

## Documenting Sociolinguistic Variation in Lesser-studied Indigenous Communities: Challenges and Practical Solutions

John Mansfield and James Stanford  
*University of Melbourne and Dartmouth College*

Documenting sociolinguistic variation in lesser-studied languages presents methodological challenges, but also offers important research opportunities. In this paper we examine three key methodological challenges commonly faced by researchers who are outsiders to the community. We then present practical solutions for successful variationist research on indigenous languages and meaningful partnerships with local communities. In particular, we draw insights from our research with Australian languages and indigenous languages of rural China. We also highlight reasons why such lesser-studied languages are crucial to the further advancement of sociolinguistic theory, arguing that the value of the research justifies the effort needed to overcome the methodological difficulty. We find that the challenges of sociolinguistics in these communities sometimes make standard variationist methods untenable, but the methodological solutions we propose can lead to valuable results and community relationships.

**1. INTRODUCTION: FOOTNOTE OR FOCUS?** Documenting sociolinguistic variation in lesser-studied indigenous languages, especially where they are spoken in rural areas or remote villages, requires methodological adaptation both in terms of data collection and in terms of analytical limitations. It is not easy to investigate variation as part of a language documentation project, causing some field linguists to set it aside as an intractable puzzle. As a result, many intriguing instances of variation end up as a footnote or parenthetical comment in descriptive grammars, rather than receiving the focused attention they deserve. For sociolinguists, there are acute challenges as well. The second author recalls sitting in a variationist sociolinguistics class in a U.S. graduate program, thinking about his prior experiences in rural China, and then asking himself “How in the world can I apply these variationist methods in an indigenous minority village in China?”<sup>1</sup>

The particular challenges of studying variation in indigenous communities have received little dedicated discussion in the literature. In this paper we tackle three key challenges, and argue that they are neither fatal nor insurmountable. Drawing from our own research, we explore practical methods for successful application of variationist sociolinguistics to lesser-studied indigenous languages. In particular, we focus on fieldwork issues in two settings: Aboriginal Australia and indigenous minority communities of China. We find that these two research settings exemplify:

---

<sup>1</sup> The authors would like to thank all the indigenous communities with whom we have done sociolinguistic fieldwork, especially the Murrinhpatha and Sui communities highlighted in this article. The hospitality we have received has been personally and scientifically enriching, and formed the basis of many observations recorded here. We would also like to thank the volume editors and reviewers for their valuable comments on earlier versions of this article.

- (i) languages with relatively limited prior literature or scholarly knowledge of variables and language structure;
- (ii) languages that are very distinct from the majority language (i.e. not dialects of a local standard language);
- (iii) languages that are structurally very distinct from Western languages;
- (iv) communities that have historically been overlooked by variationist sociolinguistics;
- (v) communities that are culturally very distinct from the West;
- (vi) communities that are relatively small and kinship-oriented, and may have socio-political reasons to distrust Westerners.

For convenience in this paper, we will refer to these communities as “indigenous communities.” Naturally, the term “indigenous” can also be applied to many other communities that differ from this in various ways. By making a list, we are not implying that research in other communities is less valuable or less challenging, nor that “indigenous” can only apply to communities with these limited characteristics. This list simply defines our focus in the present paper.

In this paper, we examine why the characteristics we have listed in (i-vi) can be hurdles for indigenous sociolinguistics in general, but especially for researchers (such as the co-authors) who are outsiders in these field sites. We applaud and strongly support the work of researchers who are “cultural insiders” and members of the lesser-studied indigenous communities that they are researching. But in this paper, although we address researchers of any cultural background, we pay particular attention (in sections 2 and 3) to the specific challenges faced by cultural outsiders. In disciplinary terms, this paper is intended for two audiences: sociolinguistic variationists, who may shy away from studying indigenous languages due to perceived challenges of data collection or analysis; and documentary field linguists, who may not have considered sociolinguistic variation to be part of their descriptive task. We explore why indigenous settings present hurdles for variationist research, but we also argue that they provide important opportunities to shake our prior assumptions and test existing sociolinguistic principles.

All things being equal, it seems reasonable to suppose that the greater the distance (sociocultural and linguistic) between a language community and the traditional sociolinguistic settings (large, industrialized, Western field sites like New York, Philadelphia, Norwich, etc.), the more likely it is that the research in the lesser-studied community will pose theoretical challenges for existing principles, along with surprising new insights and new research questions to explore. Following the Labovian tradition of naming principles (Labov, 1966/1972, 1994, 2001, 2010), we call this the *Principle of Sociolinguistic Distance*. As a general rule of thumb, it motivates our research and drives us to keep exploring exciting new field sites and languages. The comparison of data separated by such great sociolinguistic distance promises to inform more general theories of sociolinguistic variation: what types of linguistic features are used for what types of social functions in diverse language types, and what types of social meaning are attached to such variables in diverse societies? At present, sociolinguistic theory draws on data that tends to be limited in both linguistic and social dimensions.

