those facing assimilationist pressures from colonial linguistic systems. It is also true for languages spoken by communities in the process of (re)asserting or (re)establishing their identities, particularly in postcolonial situations.

This paper develops a conceptual and practical examination of biolexicography – an ethnographically-rooted and ecologically-based topical methodology for community driven, use-oriented lexicographic work – to explore this and other forms of topical lexicography. It combines lessons from Lichtenberk's engagement with the complexities of creating dictionaries for languages in transition with insights drawn from the author's ongoing linguistic and ethnographic work in the Americas and Micronesia to explore the practicality and utility of topical lexicographic approaches. Particular attention is given to (i) the relevance of this approach for languages in transition and their speakers, and (ii) the necessity of an ethnographic foundation that is critically oriented and collaboratively developed.

2. The Lexicographic Inheritance and Language Change

Throughout much of its long history, lexicography has held "a mediating position for language use" focused on "[e]xplaining foreign languages or areas of native languages which remain foreign…[or] unintelligible to the common speaker and writer" (Hüllen, 2000, p. 6). This is, in part, a product of its developmental history and trajectory. It is also reflective of the close association of lexicography with academic linguistics and of the authority that is frequently ascribed to – or claimed by – academic researchers. By explaining language with and through language (see Hüllen, 2020) and using lexicographic presentations to facilitate "correct and successful [linguistic] performance" (Hüllen, 2020, pp. 4–5), lexicography eventually emerged as a subdiscipline of linguistics that shares "basic assumptions about language with other subdisciplines" (ibid.) and is central to many documentation efforts. The lexicographic inheritance and issues of language change helped motivate this process; considering their role in the development of lexicography is thus crucial to any consideration of the development and implementation of more specific lexicographic approaches.

2.1 The Lexicographic Inheritance

Biolexicography – like other topical lexicographic approaches – represents a critical and highly contextualized response to the lexicographic inheritance. The earliest known lexicographic materials were developed in ancient Mesopotamia, where monolingual Sumerian wordlists were used for training scribes by c. 3200 BCE and bilingual Sumerian-Eblaite wordlists existed by c. 2400 BCE (Considine, 2015, p. 605). Similar materials appeared in Egypt by the 18th century BCE, in India around. 300 BCE, in Greece during the 3rd century BCE, and in Chinese compendium of glosses known as the Erya (ibid). Dictionaries of a relatively modern form first appear in China with Yang Xiong's dialect dictionary c. 18 AD and Xu Shen's dictionary of Chinese characters compiled c. 149 AD (Considine, 2015, pp. 605–606). The Western tradition arguably originated with Hesychius of Alexandria's Greek dictionary in the 5th – 6th century AD; other early lexicographic traditions appeared in the Indian subcontinent in the 6th century AD, in the Arab world c. 791 AD, in Tibet c. 814 AD, and in the Hebrew-speaking world c. 902 (ibid).

The strongly academic orientation of these early materials continued into the modern era. Early editions of the Oxford English Dictionary, for example, were developed through "impersonal processes" that produced "a more strictly descriptive format" (Kistner, 2013, p. 802). This reflected a tradition of defining dictionaries as catalogues of words and definitions. It was also a product of the persistent belief that dictionaries can provide "accurate and complete accounts of a community's (standard) language… comprised of words which have specific meanings" (Seargeant, 2011, p.1) – a view that conceptually linked dictionaries with comprehensiveness and with assumed academic (i.e., etic) abilities to identify, define, and record this scope.

This perspective is predicated on a trio of interrelated beliefs: (i) words form the core units of language, (ii) all words possess fixed and delineable meanings, and

(iii) a language is a defined system composed of a vast but defined set of words and their meanings (Seargeant, 2011). Dictionaries are thus tasked with recording languages "in the *fullest possible* detail" (Seargeant, 2011, p. 3, emphasis added) and viewed as "authorit[ies] that tell [] us whether a certain locution is actually part of the language" (Dolezal, 2006, pp. 695). The processes of their creation are similarly imbued with legitimizing power, since dictionaries and dictionary-making often symbolically function as evidence of linguistic realness and validity among both speakers and outside observers (see e.g., Corris et al., 2000). Lexicographers thus have a tendency to "want to include as much information as is known" (Corris et al. 2000, 341), and lexicographic work – particularly in contexts of endangerment and attrition – often seeks to preserve a language in its entirety in dictionary form "for future study of revival…[by] linguists and researchers" (Corris et al., 2000, pp. 341, 330). However, "this wealth of information can be counterproductive for users" (Corris et al., 2000, p. 341).

