

# Building the lexicon for awakening languages

Daniel W. Hieber

University of California, Santa Barbara

## 1 Introduction

Just as language endangerment involves a reduction in the domains of use for a language, so does language revitalization entail an expansion of the language into new (or previous) social domains. Therefore a major task facing language revitalization efforts is the need to augment the lexicon to accommodate these new domains. Yet little guidance or training is available to communities or the linguists working with them on how this is to be done. This is especially unfortunate because coining new terms and other means of expanding the lexicon are not at all straightforward tasks, either linguistically or socioculturally. The process can be fraught with conflicting language ideologies, and should ideally be informed by the lexeme- and phrase-building processes that already exist in the language. This chapter provides essential guidance on how to avoid these pitfalls with successfully expanding the lexicon into new domains. It is written primarily for language revitalization teams and the linguists that work with them.

Expanding a lexicon is not simply a matter of creating new words for new things. In many cases, as will be seen, this would be an inappropriate way to accomplish the task. It must always be kept in mind that *the central goal of revitalization-oriented lexicographic work is to accommodate new uses of the language in a manner commensurate with the wishes and language ideologies of the community*. Occasionally this entails creating new bits of language, but much more often it entails giving new uses to old words instead. Each language already has within it the remarkable potential to accomplish whatever social aims a human mind can put it to. Expanding the lexicon is simply a matter of teaching speakers how to do this, given the linguistic resources at their disposal.

It is important to note that neologisms are not just words, but can be any means of referring to a concept or thing. In Western Apache, place names consist of phrases that physically describe the location in some way, such as *tséká' tú yahilí'* 'water flows downward on top of a series of flat rocks' (Basso 1996:27). This is not a noun or even a nominalized phrase, but rather a description, yet this is the canonical method of referring to places. So this chapter is not so much about expanding the 'lexicon' as expanding the whole set of conventionalized

phrases, constructions, and terms that speakers have at their disposal; but for simplicity's sake I will continue to use the term 'lexicon'.

I distinguish between two types of neologisms. One type is WORD CREATION, which I define here as the creation of new morphemes through novel combinations of strings of phonemes, or what one often means by 'making up words'. This is done frequently in product naming, branding, and marketing, and is the origin of names such as *Kodak* and *Verizon*, but is rarely done in language revitalization projects, for reasons I will return to later. A second type of neologism I refer to as a COINAGE, which makes use of morphemes and words already in the language, but in novel ways. The recent English coinage *hashtag* is one such example. It contains two morphemes already present in English at the time of its creation (*hash*, referring to the symbol <#>, and the word *tag*), but combines them in a novel way with a novel meaning.

## 2 Issues & Ideologies

The ways in which a language revitalization team chooses to expand the lexicon, and whether those choices will lead to success or failure, depend crucially on the language attitudes and ideologies of both the dedicated language team and the community at large. At the heart of all this is the issue of *authenticity* – what counts as authentic language according to each member of the language team and the community?

The answer to this question varies tremendously from community to community and person to person. At one extreme is an ideology that says that all new ways of using the language are inauthentic, and is reflected in the prescriptivist ideologies of those who find all linguistic change to be a corruption of the language. In its extreme form, this ideology also prohibits borrowing or code-switching – the language should be pure and unmixed or not used at all. This ideology is a death knell for any revitalization project, for if the language cannot be used in novel ways, it can never expand into new domains. At the other end of the spectrum is an ideology which does not give any special status to the heritage language, so that there is no imperative to adhere to how the language actually is or was. Using words or even the phonology of another language is considered completely legitimate.

Language ideologies regarding authenticity typically lie somewhere between these two extremes. For example, the Mohawk Language Standardization Project states in its recommended writing guidelines that new words “are to be formed by function, activity, or characteristic”, but that loan words from other languages are acceptable as well (Jacobs, Thompson & Leaf 1993, cited in Grenoble & Whaley 2006:92). This demonstrates a preference for using existing Mohawk words over loanwords, presumably reflective of an ideology that

this is more authentic Mohawk than, say, English words pronounced with Mohawk phonology. Hinton & Ahler (1999:62) report a much different ideology for some participants in Karuk revitalization efforts, which one speaker summarizes by saying, “I am interested in communication, not preservation.” This speaker was less concerned with pronunciation and grammatical accuracy than simply using the language whenever possible.

Revitalization teams should think hard on what they consider to be authentic language before embarking on a project to expand the lexicon, and ask themselves whether these ideologies are in line with their broader revitalization goals. This is true regardless of whether revitalization efforts will involve a centralized language planning committee that coins new terms, or adopts a distributed, bottom-up approach instead. If one goal of a revitalization project is to see speakers using the language on the internet, would it be considered inauthentic for a speaker to use the word *Skype* as a verb in a blog post (e.g. *I’m going to Skype with them later*)? If so, the language team should be prepared to coin terms for a large variety of tasks and concepts relating to electronically-mediated communication, or have some plan in place for teaching speakers/learners how to utilize existing linguistic resources to this end. On the other hand, if the revitalization team encourages borrowings for technical terms, will speakers find it strange or inappropriate to hear speech peppered with terms like *Skype*, *gif*, *Facebook*, *newsfeed*, and *Snapchat*? A compromise position is to recommend that speakers make free use of jargon words like these, but integrate them phonologically and structurally into the language. Again, this requires educating new speakers on how this is done in the language. As an example, both Tlingit and Chickasaw have borrowed the word *Skype* as a verb stem that takes normal verbal inflection and fits the phonology of the language, so that the phrase ‘we Skyped him’ in Tlingit becomes *ash wutusikáayp* (Crippen & Twitchell 2013; Hinson 2014).

