

“Unknown Unknowns” and the Retrieval Problem in Language Documentation and Archiving

Gary Holton

University of Alaska Fairbanks

“[T]here are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns—the ones we don’t know we don’t know.”

– Donald Rumsfeld (Feb 12, 2002)

“He who knows not, and knows that he knows not, can be taught; teach him.”

– Anonymous (quoted in Maurus 1987)

One of the major motivations driving the field of documentary linguistics is the need to create a lasting record of language that can be (re)used by both speakers and linguists. However, the mere act of language documentation does not guarantee that the products of documentation are accessible. This retrieval problem can result in a false belief that a language has been adequately documented—what I refer to as an unknown unknown. This paper illustrates unknown unknowns with examples drawn from the field of place names documentation, touching briefly on unknown unknowns in other areas of language documentation. The paper concludes with some suggestions as to how to mitigate against the retrieval problem.

1. INTRODUCTION. As the field of documentary linguistics continues to expand, more and more attention has been paid to the nature of language documentation itself. Just what does it mean to document a language? What counts as language documentation? For example, some would assert that the traditional Boasian trilogy of grammar, text, and lexicon is not sufficient to capture the full range of linguistic competence of a native speaker, preferring to categorize the former as description rather than documentation. Indeed, one of the seminal papers in this area famously draws a distinction between these two methodologies (Himmelman 1998). These debates will no doubt continue, and competing views will continue to be aired in publications such as this one. But regardless of one’s view as to what counts as language documentation, there remains another lingering problem. Namely, how do we know that a language has been adequately documented?

While this may seem a trivial question, I would claim that it lies at the heart of our field. If nothing else, most funding applications for endangered language work require the applicant to justify the need for the proposed documentation, based at least in part on an assessment of the adequacy of existing documentation. Yet in practice, determining whether a language has been adequately documented turns out to be far from trivial. Some standard answers include: “language X has a dictionary and/or grammar”; “linguist Y worked on language X”; and “funding agency Y supported a project on language X.” But while these answers are readily operationalizable, they aren’t very satisfactory because they don’t tell us anything about the state of the actual documentation. A better answer is that a language

should be considered adequately documented when primary data—recordings, field notes, text transcriptions, lexical databases, etc.—are available and accessible.

Contrary to popular assumptions about language archiving, it can be surprisingly difficult to access the results of language documentation. I refer to this difficulty as the retrieval problem. The retrieval problem is a multi-faceted one involving several different types of barriers to retrieving language documentation. Most often, these barriers arise due to failures to adequately archive or curate language documentation. For example, linguists may fail to properly archive data in the first place. Archives may fail to maintain appropriate linkages between various parts of the documentation. Preservation formats may become inaccessible. A failure at any one of these points can result in an insurmountable barrier to the future retrieval of that documentation.

In the case of recent language documentation, the retrieval problem may arise deliberately. For example, language documenters intending to compile the results of a documentation project for a doctoral dissertation may place a sunset clause on archival deposits, restricting access for a defined period of time. In the case of older materials, the retrieval problem may be due to physical constraints. Much early language documentation sits in archives waiting to be processed and discovered. But these are known and surmountable hurdles.

A less trivial version of the retrieval problem occurs when the barriers to access are not so easy to overcome. It is here that the Rumsfeldian terminology—famously defended by Geoffrey Pullum—becomes useful.¹ The most straightforward cases are those languages for which we know little or no documentation exists. These can be referred to as known unknowns. They are the blank spots on the map. When I first began a project to document Western Pantar (ISO 639-3 code: lev), a Papuan outlier language of southern Indonesia, the only extant documentation consisted of a 117-item wordlist gathered in the provincial capital. This was a clear case of an inadequately documented language. It was a known unknown: a language about which we know we don’t know. The known unknowns are the major focus of documentary linguistics today. They are the tasks laid out for us. They are the motivation for developing the explicit discovery strategies for endangered languages discussed here in this journal. We don’t know enough about these languages. And since we know that we don’t know enough, we can develop strategies for documenting these languages.