Despite these consequences, the dominance and relative inflexibility of these beliefs has also encouraged the development of new and more customized approaches to dictionary creation. This is particularly true in field-based work and in cases of language attrition and change, since these situations do not always – or often – fit neatly within traditional lexicographic methods and formats (see e.g., Cablitz, 2011; Haviland, 2006; Mosel, 2004, 2011; Vamarasi, 2014). It is also apparent in more general discussions about the formatting of dictionary entries (see e.g., Atkins & Rundell, 2008; Field, 2009; Ivanishcheva, 2016; Kroskrity, 2015), the treatment of loanwords (see e.g., Crowley, 1992; Kroskrity, 1993, 1998; Lichtenberk, 2003), and the particularities of orthographic representation (see e.g., Mosel, 2004). Considered together, these discussions reflect a growing awareness of the need for contextualized approaches sensitive to the needs and experiences of individual communities.

2.2 Lexicography and Language Change

Consideration of – and engagement with – languages undergoing attrition or change further illustrates the need for contextualized lexicographic approaches sensitive to the needs and experiences of individual communities of linguistics practice. In these cases, community conceptualizations of what constitutes fluency – and of what makes someone a speaker – often differ markedly from those of academic linguists and lexicographic researchers. Knowledge of indigenous words is often viewed as evidence of speaker status, even if those words are not actively used or are used only in the context of discourse that primarily utilized a different linguistic tradition. The identities of these words reflect a community's culture and their sociocultural, economic, and political positions within a broader postcolonial landscape. They are also a product of cultural and linguistic politics and values, in academia and beyond, and of agentive processes of identity (re)assertion.

The complex and contextualized nature of language attrition and change make it necessary to critically engage with the lexicographic process. This necessarily begins with audience identification and with the prioritization of different audiences. There

are two general audiences relevant to any lexicographic production: (1) an emic audience made up of speakers and/or members of descendent communities, and (2) an etic audience consisting of academics and other outsiders with an interest in a particular linguistic system. The emic audience should, through linguistic patrimony and ownership, be given priority, though etic audiences are often given equal or greater consideration. Lictenberk (2003) devoted considerable attention to how his dictionary of Toqabaqita would be used. He identified two groups of users – local Toqabaqita people and academics – and believed his dictionary could address the needs and desires of both audiences. However, his concern with the limitations of emic audiences and his belief that it is "unrealistic to think that the dictionary will be frequently used by the local people" because its main value for them will be the recognition garnered by its existence suggest an implicit – and cautionary – focus on etic users (Lichtenberk, 2003, pp. 398–400).

Processes of lexicographic selection and orthographic standardization are equally crucial. Word selection should be culturally motivated, contextualized, and reflective of the language as it exists within a particular speaker community. Loanwords regularly used by speakers should be included regardless of their origins; those that co-occur with indigenous alternatives should be recorded as distinct synonyms. In all cases, the orthography used for these words should match conventions used by speakers, even when they appear illogical or imprecise to professional observers, and a user-friendly orthographic system should be developed in collaboration with community members when no pre-existing system exists. Integrating these practices increases the contextual accuracy and appropriateness of lexicographic materials while also making them more reflective of emic speaker perspectives (see e.g., Haviland, 2006; Pawley, 2001). Selectively applying them, however, can limit the accessibility of lexicographic materials – particularly for emic users – and exacerbate issues in the lexicographic process. These include notions of 'pure' or 'good' linguistic forms, resistance to lexical and phonological change, a desire to edit words deemed overly reflective of a language's colonial experiences, and qualitative judgments of the language used by speakers.