Most standard guides to fieldwork methods (e.g. Bowern, 2015; Chelliah & de Reuse, 2010) do not include variationist sociolinguistic analysis as a major topic,

though there are a few pages dedicated to variation in Sakel & Everett (2012, pp. 20–24, 100). Conversely, most guides to sociolinguistic data collection (e.g., Feagin, 2002; Tagliamonte, 2006) do not focus on indigenous settings, although Mallinson, Childs & Van Herk (2013) includes two vignettes on this topic. Of course, this situation merely reflects the expectations of the two respective fields: indigenous language documentation and sociolinguistics are widely assumed to be disjoint domains. For example, a survey of the most recent decade (2006–2015) of the journal *Language Variation and Change* shows that 86% of the published articles were primarily based on data from European languages, and only 7% of the articles focused on indigenous minority languages (Stanford 2016). For each annual volume of about 15 articles in *Language Variation and Change*, there was an average of just one article on an indigenous minority language. Nagy & Meyerhoff (2008) and Smakman (2015) report similar trends in other sociolinguistic journals and handbooks.

Fortunately, this situation is changing for the better. In recent years, field linguists have expressed increasing interest in applying variationist methods to endangered languages, while increasing numbers of variationist sociolinguists are working on lesser-studied languages. Nicholas Evans' Wellsprings of Linguistic Diversity project at Australian National University, for example, is a seven-year project dedicated explicitly to exploring sociolinguistic variation in undocumented or underdocumented indigenous languages of the Australia/Pacific region. Another recent example is the Linguistic Society of America workshop "Documenting Variation in Endangered Languages" (Hildebrandt, Silva & Jany, 2016), which led to the present journal issue. Similarly, Meyerhoff (in press) shows the importance of creating a "linguistic symphony" of multiple variables and blending variationist approaches with endangered language documentation and description. Sankoff (1980) blends quantitative variationist research with language description in her work on Papuan languages, and a sample of other early work on variation in lesser-studied languages includes Miller (1965), Sutton (1978), Foley (1980), Dorian (1981), Hill & Hill (1986), Smith & Johnson (1986) and Hill (1996). In recent years an increasing number of variationist studies in indigenous languages have appeared. Examples include Hildebrandt (2003, 2005), Stanford (2007), Jones & Meakins (2013), Mansfield (2014), Meyerhoff (2015b), Stewart (2015), Suokhrie (2016), as well as Clarke (2009), Meyerhoff (2009), Nagy (2009), O'Shannessy (2009), and the other chapters in the edited volume on this topic (Stanford and Preston 2009). Further signs of progress in variationist study of indigenous languages are visible in the growing *NWAV Asia-Pacific* conference series and the journal *Asia-Pacific Language Variation*, both of which emphasize lesser-studied languages in non-European settings. These publications and conferences are helping to widen the typological scope of quantitative research on language variation, but they still represent only the tip of the iceberg of linguistic and social structures.

### *The present study*

This paper examines what we believe are the three most challenging aspects of documenting variation in lesser-studied indigenous languages:<sup>2</sup>

---

<sup>2</sup> For readers looking for basic information about practical quantitative and field methods for variationist sociolinguistics in general, we recommend Meyerhoff (2015a), Tagliamonte (2006), Thomas

- (1) being a cultural outsider,
- (2) identifying variables with limited prior knowledge, and
- (3) conducting sociolinguistic analysis with limited data.

The first affects fieldwork practice in itself: the difficulties an outsider researcher faces in attempting to unravel local intersections of language and culture, in which typical variationist sociolinguistic methods may sit uneasily. Sociolinguistic research always inserts the researcher to some degree as a participant in social interactions, power dynamics, and cross-cultural interventions (Eckert, 1989; Milroy, 1987). In the case of sociolinguistics in indigenous communities, the cross-cultural gaps may be particularly wide. This is not a matter of there being anything intrinsically “exotic” about the indigenous community, but rather the discrepancy between some indigenous cultural norms and the expectations implied in standard variationist methodology. As for the second challenge, we note that identifying sociolinguistic variables is not an issue for cosmopolitan languages where major patterns of variation are often already known from prior work. But for indigenous languages, the morphosyntax and phonology tend to lack research attention, much less the relevant sociolinguistic variables. The fieldworker must balance grammatical analysis with identifying sociolinguistic variables, potentially starting from scratch on both tasks. Finally, with respect to the challenge of limited data, variationist approaches typically presume a large dataset from which to derive analyses of variability, but there are various impediments to reaching this requirement in indigenous language research.

These challenges of indigenous sociolinguistics may be intimidating for both outside researchers and local community members to various degrees, but the experience can be rewarding on both sides. Since Labov’s (1966/1972) foundational variationist work, scholars have amassed 50 years of research on the social life of language, especially in urban, Western societies. A nuanced and rigorous body of work has accreted on the interaction of social structure and linguistic structure (e.g. Labov, 1994, 2001). But the same topics remain only very marginally studied in indigenous communities. Studies of indigenous communities are needed to test the limits of sociolinguistic theory, to reveal which parts of it are merely artifacts of the Western metropolis, and which are truly general principles (Stanford, 2016; Nagy, 2009; Meyerhoff, in press; Stanford & Preston, 2009). Such studies can also build a “sociolinguistic typology”, for which at present we have broad hypotheses (e.g. Trudgill, 2011). But with proper empirical sources, this work might be developed to parallel grammatical typology (e.g. Dryer & Haspelmath, 2005).