A more insidious problem involves *unknown unknowns*. I argue here that there are indeed unknown unknowns in the field of language documentation. These unknown unknowns arise most commonly when we mistakenly believe a language, or a certain aspect of a language, to be adequately documented and archived. Extended fieldwork with a particular language or language family can lead to the impression of comprehensive documentation, when in practice such documentation may not be readily retrievable. If we think that a language is adequately documented when it is in fact not, in effect we don’t know that we don’t have adequate documentation. The language is an unknown unknown.

In this paper I focus on a specific type of unknown unknown which I encountered recently in the course of my own research. This case study serves to illustrate the ways

¹ Pullum defends Rumsfeld’s language in a blog entry dated 12/2/2003: <http://itre.cis.upenn.edu/~myl/language/og/archives/000182.html>

in which unknown unknowns arise in language documentation. Following this example, I discuss briefly some other types of unknown unknowns. Finally, I conclude with some suggestions for overcoming the problem of unknown unknowns in language documentation.

2. A CASE OF “UNKNOWN UNKNOWNs”: ALASKA NATIVE PLACE NAMES. I first became aware of the problem of unknown unknowns while working on a project to revise and update the *Indigenous Peoples and Languages of Alaska* map, first published in 1974 and last revised in 1982. As part of the revision process, we decided to include indigenous place names for some 270 Native villages and a few geographic features, including rivers, lakes, and islands (see Figure 1).



FIGURE 1: Detail of revised Alaska Native language map (Krauss et al. 2011)

This was in no way intended to be a comprehensive place names inventory for Alaska. Rather, the project had two primary objectives: (i) to revise some language names and spellings which had acquired pejorative senses, and (ii) to create a digital data set for the language boundaries, to be published in the Alaska State Geospatial Data Clearinghouse.² The decision to include place names was something of an afterthought. The plan for the revised Alaska map was to include just the prominent names for existing and historical villages and major hydrographic features across the 20 Native languages of Alaska, not obscure names for minor geographic features. These were just the “easy” names.

By way of comparison, I calculate that a comprehensive inventory of Alaska Native place names would number upwards of 39,000 names, far exceeding the 270 names included on the revised Alaska map. This figure is arrived at by extrapolating from the density of place names documented in Kari’s (2008) Ahtna Place Names Lists—the most com-

² <http://www.asgdc.state.ak.us>

prehensive published documentation of place names for any Alaskan language. The state of Alaska covers some 170 million hectares (170 m ha); thus the density of names on the revised map is just 1.5 per million hectares. (1 m ha is about the size of Connecticut or Cyprus.) In order to put this in perspective, consider that Kari (2008) lists over 2,100 names in an area of 9 m ha, a density of 234 names per million hectares (0.06 per square mile). This figure seems fairly typical for most regions of Alaska. Although Kingston (2009) calculates an extraordinarily high name density for King Island, a small island in Bering Strait in Western Alaska, the restriction to the island land mass may inflate the name density by concentrating names within a bounded region. Hunn (1994) reports name densities for Alaska Athabaskan regions slightly lower than those found in the Ahtna region, while the density among eastern Canadian Inuit is slightly higher. So it seems reasonable to take the published Ahtna density figure as representative of what might be expected across Alaska as a whole. This means that the expected true name density is nearly 150 times that which we intended to represent on the revised *Indigenous Peoples and Languages of Alaska* map.

My expectation was that most, if not all, of these 270 “easy” names would be found in existing published and archival documentation. We already knew the common English names for the locations; we had merely to look up those names in existing place names lists. This was not an unreasonable expectation, given the long history of documentation of Alaska Native languages. Alaska may be geographically remote, but it is far from linguistic *terra incognita*. Decent language maps appeared already in the 18th century (Shelikhov & Pierce 1981). Modern linguistic documentation increased rapidly in the 1970s as shifting educational policies first tolerated and then encouraged Native language use and development. Leading scholars of Alaska Native languages—especially Michael Krauss, James Kari, Steven Jacobson, and Jeff Leer at the Alaska Native Language Center—have been working to document these languages for 30, 40, even as many as 50 years. All of this documentation has been deposited with the Alaska Native Language Archive, to which I had unfettered access. Surely with such a wealth of comprehensive documentation, I would easily be able to locate the necessary references to fill out a simple map of 270 prominent Alaskan place names. How wrong I was!