In Lichtenberk's (2003) work with Toqabaqita, for example, issues of lexical variability – particularly loanwords – and orthographic representation are central foci. His treatment of loanwords is based on the degree to which loanwords fit Toqabaqita phonological and phonotactic patterns rather than on the use patterns of these words by speakers (ibid.). Loanwords are included if they fit these patterns or have been modified to fit them, though words without modification are still excluded if they belong to a set in which other words are not modified (e.g., numbers or day names) or if they only occurred once (Lichtenberk, 2003, pp. 394–397). Lexical variability is thus governed by phonological and phonotactic parameters rather than by patterns of use, motivating an academically oriented definition of the Toqabaqita lexicon. Lichtenberk's (2003) orthographic conventions improve the consistency and phonetic accuracy of the writing system but differ – sometimes significantly – from those used by Toqabaqita people, thus further reducing their ability to utilize the dictionary.

3. Biolexicography: A Multispecies, Use-Driven Approach

Biolexicography is an ethnographically rooted and ecologically based topical methodology for contextual and use-oriented lexicographic work being developed through ongoing collaborative and community-driven work with communities in Latin America and western Micronesia.[1] It represents a highly contextualized response to the orthodox beliefs imbedded in the lexicographic inheritance and to the pressures of language attrition and change in postcolonial communities. It also reflects the reconceptualization of culture, identity, and language within these communities and acknowledges the complex relationships that exist between these categories and the more-than-human world. The utility of biolexicography is rooted in the linguistic and sociocultural significances of words pertaining to animal and plant species – collectively described as biolexica – and in its emphasis on communities of practice, its ethnographic foundations, and its cognizance of identity (re)formulation in postcolonial situations. Each of these factors will be considered in turn.

3.1 Biolexica

Within the framework of biolexicography, words pertaining to animal and plant species are collectively known as biolexica. The use of this term instead of more semantically restrictive labels such as zoonyms allows the inclusion of words identifying both plant and animal species, thus allowing the methodology to better align with systems of traditional ecological knowledge and classification. It also facilitates the incorporation of words describing (i) parts of animals and plants, (ii) aspects of the life cycles and behaviors of animals and plant species, (iii) species or species groupings, and (iv) human behaviors related to or based on other species. Biolexica thus catalogue patterns of human interaction with more-than-human worlds – including, among other things, culturally defined relationships between animals and space, conceptualizations of wild versus domestic, cultural typologies of the natural world, and the potential relationships between humans and animals – and record traditional ecological knowledge. They also aid in the reconstruction of precolonial environments, foster ecological engagement in contemporary communities, and can be used to reassert or reestablish real or symbolic connections to traditional homelands and environments. By extension, biolexica can also motivate or maintain interest in protecting and maintaining these spaces and their diverse inhabitants.

3.2 Communities of Practice

Lexicographic work should reflect a specific and well-defined context since dictionaries are "artifact[s] *designed* with care to fit *precise specifications*" (Atkins, 2008, p. 35). These specifications are determined by the relevant community of practice – defined here as a speaker community or descent community with interests in developing, maintaining, and/or using a particular language – and guided by their desires, needs, and interests. Such community-driven work requires awareness of, and sensitivity to,

---

[1] Community elders and other community members request that their identities and specific locations remain anonymous as a condition of access and ongoing work.

the contexts of these communities – cultural, economic, historical, political, and social – and the ways in which they view and use their language within them. This allows the parameters of linguistic use to be defined by speaker communities in contextually appropriate ways that encourage community engagement with lexicographic work. It also facilitates the integration of dictionaries and other lexicographic products into broader cultural and linguistic systems in ways that increase their relevance and utility for speakers.

Such purpose-driven lexicography is particularly significant for speakers of languages undergoing attrition or change. Although many lexicographers working with these languages agree that "dictionaries for endangered languages…differ mostly in the amount of information in certain parts of the microstructure," (Corris et al., 2000, p. 2), there are others (e.g., Cablitz, 2011; Field, 2009; Ivanishcheva, 2016) who advocate for more contextually and culturally sensitive approaches in which these materials also differ in the nature of their content and structure. Biolexicography is one such approach; its focus on biolexica significant in a particular community of practice leads to distinct lexicographic content structured around culturally significance relationships between people, language, and environment. By reflecting the cultural, linguistic, and sociopolitical worlds of a community, biolexicography and other contextualized approaches capture the "cultural constructs embedded and reflected in the…language" (Field, 2009, p. 296). They also preserve "unique facts about the culture, and language as part of the culture, of an indigenous minority" (Ivanishcheva, 2016, p. 83) and incorporate the perspectives of the speaker community.

