2004. In Proceedings of CATaC '04.

REPRESENTATIONS OF INDIGENOUS LANGUAGE GROUPS OF NORTH AND SOUTH AMERICA ON THE WORLD WIDE WEB

In whose voice?

SUSAN C. HERRING AND ZILIA C. ESTRADA
Indiana University, Bloomington, USA

Abstract. The Web has been claimed to support indigenous languages and cultures by giving members of such groups a voice. As yet, however, this claim has not been systematically examined on a large scale. This paper presents the findings of a preliminary survey and content analysis of Web sites about small indigenous linguistic groups in North and South America. Initial observations reveal both a low overall use of indigenous languages, and a predominance of English and North American perspectives on sites about groups from both continents. At the same time, the North American groups more often than the South American groups "speak" on the Web in their own culture's voice, albeit in English.

1. Introduction

The World Wide Web has been claimed to permit speakers of indigenous languages to create content, thereby promoting their agendas and enhancing the vitality of their languages and cultures (Buszard-Welcher, 2001). As yet, however, we know relatively little about the extent to which the Web is actually benefiting indigenous languages. Most statistics about the "languages of the Internet" are based on population demographics, rather than on studies of actual Internet use (Paolillo, 2003). Conversely, studies of Web use by minority language groups tend to be individual case studies (e.g., Enteen, 2002), rather than generalizable overviews. To begin to address this gap, we report in this paper on the initial findings of a survey and content analysis of Web sites about indigenous language groups in North and South America.

2. Data and Methods

The focus of this study is languages with between 5,000 and 20,000 speakers, as reported in the Ethnologue (2004). Languages of this size can be considered 'endangered' (cf. Grimes, 1995); without intervention, they are at risk of dying out within the next 100 years (Nettle & Romaine, 2001). They are therefore of special interest in discussions of the effects of the Web on linguistic diversity, in that the Web could potentially make a difference to their survival.

We chose North and South America in order to compare the effects of an English-dominant context with a context in which another "big" language—in this case,

Spanish—is dominant. We excluded any language from our sample that is not predominantly spoken in an English region in North America (e.g., languages spoken in French regions of Canada) or in a Spanish region in South America (e.g., languages spoken mostly in Brazil, where Portuguese is the majority language). A total of 15 languages in South America, and 12 languages in North America, met these criteria.

For each of the 27 languages, a Web search was conducted using the Google search engine, first for the name of the language (e.g., 'Hopi'), then for the name plus the English word 'language' (e.g., 'Hopi language'), and then for the language name plus the Spanish word 'lengua' ('language'). For the purpose of this preliminary analysis, the number of hits retrieved for each search was noted, and the content of all the relevant hits in the first two pages (N=20) was examined. This sampling method does not identify all available web sites about any language group; rather, it was intended to identify the content about the languages that is most readily accessible using obvious search terms, and thus that Web searchers are most likely to encounter.

Content analysis methods (Bauer, 2000) were employed to analyze two aspects of the content of the relevant Web sites identified: language(s) used, and evidence of indigenous "voice." For language, we coded the overall language of the Web site (indigenous, English, Spanish, or 'other') and noted how much (if at all) and for what purposes the indigenous language was used. For "voice," we coded the presence of content features such as: historical accounts from the perspective of the people; information about tribal governance; language preservation efforts (including schools, courses and/or availability of language-learning materials) led by natives; community newsletters or newspapers; and discussion forums in which people self-identified as members of the indigenous group. These categories were intended to measure evidence of indigenous presence—web content created by the people to meet the needs of the people—even in sites where no use of the indigenous language was found.

3. Initial Observations

Below we summarize some initial observations based on Web sites associated with nine South American language groups (Guambiano, Guarayu, Huambisa, Michiguenga, Maquiritari, Sanapaná, Trinitario, Tsimané, and Yagua) and six North American language groups (Apache, Cherokee, Hopi, Inuktitut, Micmac, and Zuni).

3.1. LANGUAGE USE

The first and most general finding is that indigenous languages are used very little on the Web sites identified by our sampling method. Most sites that pertain to the people group, with the exception of online dictionaries, contain no words in the indigenous language. When words in the language are used, they tend to be simple greetings (or in several cases, the name of a community newspaper), and they are usually translated into the dominant language. A notable exception is Inuktitut, which boasts an online newspaper with Inuktitut entries (which require an Inuktitut font to be installed on the reader's computer) alongside their English translation, and which contains links in the sidebar to untranslated Inuktitut documents. Another site offers a feature that allows the user to browse in Inuktitut for Inuktitut language movies, of which 11 are listed. The reasons for Inuktitut's special status are considered below.

A second overall finding is that most of the Web sites, including those about South American groups, are in English. Fewer than one-quarter of the South American group sites examined thus far are in Spanish. This is because the sites are mostly produced in North America by English-speaking missionaries, linguists, biomedical researchers, and travel professionals, and describe the people group or the language from an external perspective. A few of the retrieved web sites contain prayers written in South American indigenous languages (e.g., Huambisa), but they are translations from English of Christian prayers. In contrast, the North American sites include a number authored by native peoples (see also Buszard-Welcher, 2001). These are more likely to contain occasional words in the indigenous languages (e.g., to refer to cultural concepts), although most of the writers do not appear to be able to speak or write the languages, and some groups—such as the Hopi and the Cherokee—refer to the language as the language of their ancestors, rather than as their own.