In fact, it turns out that for most of the twenty Alaska Native languages, there is no readily available authoritative reference for place names. This remains the case, in spite of the fact that the modern documentation of Alaska Native languages extends back at least half a century to the arrival of Michael Krauss in Alaska in 1960. In particular, the Ahtna example cited above is exceptional. No similar published documentation exists for any Alaska Native language.

For every one of the twenty Alaska Native languages, at least some of the names I was looking for were not available in archival sources. In some cases multiple linguists and speakers had to be consulted. In other cases existing published and archival sources were in conflict, so it was not possible to determine from the documentation which name was correct. The correct names could only be retrieved by consulting with Native speakers or with linguists who had worked with the language. This may sound like standard fieldwork, but bear in mind that this was not supposed to be fieldwork; these names were “supposed” to have been documented already—or at least I had assumed them to be. These were common names which should have been relatively easy to retrieve from archival sources. In some cases even existing published sources turned out to be erroneous. For example, the name

for the village of Deering appears incorrectly as *Ipnaitchiaq*, rather than the correct name, *Ipnatchiaq*, on Krauss’s widely circulated 1995 map of the Inuit-Yupik-Aleut languages. Lacking available sources of documentation, what had begun as a simple reference project turned into a primary research project.

One further peculiarity of indigenous place names research is worthy of note here. During the course of my research, it was suggested that I consult published gazetteers such as Orth’s (1967) *Dictionary of Alaska Place Names* and Bright’s (2004) *Native American Place Names of the United States*. But these sources are inadequate for a place names project based on indigenous names. Orth (1967) only includes English names, noting indigenous etymologies where relevant or known. Bright (2004) is limited to English names of Native origin, taking officially recognized names as its point of departure. This can be useful for names that are (more or less) transparently of indigenous origin. For example, the Gwich’in village known in English as Venetie very clearly derives from the Gwich’in name *Vijhtaii*; hence, the latter can be found correctly transcribed underneath the entry for Venetie in Bright (2004). In contrast, the nearby settlement known in English as Arctic Village is referred to in Gwich’in as *Vashraqi K’qq*, literally ‘bank creek.’ In this case the indigenous name is unrelated to the English name, thus Arctic Village is not found in Bright (2004). To my knowledge, there exists no comprehensive reference for Alaska Native place names, in spite of many years of documentation efforts.

3. WHERE DID ALL THE PLACE NAMES GO? So where did all the names go? That is, how could we think that documentation existed when it in fact did not? There are a number of reasons, some unique to place names research and others more general. Among the former is that most place names work is not actually published, or is published only in “gray” literature, such as agency reports which are not widely circulated. Like other forms of lexical documentation, place name lists may exist for years or decades in draft form. Place name reports are often tentative and don’t represent the final product. In turn, this work tends not to be archived. Thus, while we might know that the research has been done, the results may be difficult or impossible to locate.

Another problem concerns the interdisciplinary nature of place names. This means that, often, place names documentation is undertaken by non-linguists, sometimes—but not always—in partnership with a linguist. In Alaska, place names documentation is often undertaken under the aegis of state and federal agencies such as the Alaska Department of Fish and Game or the National Park Service. These agencies have a legitimate interest in documenting the cultural history of the regions they oversee, and place names are an obvious part of that cultural history. But because they do not employ linguists, these agencies may not view such work as language documentation, nor may they understand the larger issues of endangered language documentation. This results in great variation in transcription styles and linguistic accuracy, complicating the task of resolving discrepancies between sources.

A related issue is that much place names research is limited to a specific region, language, speaker, or jurisdiction, leading to large gaps in data sets. Rather than a single lexical data set, we often find myriad overlapping data sets representing names in a defined administrative region, or names known by a certain speaker. For example, a recent publication of Alutiiq names focuses on those names known by two particular native speaker

experts (O'Leary 2009). These lists reveal much about traditional knowledge, but they are by design not comprehensive. The existence of published documentation may easily lead to the false belief that comprehensive documentation exists.