3.3 Ethnographic Foundations

Biolexicography, like many topical lexicographic approaches, is strongly rooted in ethnographic fieldwork since "a great deal of cultural knowledge is part of the linguistic competence of speakers" (Cablitz, 2011, p. 447). Its methods are culturally motivated and contextualized; its products incorporate cultural knowledge and practices in ways that articulate the sociocultural contexts of a language with the historical, economic, and sociopolitical landscapes of its speakers. The integration of lexicographic and ethnographic methodologies is thus crucial. Simply being in the field for purposes of data acquisition does not allow for the development of nuanced and contextualized understandings of words, their meanings, or their uses; engagement with, and participation in, the field are required. This is particularly true of biolexicography, since identifying and understanding biolexica – and, crucially, their relationships to other cultural and linguistic systems – necessitates engagement with contexts in which the traditional ecological knowledge associated with these words is made manifest through the intersection of sociocultural and linguistic behaviors. When ethnographic fieldwork is foundational to lexicographic endeavours in this way, the facts that "[c]ulture permeates lexicons" and "cultural traditions are sometimes inseparable from linguistic form" are recognized in ways that benefit languages and their speakers (Field, 2009, p. 300). Ethnographic perspectives thus facilitate the lexicographic selection of words and help define the foci of topical

lexicographic approaches: they reveal which words are active participants in speakers' cultural lives, which relate to constructions and conceptualizations of identity, and – by extension – which are *functioning* parts of the language as it is used in a particular community.

3.4 Implications for Identity and the More-Than-Human World

Biolexica – like the lexicographic materials that record them – provide a means of asserting cultural and linguistic competency and identity through affiliation with particular environments, even when more expansive knowledge of cultural and linguistic systems has been lost or eroded. The knowledge and use of biolexica reflect an individual's access to the cultural knowledge associated with them and, by extension, indicate sufficient engagement with the relevant ecological systems to use these words appropriately. They are also often viewed as evidence of a connection with history and with the legitimacy that derives from it. Postcolonial indigenous communities often feel "a...need to demonstrate that they exist, have been present for generations, and have co-shaped the culture and cultural landscape of a certain place" (Jordan, 2012, p. 129); this need can be addressed through use of biolexica recorded in lexicographic materials and through access to – or possession of – these materials. By linguistically linking themselves with their environment in this way, communities more permanently connect themselves to particular locations and provide evidence of historical existence.

Biolexicography also has implications for conservation efforts and sustainability initiatives, which frequently parallel efforts for maintain and preserve cultural identities. Biodiversity is generally more stable in areas managed by indigenous populations, and the practices that facilitate this – including the creation of diverse and species-rich spaces, the active management of land, the pursuit of ecological restoration, and the diachronic monitoring of ecosystems – are often rooted in lexicographically preserved systems of traditional ecological knowledge. The ethnographic elicitation and lexicographic recording of biolexica can trigger these efforts, guide their methodologies, and contribute to their maintenance, particularly when cultural and practical information is recorded in the definitions, etymologies, example sentences, and/or supplemental discussions that accompany them.

4. Concluding Remarks

Biolexicography provides an ethnographically rooted and ecologically-based topical methodology for community driven, use-oriented lexicographic work that can be used to explore the utility of topical lexicography. The lexicographic inheritance and engagement with issues of language attrition and change motivated its development; their role in the development of lexicography more generally provides a useful framing for considerations of this and other topical lexicographic approaches. The utility of biolexicography is rooted in the linguistic and sociocultural significances of biolexica; its emphasis on communities of practice, ethnographic foundations, and connection to processes of identity (re)formulation facilitates the integration of its

products into broader cultural and linguistic systems. The ongoing development of this and other topical approaches to lexicography will continue to illustrate their significance, particularly in contexts of attrition and change, while also centering speaker communities in the lexicographic process.

References

Atkins, S. (2008). Theoretical Lexicography and its Relation to Dictionary-making. In T. Fontenelle (Ed.), *Practical lexicography: A Reader* (pp. 31–50). Oxford University Press.