3.2. NATIVE VOICE

The representations of North and South American indigenous groups differ in the degree to which a native perspective or "voice" is present in the Web sites. For the South American groups, the principal evidence of an indigenous perspective is found in Confederación de Pueblos Indigenas (Confederation of Indigenous Peoples) websites which provide information about Indian issues, typically for several ethnolinguistic groups living within a nation state combined. These sites are in Spanish. Similar sites (in English) are found for North American groups, along with central information sites for individual tribes, community newsletters, bulletin boards, and other practical resources that were not found for the South American groups using our search methods.

A further difference is that North American Indians appear to be more likely to create websites containing cultural and historical information about their group, sometimes with the explicitly-stated goal of presenting a native perspective, and sometimes with a political agenda to generate support for native people's rights. Such sites may include images, myths and stories that are sources of indigenous pride. The site creator typically identifies him- or herself as a full-blooded or partial-blooded member of the ethnic group, and refers to the group using first-person plural pronouns ('we', 'our', etc.). A male voice on one such Cherokee site speaks a Cherokee greeting when the site is accessed; one Apache site plays continuous native music.

The last observed difference concerns efforts to promote learning of the indigenous language targeted primarily at members of the ethnic group who wish to reclaim their linguistic heritage. The Official Site of the Cherokee Nation announces free Cherokee language classes (off-site) and links to online language learning resources, including interactive language games for children. Apache (and other indigenous language) instructional tapes and CDs are for sale at one commercial site owned by a Sioux woman and "endorsed by the Apache Nation;" the stated goal of the site is "to promote pride in our Native American heritage." No such sites were identified for the South American groups using the same search methods. Finally, for languages written in a non-roman script (e.g., Cherokee, Inuktitut), a number of sites provide description of the writing system and how the symbols are pronounced, and others offer special fonts; some of these sites, however, appear to have been created by linguists to assist indigenous peoples in preserving their languages.

4. Interpretations

The above observations are based on non-exhaustive Web searches for a small number of languages in one part of the world; therefore, our interpretations at this stage are necessarily tentative. However, if further research bears out a pattern in the Americas of negligible authentic use of small indigenous languages on the Web and a predominance of English language and North American perspectives, it would constitute evidence against the effectiveness of the Web in supporting linguistic diversity. At the same time, we believe that the uses of the Web by North American indigenous groups to pursue their political and cultural agendas hold some promise for language revitalization, in as much as political empowerment can lead to increased rights which in turn can legitimize the language and culture of a people group. This appears to be the case for the Inuktitut language, which has received legitimization—including the creation of a "language commissioner"—by the granting in 1999 of a semi-autonomous territory, Nunavut, for Inuktitut speakers in northern Canada (Office of the Languages Commissioner of Nunavut: 2002).

It is not yet clear to what factors the other observations reported here should be attributed. Is the sparse use of indigenous languages of the Americas on the Web due to the moribund status of the languages, to reluctance (or lack of access) on the part of indigenous peoples to use the Web, or to something else? Does the relative paucity of Spanish-language websites about indigenous groups reflect a political stance towards indigenous peoples in South America, a lesser use of the Web in general, or something else? To what extent do the number of speakers of a language, its geographical isolation, the extent to which its speakers orient towards a tribal versus a pan-Indian identity, and the historical recognition (or lack thereof) by a nation of the rights of indigenous peoples explain variation in their use of the Web? The continuation of the analysis begun here will address these broader interpretive issues.

References

Bauer, M.: 2000, Classical content analysis: A review, in M. Bauer and G. Gaskell (eds.), *Qualitative Researching with Text, Image and Sound*, Sage, Thousand Oaks, CA, pp. 131-151.

Buszard-Welcher, L.: 2001, Can the Web save my language? in L. Hinton and K. Hale (eds.), *The Green Book of Language Revitalization in Practice*, Academic Press, San Diego, pp. 331-348.

Enteen, J.: 2002, Spatial conceptions of URLs: Tamil Eelam on the world wide web, Conference on cultural attitudes towards technology and communication (CATaC), Université de Montréal, July 12-15.

Ethnologue: 2004, *Languages of the World*, Retrieved February 11, 2004 from: http://www.ethnologue.com/

Grimes, J. E.: 1995, Language endangerment in the Pacific, *Oceanic Linguistics*, 34, 1-12.

Nettle, D. and Romaine, S.: 2000, *Vanishing Voices: The Extinction of the World's Languages*, Oxford University Press, Oxford.

Office of the Languages Commissioner of Nunavut: 2002, History of the office. Retrieved February 12, 2004 from: http://www.langcom.nu.ca/english/about/history.html

Paolillo, J. C.: 2003, *Language Diversity on the Internet*, Report submitted to the UNESCO Institute for Statistics as input for the World Summit on the Information Society.