While the issues mentioned above may be (at least in part) unique to place names research, there are at least two issues of broader relevance which may lead us to believe that documentation exists when in fact it does not. The first concerns the project-oriented nature of much place names research. Like much linguistic research today, place names research tends to be embedded within much larger projects. Within the larger context, place names research may not have a high priority. Lacking a developed framework for distribution and archiving, such work is often abandoned at the end of the project, with no attempt made to archive the results.

The second problem has to do with the nature of the data themselves. Like much linguistic data, place names data are fragile and easily lost, so that even where documentation has been collected, it may no longer exist as a result of decay of the original data. I discuss each of these issues in turn in the following two subsections.

3.1 PROJECT-ORIENTED PLACE NAMES RESEARCH. Place names research and documentation is often embedded within larger research projects focused more broadly on cultural documentation. Often this type of research is described under the rubric of traditional ecological knowledge (TEK). The goal in such projects is not merely to collect a list of names, but also to research the indigenous knowledge associated with those names. Researchers may be more interested in learning the stories associated with places and the ways those places connect people to the land. The emphases on discovering what types of foods were harvested at a certain location, or how a location served as a trail marker, may take precedence over linguistic documentation. In many ways, this tendency reflects the indexical nature of place names. For most non-linguist researchers and community members, names are signs: they immediately point to a location, a story, or a memory. The linguistic content of the name is easily overlooked in the quest for the deeper meaning which lies behind the name. So it is not surprising that the names documentation itself may be lost in the context of a TEK project.

It is not too difficult to find examples where this has occurred, but I am hesitant to single out just one project here. I maintain that the loss of place names documentation within a larger TEK project is not due to a shortcoming of any one particular project, but is rather a result of our tendency to believe that documentation has been completed when in fact it has not. To illustrate this phenomenon, consider the case of one recent project which documented cultural geography, biogeography, and traditional ecological knowledge in a certain linguistic region. I will refrain from explicitly identifying the project. This project was exemplary in both its community-based approach and its interdisciplinary collaboration. Conceived, designed, and led by community members, it was a prime example of what Czakowska-Higgins has described as community-based research, in that the research was conducted "for, with, and by" the community (2009:24). The project integrated teams of cultural experts, anthropologists, linguists, botanists, and geographers in a true spirit of interdisciplinary collaboration.

After more than five years of funding, the primary public outcome of the project is a media-rich web site. Visitors to the web site navigate aerial images with marked locations.

Zooming in reveals more detail and additional locations. Clicking on a location reveals a page containing a close-up image of the location and some additional information, including a transcription of the name of the location; a literal translation of the name; a short description; and the name of the recorder. In addition, users can click to hear a recording of the name being pronounced by a native speaker. Notably absent from this list of information is identification of the geographic locations. This can be approximated by examining the locations of dots on the aerial images, but these locations are imprecise. More troubling than the lack of precision is that there is no way to retrieve a simple listing of names and locations. The link between the names and locations relies on a proprietary software tool (Adobe Flash) which maps particular pixels on the static aerial image to a particular name. There is no link between the name and a definite geographic location, such as that defined using latitude and longitude coordinates.

Providing precise geographic information need not be an onerous task. Many recent projects have collected location data directly in the field using geographic positioning systems (GPS). In such cases, location data can be provided simply by recording the latitude and longitude (and corresponding coordinate system) from the GPS device. Legacy data recorded on paper maps can be scanned and then geo-referenced using Geographic Information Systems (GIS) software such as ArcGIS or GRASS. This process assigns a geographic location to each point on the scanned image of the map, making it possible to determine precise geographic coordinates for each place name represented on the map. In contrast, simply marking a location on a digital image using Flash or another graphics tool merely locates a name on the image, without yielding a geographic location. (For more information on georeferencing, see Chapman & Wiczorek 2006.)