Cablitz, G. H. (2011). Documenting Cultural Knowledge in Dictionaries of Endangered Languages. *International Journal of Lexicography 24*(4), 446–462.

Considine, J. (2015). A Chronology of Major Events in the History of Lexicography. In P. Durkin (Ed.), *The Oxford Handbook of Lexicography* (pp. 605–624). Oxford University Press.

Corris, M., Manning, C., Poetsch, S., & Simpson, J. (2000). Dictionaries and Endangered Languages. In D. Bradley, & M. Bradley (Eds.), *Language Endangerment and Language Maintenance: An Active Approach* (pp. 329–348). Taylor & Francis Group.

Crowley, T. (1992). *A Dictionary of Paamese.* Australian National University.

Dolezal, F. (2006). World Englishes and lexicography. In B. Kachru, Y. Kachru, & C. Nelson (Eds.), *The Handbook of World Englishes* (pp. 694–708). Blackwell.

Duval, T., & K. Kuiper. (2001). Maori Dictionaries and Maori Loanwords. *International Journal of Lexicography 14*, 243–60.

Field, M. C. (2009). Metaphor, Mythology, and the Navajo Verb: The Role of Cultural Constructs in the Lexicography of Endangered Languages. *Anthropological Linguistics 51*(3/4), 296–302.

Haviland, J. (2006). Documenting Lexical Knowledge. In J. Gippert, N. P. Himmelmann, & U. Mosel (Eds.), *Essentials of Language Documentation* (pp. 1–30). Mouton de Gruyter.

Hüllen, W. (2000). *English Dictionaries, 800-1700: The Topical Tradition.* 1st ed. Oxford University Press, Incorporated.

Ivanishcheva, O. N. (2016). Dictionaries of Critically Endangered Languages: Focus on Users. *Journal of Linguistics 67*(1), 73–86.

Kistner, K. (2013). 'A Word Factory was wanted': Organizational objectivity in the making of the Oxford English Dictionary. *Social Studies of Science 43*(6), 801–828.

Kroskrity, P. V. (1993). *Language, history, and identity: Ethnolinguistic studies of the Arizona Tewa.* University of Arizona Press.

Kroskrity, P. V. (2010). Language ideologies – evolving perspectives. In Ju. Jaspers, J.-O. Östman, & J. Verschueren (Eds.), *Society and language use, Handbook of Pragmatics Highlights, Vol. 7* (pp. 192–211). John Benjamins.

Kroskrity, P. V. (2015). Designing a Dictionary for an Endangered Language Community: Lexicographical Deliberations, Language Ideological Clarifications. *Language Documentation & Conservation, 9*, 140–157.

Lichtenberk, F. (2003). To List or Not to List: Writing a Dictionary of a Language Undergoing Rapid and Extensive Lexical Changes. *International Journal of Lexicography 16*(4), 387–401.

Mosel, U. (2004). Dictionary Making in Endangered Speech Communities. In P. K. Austin (Ed.), *Language Documentation and Description, Vol. 2* (pp. 39–54). School of Oriental and African Studies.

Mosel, U. (2011). Lexicography in Endangered Language Communities. In P. K. Austin, & J. Sallaback (Eds.), *The Cambridge Handbook of Endangered Languages* (pp. 337–353). Cambridge University Press.

Pawley, A. (2001). Some Problems of Describing Linguistic and Ecological Knowledge. In L. Maffi (Ed.), *On Biocultural Diversity: Linking Language, Knowledge, and the Environment* (pp. 228–247). Smithsonian Institution Press.

Seargeant, P. (2011). Lexicography as a philosophy of language. *Language Sciences 33*, 1–10.

Vamarasi, M. (2014). The creation of learner-centered dictionaries for endangered languages: a Rotuman example. *Journal of Multilingual and Multicultural Development 35*(2), 167–178.

Contact information

Kathryn Hudson
University at Buffalo
khudson@buffalo.edu

This paper is part of the publication: Despot, K. Š., Ostroški Anić, A., & Brač, I. (Eds.). (2024). *Lexicography and Semantics. Proceedings of the XXI EURALEX International Congress.* Institute for the Croatian Language.

461