Revitalization efforts also need to take into account community reactions to the effects of decisions made by language teams. Prominent Māori language activist Tīmoti Kāretu often states that coining new terms via committee is unnecessary for Māori language revitalization, reasoning that one has merely to go ask speakers to find out that they are already using terms for the concept. A bigger concern, in his opinion, is a proliferation of terms that few people use, and that won’t be understood by existing speakers of the language. This raises two crucial issues for lexical expansion projects: First, centralized approaches where neologisms are created by a dedicated team always face the problem of dissemination. Coining the terms is just the first step. How will the language team educate new speakers about these terms? Including them in any dictionary projects is one straightforward step, but other strategies include word-of-the-day/week programs via various media outlets, or asking language instructors to include the new terms in their curricula. Just as important as educating new speakers is informing

existing speakers about the new terms, which raises the second issue: mutual comprehensibility between speakers, and in particular between existing and new speakers. Neologisms are often opaque, and only make sense once explained. Spolsky & Boomer (1983:235) report one coinage by a Navajo speaker of the term *saad ahááh deidinili*, that translates literally as, ‘one who links conversations together’. It is unlikely that anybody would guess that this refers to a telephone operator without an explanation or sufficiently strong context to inform them. Moreover, when traditional speakers – who are typically older and regarded as most knowledgeable about the language – are confronted with terms that are unfamiliar, confusing, or that they simply don’t know, this has the potential to foster negative attitudes towards new uses of the language. And because the new speakers now command phrases that traditional speakers do not, this even challenges the elders’ status as an authority on the language. A common reaction is to reject these neologisms as inauthentic, creating linguistic rifts within the community. One of the best ways to address this potential hurdle, aside from fostering a general attitude of open-mindedness among both new and existing speakers alike, is to coin terms that are as faithful to traditional patterns of coinage as possible in both form and content. Then, when traditional speakers encounter the new terms, it will be in grammatical constructions and ways of speaking that are familiar to them, and which they are most comfortable with. Again, the best solution to this problem is to coin terms that are faithful to traditional patterns in the language.

Faithfulness to the patterns of the traditional language goes far beyond the mere assembling of morphemes into a grammatical utterance. It also includes faithfulness to traditional values, practices, and ways of viewing the world (Hinton & Ahlers 1999). For example, learning place names in Western Apache is a crucial part of one’s linguistic and cultural socialization, because each place is associated with a specific moralizing tale about “persons who suffer misfortune as the consequence of actions that violate Apache standards for acceptable social behavior.” (Basso 1984:35–36). A failure to use these place names when they would be called for, or coining some neologism in their stead, would only alienate traditional speakers, and prompts statements such as that by Apache speaker Ronnie Lupe: “Our children are losing the land. It doesn’t go to work on them anymore. They don’t know the stories about what happened at these places. That’s why some get into trouble.” (Basso 1984:21). Under the umbrella of “ways of viewing the world” sit things like traditional metaphors. In Hupa, journeying is an important metaphor for both life and marriage, giving rise to expressions like *minejit na’asiya* ‘center his walking’ (i.e. ‘middle-aged, middle of life’) and *nayaseL* ‘they are walking together’ (i.e. ‘they are married’) (Hinton & Ahlers 1999:65). Hinton & Ahlers note that this metaphor could be exploited for many related terms, such as ‘divorce’. Again, neologisms based on such

metaphors will be much more accessible to community members who have been socialized into traditional culture and language, while simultaneously providing cultural insights for new speakers as well.

In sum, for any revitalization project it is vital to keep in mind the language ideologies of community members, their possible reactions to new uses of the language, and the important cultural values and practices that ought to be reflected in the language. I turn now to the different broad strategies that one can take towards expanding the lexicon.

### 3 General Approaches

Broadly speaking, there are two approaches that revitalization teams can adopt in expanding the lexicon: *CENTRALIZED* and *DECENTRALIZED*. I prefer these terms to the more common terms *top-down* and *bottom-up* because they avoid any implication or imposition of social hierarchy where none may actually exist. Centralized approaches are those where a dedicated team of language experts or otherwise duly-appointed individuals make decisions regarding the lexicon and in some way disseminate those decisions to the rest of the speech community.