An additional problem with the above-mentioned place names project is that the data are stored on a web server rather than in an archive. This means that continuity of the data will be jeopardized with each future migration of the hardware and software which support the web server. Many language documentation project web sites have become inaccessible due to technology migration (cf. Holton 2010). The loss of information on a project web site would not be such a bad thing if the data were archived and accessible elsewhere. But in the particular case described in the preceding paragraph—and in very many other projects as well—the data are available nowhere else. In the case described here, the place name data have never been published, nor to my knowledge have they ever been deposited with a public archive. It should be noted here that depositing with an archive is not the same as making the data freely available. Some projects may wish to deposit documentation with an archive while maintaining access restrictions, but even this has not been done in the case described above.

Instead, the existence of an attractive project web site gives the impression that the documentation has been completed. Even though users may not be able to access the documentation directly on the web site, they may easily get the impression that the data are safely stored away in another location. Yet the underlying data on which the project web site was based may have already disappeared. In any case, they are not accessible. They are unknown unknowns.

3.2 FRAGILE DATA: LINKING NAMES AND LOCATIONS. An additional problem with place names research is that the data tend to be fragile and easily lost. Often names are

recorded directly on a paper map, so that the link between name and location relies on the physical positioning of a label. As Figure 2 illustrates, labels can easily become separated from the paper map as they age.



FIGURE 2: Name labels separating from a place names map (courtesy Alaska Native Language Archive)

Even worse, names may be indicated not with a textual indicator, but with a code or number on a paper map. This code is then linked to an external document which notes both the number and the corresponding name. This approach has the advantage of allowing much additional information about the name to be recorded, as the external document is not constrained by physical space on the map. However, once the key or index document is separated from the map, the data become degraded in that the names—though in this case not the locations—are lost. At the Alaska Native Language Archive, approximately half of the place names maps that are marked up in this way lack an accompanying key.

The most obvious solution to this problem involves digitization of the paper maps to create a GIS database, as discussed in section 3.1. This approach effectively associates geographic coordinates (latitude and longitude) to each place name.³ But digitization doesn't necessarily solve the problem. In particular, attempts to sort out this problem using GIS technologies may also fail. In Alaska most GIS-based place names research is done by government agencies, and the focus tends to be on the geographic rather than the linguistic data. This results in database structures that emphasize locations over names. Due to a lack of understanding of Native orthographies and symbols, database creators often use numbered indices rather than names in the database tables. The example in Table 1 is typical.

³ While place names are typically modeled as point data in GIS databases, it is also possible to employ more complex topologies, including lines and polygons.

OBJECT ID	PLACE	Longitude	Latitude
847	1	-151.6036954	65.1809537
848	1	-151.5456042	65.1820155
1054	1	-146.7306593	63.0148440
1529	1	-152.8084079	63.7119276
1566	1	-151.3872265	62.9260095
4	2	-153.8660532	62.1637877
443	2	-152.2580874	61.5283609
846	2	-151.7272196	65.1016275
1192	2	-148.1951629	63.7526432
590	3	-151.0279645	63.0926017
591	3	-150.9517907	63.0929129
855	4	-150.9198858	64.9086955
1556	4	-152.2828937	65.1673213

TABLE 1: Excerpt from a GIS database table with non-unique name identifiers in the PLACE column

In this database table, each place name is precisely located with latitude and longitude coordinates. Like the paper maps discussed above, the names are indexed using numbers. In this case the relevant numbers are found in the column labeled PLACE. Presumably, these numbers are indexed to a published report. However, closer examination of the database shows a many-to-one relationship between the locations and the names. Here, five distinct locations are associated with a name indexed by the number 1, as is evident in the five occurrences of the number in the place column of the table. This reflects the fact that the region of concern involved five distinct languages with overlapping territories, and each language was given its own index, starting with 1. Since the languages are not identified in the table, it is not possible to determine which number 1 belongs with which language. So even if we had access to the printed report, it would not actually be possible to identify which location goes with which name.

Knowing that place names have been documented often leads us to the conclusion that the resulting place names data have been adequately preserved. What these examples illustrate is that the mere fact that documentation has been undertaken does not ensure that the documentation will endure. Place names data—digital or not—are fragile. Unless steps are taken to properly archive the data, they can easily be lost. When we assume that the mere act of documentation implies the existence of documentation, place names data become unknown unknowns.

4. BEYOND PLACE NAMES. Those familiar with other parts of the world outside Alaska might argue that place names hold a particular significance which results in a reticence toward publication. I don’t deny this, though I point out that within the Alaskan context the situation (at least today) is much different. Indigenous land claims were largely settled with the Alaska Native Claims Settlement Act in 1971, and—contra Australia and Canada—place names play no major role in current territorial disputes in Alaska. That said, Alaskans

may have other reasons for not publicizing place names. In particular, ongoing legal conflicts relating to the appropriation of material and non-material culture may lead to a reluctance to share information about place names. But such attempts to justify the non-public nature of language documentation do nothing to change the end result. Documentation which is not accessible is of little use to the overall progress of language documentation. Further, when such documentation is not archived properly, it may just as well not exist.

More significantly, the existence of unknown unknowns is by no means restricted to place names studies. In this paper I have focused primarily on place names because that is the subject of my current research, and because that is where I first became aware of the retrieval problem. However, similar situations can be identified with many types of linguistic data. One need look no further than the linguist’s field notebook—the very foundation of linguistic fieldwork and the documentary enterprise. Filled with observations and varying levels of meta-analysis, the field notebook spans the gulf between documentation and description. From an archivist’s point of view, field notebooks are a treasure revealing the linguist’s thought processes as the language is revealed to her. Scribbles and corrections and mistakes trace the path of the linguist’s experience.

From the analyst’s point of view, field notes reveal the context in which the data were collected, thereby improving the quality of secondary analyses. To take a simple example, field notes may reveal which sounds a linguist was struggling to transcribe, permitting a better exegesis of transcribed texts. Yet, while we as a discipline generally acknowledge the importance of field notes, linguists often fail to deposit them with archives. In language archives such as the Alaska Native Language Archive, it is common today for linguists to archive typed versions of field notes which have been highly processed and redacted. Thus, in many cases we may believe that field notes exist, when in fact they cannot be retrieved. They, too, are unknown unknowns.

While the reluctance to archive original field notes is not a new problem, it may be exacerbated by today’s emphasis on data processing. Recordings need to be transcribed and time-aligned. Lexical data need to be Shoeboxed and glossed. The tendency, then, is to archive only this semi-processed documentation, which may have been affected by a significant amount of normalization and processing. For example, while lexical documentation may be collected by hand with pencil and paper, these data are often typed and entered into a lexical database prior to archiving. In the process, the original collector’s annotations and close transcriptions may be irretrievably lost. In turn, in the process of moving from database to published dictionary, the original database may be discarded or forgotten. This is the case, for example, with the *Koyukon Athabaskan Dictionary* (Jones & Jetté 2000). While the published document can be found in the Alaska Native Language Archive, the lexical database from which it was derived has not been deposited. We tend to assume that some larger lexical database backs up the published dictionary somewhere. But how can we be sure? Lexical data and lexical databases are very often unknown unknowns.

In fact, once we acknowledge the existence of unknown unknowns in one domain of language documentation, the possibility of unknown unknowns in other domains becomes difficult to deny. Published maps of language boundaries reflect prior mapping of linguistic isoglosses, yet the original isogloss mapping may not be archived or accessible. Descriptive grammars based on corpora imply the existence and accessibility of the associated corpora, when in fact they may not be accessible. It turns out that, in this respect, place names

documentation is not so atypical after all. Instead, the case of place names documentation nicely illustrates the problem of unknown unknowns in language documentation. Just as a pretty (but selective) place names map can lead us to believe that complete documentation of place names exists, a published grammar gives the impression of complete documentation of grammatical structures—when in fact the underlying documentation may be inaccessible. But inaccessible documentation is, for all practical purposes, no documentation at all.