Decentralized approaches encourage new and/or existing speakers to create their own neologisms instead. The role of the language team in decentralized approaches can either be to document and disseminate the terms that community members are using, or to let speakers adopt whichever terminology they prefer as neologisms spread through the community, until speakers standardize on just one or several terms. In this second approach, where the adoption of neologisms is largely emergent, the primary task of the revitalization team is to educate new speakers on the methods of coining new terms that reflect the ideologies of the community (§2).

Crippen (2013) speaks of the difference between these two approaches in terms of ‘fiat’ (centralized) versus ‘consensus’ (decentralized), and points out several drawbacks of centralized approaches. First is that the language team may become viewed as the only authority on what is considered correct for the language. People may be reluctant to use the language if they feel they can’t use it correctly. Second is that a firm insistence on standards may actually retard language growth. Crippen talks of the need to “let go” of the language into the community. The primary goal of most revitalization projects, after all, is for speakers to pass on and grow their language on their own. Additional risks for centralized approaches is that the language team might coin terms where appropriate ones already exist in the community, or that not everyone will like the new terms.

Still, it is important to realize that centralized lexicon expansion is not necessarily the same thing as ‘a firm insistence on standards’. A compromise position, mentioned earlier, is for the

language team to collect and report on neologisms in the community rather than dictate them. The difference is essentially one of prescriptivism versus descriptivism. As a more specific method, the language team might encourage weekly contests for community members to coin new terms for specific concepts. The language team could then choose the best neologism based on whether it adheres to both the grammatical conventions in the language and traditional values and ways of seeing the world. In many communities though, the support for language revitalization is not yet broad enough to receive sufficient participation in such a contest.

Decentralized approaches are not without risks either. The most obvious one is that mass second language acquisition tends to affect the grammar of the language, usually discussed in terms of ‘imperfect learning’ of the traditional language. This has been well documented for Dyirbal, where morphological ergativity is no longer present among the language of younger speakers (Schmidt 1985). Again, the most effective way of mitigating against this risk is educating new speakers about their language.

Another matter entirely is when the language has not been spoken in the community for some time, and there remain no native speakers. In this case, the team doing language work is the only source of information on the language. Even then, decentralized approaches are possible if the language team takes as its focus instructing language learners on how to form new terms in linguistically and culturally appropriate ways. A major advantage of this approach is that it relieves the language team of the burden of having to coin hundreds of new terms, allowing them to focus on the task of making information on the grammar and previously-documented lexicon available instead. In this way lexical expansion can be both centralized and decentralized at the same time, through a division of labor between the language team and the community of learners.

Regardless of the strategy a language team adopts, the goals of the revitalization effort should be kept prominently in mind. If the goal is for speakers to learn how to be creative with the language, and adapt it to new uses themselves, then the language team should make its focus providing language learners with the linguistic resources they need to take the language into new domains. Small language teams cannot do everything themselves, but language learners also need somebody to turn to as an authority on the language. This will almost always entail some division of labor between the language team and the community of learners.

#### **4 Conclusion**

Expanding the lexicon of an awakening language into new domains is never an easy task. This paper has attempted to provide guidance that will help make this process easier for language revitalization teams and linguists alike. It has examined the many sociopolitical and

practical factors that go into the process of lexical expansion, and suggested proactive ways of avoiding potential issues that may even improve the outcome of lexical expansion projects. Through an efficient division of labor between dedicated language activists and the community of speakers at large, the task of expanding the language into new domains does not have to be an insurmountable one.

## References

- Basso, Keith H. 1984. "Stalking with stories": Names, places, and moral narratives among the Western Apache. *Text, play, and story: The construction of reconstruction of self and story*. Prospect Heights, Illinois: Waveland Press, Inc.
- Basso, Keith H. 1996. *Wisdom sits in places: Landscape and language among the Western Apache*. Albuquerque: University of New Mexico Press.
- Crippen, Dzéiwsh James A. & X'unei Lance Twitchell. 2013. Developing consistency by consensus: Avoiding fiat in language revitalization. *ICLDC 2013*. Honolulu.
- Grenoble, Lenore A. & Lindsay J. Whaley. 2006. *Saving languages: An introduction to language revitalization*. Cambridge: Cambridge University Press.
- Hinson, Joshua. 2014. Anompa Himitta': Chickasaw neologisms. *SILLA 2014 Winter Meeting*. Minneapolis.
- Hinton, Leanne & Jocelyn Ahlers. 1999. The issue of "authenticity" in California language restoration. *Anthropology & Education Quarterly* 30(1). 56–67.
- Jacobs, Annette Kaia'titahkhe, Nancy Kahawinonkie Thompson & Minnie Kaia:khons Leaf (eds.). 1993. The Mohawk language standardization project.  
<http://www.edu.gov.on.ca/eng/training/literacy/mohawk/mohawk.html>.
- Schmidt, A. 1985. *Young people's Dyrbal: An example of language death from Australia*. Cambridge: Cambridge University Press.
- Spolsky, Bernard & Lorraine Boomer. 1983. The modernization of Navajo. In Juan Cobarrubias & Joshua A. Fishman (eds.), *Progress in language planning: International perspectives*, 235–252. Berlin: Walter de Gruyter.