5. SOLVING THE RETRIEVAL PROBLEM. There remains the question of how we can mitigate against the retrieval problem. I maintain that the most important step we can take is to acknowledge the existence of unknown unknowns. Once we acknowledge that language data may be inaccessible, we can take the necessary steps to address the issue. In the case of the place names documentation described in this article, we have already had some success in identifying keys to maps and resolving ambiguities in name databases. We have been able to achieve this success in large part because many of the original data collectors are still alive and available to assist with these efforts. As more time passes, this window of opportunity will surely close, rendering the data much less useful, if not useless. But while this window remains open, we should certainly work to consolidate existing documentation.

At the same time, we should also be more adamant about encouraging the practice of archiving primary data. We cannot force people to make documentation publicly available, but we can certainly encourage this as best practice. Funding agencies hold the biggest sway in this regard, and they could do a better job of ensuring that research they fund is properly archived and accessible. Data management plans should include a clear statement of precisely what materials will and will not be archived. Further, funding agencies could make additional efforts to support work on languages presumed to be documented, but for which the relevant data remain inaccessible. That which is documented but inaccessible should be considered undocumented.

But in the end, it is up to linguistics as a discipline to discover these unknown unknowns and reclassify them either as *known knowns* (if possible) or as *known unknowns*. Knowing what we don't know allows us to make progress in filling gaps in documentation. To do this, we need to know just what we *do* know. We should guard against the temptation to assume that we could, with a little effort, uncover the documentation. The ability to retrieve material from a recognized language archive should be the *sine qua non* supporting the assertion that a language has been adequately documented.

REFERENCES

- Bright, William. 2004. *Native American Place Names of the United States*. Norman: University of Oklahoma Press.
- Chapman, Arthur D. & John Wieczorek (eds.). 2006. *Guide to Best Practices for Georeferencing*. Copenhagen: Global Biodiversity Information Facility.
- Czaykowska-Higgins, Ewa. 2009. Research Models, Community Engagement, and Linguistic Fieldwork: Reflections on Working within Canadian Indigenous communities.

- Language Documentation & Conservation* 3(1). 15-50.
- Holton, Gary. 2010. The role of information technology in supporting small and endangered languages. In Peter K. Austin & Julia Sallabank (eds.), *The Cambridge Handbook of Endangered Languages*, 371-399. Cambridge & New York: Cambridge University Press.
- Hunn, Eugene. 1994. Place-names, population density, and the magic number 500. *Current Anthropology* 35(1). 81-85.
- Jones, Eliza & Jules Jetté. 2000. *Koyukon Athabaskan Dictionary*. Fairbanks: Alaska Native Language Center.
- Kari, James. 2008. *Ahtna Place Names Lists*, 2nd ed. rev. Fairbanks: Alaska Native Language Center.
- Kingston, Deanna. 2009. “Almost every place, every rock, had a name”: A consideration of place-name density on King Island, Alaska. *AlterNative* 5(1). 6-25.
- Krauss, Michael E. 1974. *Native Peoples and Languages of Alaska*. [Map] Fairbanks: Alaska Native Language Center. Revised 1982.
- Krauss, Michael E. 1995. *Inuit Nunait*. [Map] Fairbanks: Alaska Native Language Center.
- Krauss, Michael E., Gary Holton, Jim Kerr, & Colin T. West. 2011. *Indigenous Peoples and Languages of Alaska*. [Map] Fairbanks & Anchorage: Alaska Native Language Center & Institute of Social and Economic Research.
- Maurus, J. 1987. *Wisdom Stories*. Bombay: Saint Paul Society.
- O’Leary, Matthew B. 2009. Two Koniag place-name lists. In Kenneth L. Pratt (ed.), *Chasing the Dark*, 98-106. Anchorage: Bureau of Indian Affairs (BIA) ANCSA Office.
- Orth, Donald. 1967. *Dictionary of Alaska Place Names*. Washington: U.S. Government Printing Office.
- Shelikhov, Grigorii I. & Richard A. Pierce. 1981. *A Voyage to America 1783–1786*. Kingston, Ontario: Limestone Press.

Gary Holton
gmholton@alaska.edu