As for language documentation, indigenous sociolinguistics also promises to enrich the field of study by broadening its scope to cover the social dimension of language, alongside the formal or grammatical. But social dimensions should not be too sharply divided from the grammatical (Meyerhoff, in press). Many linguistic theories view the language faculty as a dynamic, usage-based system (e.g. Bybee, 2006; Goldberg, 2006; Hopper, 1987; Pierrehumbert, 2001; Tomasello, 2003). This implies that interactional and social needs, the drivers of language use, are at a deeper level inseparable from the grammar (Enfield, 2002). Therefore, if we are to further our understanding of grammatical systems at a global scale, we must complement this by furthering our understanding of the sociolinguistic dynamics that drive the evolution of those systems.

---

(2011), Mallinson, Childs & Van Herk (2013), Di Paolo & Yaeger-Dror (2011) and other existing excellent sources.

These exciting possibilities for continuing progress in variationist research with lesser-studied indigenous communities need not be blocked or limited due to methodological challenges. In the following, we outline three challenges that we have faced in our own research and then suggest practical solutions for each one.

**2. THE CHALLENGE OF BEING A CULTURAL OUTSIDER.** Being a cultural outsider may restrict access to information and the diverse types of speech and speakers needed for a thorough understanding of language variation. This issue has been recognized in variationist sociolinguistic research on highly developed industrialized societies (e.g., Labov 1972a; Trudgill 1972), but presents special challenges for indigenous sociolinguistics. We first lay out the nature of this challenge, then discuss ways in which it can be solved by deeper engagement with cultural insiders, and preferably, active participation by native speakers and community members as researchers (Smith 1999). For argument's sake, we first lay out a "worst case scenario", portraying the full extent of challenges that may be faced by outsider researchers.

Research in indigenous communities begins by engaging with a community, and seeking local approval for research. For outsider researchers, often this will happen through an intermediary – perhaps another researcher who has had some contact with the community, or perhaps the prospective sociolinguist has already researched a nearby area. It cannot be assumed that the community will immediately embrace sociolinguistic investigation, and the researcher must take a patient and open-minded approach. Building initial community relationships, understanding, and trust should be part of the research plan. If the community gives approval for research, and the logistical challenges of access and accommodation are resolved, the outsider researcher now faces another challenge: not speaking the language. Many indigenous languages, being documented little or not at all (Evans 2009; Krauss, 1992; Nettle & Romaine, 2000), are largely unknown to outsider researchers. Not knowing a language puts a sociolinguist at a great disadvantage. There may or may not be a contact language that the researcher can use with indigenous speakers, but even if there is, this constitutes a narrow channel of communication. The outsider researcher therefore faces serious difficulties in trying to explain objectives, propose research, discuss local language matters or indeed elicit speech. Moreover, a non-fluent researcher is not able to participate in everyday forms of language use. In the next section, we will return to the problem of being a non-speaker, with respect to the challenge of identifying sociolinguistic variables.

Outsider researchers may need significant time to learn local cultural norms, but in addition, these norms may be directly prohibitive of standard sociolinguistic field methods (Mansfield, 2014, p. 12ff.). Sociolinguists are interested in exploring language variation and change, but indigenous communities may not necessarily approve of data collection with the necessary range of speakers. Some indigenous communities emphasise multiplex, personal relationships: if the researcher is to engage with such a community, the relationship will likely not be just as a scientist, but also as a co-resident, a friend, and perhaps as a source of income. The community may even decide to relate to the researcher in terms of kinship or clan affiliation. In Australian Aboriginal communities, it is fairly standard for non-Aboriginal linguists to be inducted into kin relationships, and the clan, totem and sub-section groupings associated with them. The linguist is now affiliated to one particular part of the speech community, and has entered into classificatory kinship obligations. This often implies working with one kinship group, *rather than* with others (Mahood, 2012; Mansfield,

2014, pp. 16–17). This is the case among the Sui people of rural China as well. Given such personal commitments, sociolinguistic sampling then becomes a balancing act of meeting the expectations of the researchers' adoptive kin versus the expectations of a peer-reviewed journal.

There may be even stronger constraints against working with people according to gender, age or status. Young people's speech may be frowned upon, and some elderly people may be too revered to work with an outsider who has no local status. But the most common constraint of this type is against working with the opposite sex (Bower, 2015, p. 151). In various indigenous societies of Australia, Papua New Guinea, India, the Middle East and others, adult men and women do not mix unless they are close kin or romantically entwined. Therefore attempting to conduct elicitation or transcription sessions with a speaker of the opposite sex may cause acute embarrassment, or at worst rouse jealousy. This may limit research to just one half of the speech community, losing an important dimension of variation (e.g. Bradley, 1998; Haeri, 1996). While this is an issue for any linguistic field project, it is especially problematic for sociolinguistic research, where gender plays a crucial role in the analysis.

There may additionally be types of speech, certain stories, or topics that cannot be recorded. In Australian Aboriginal society, language and information are valuable commodities subject to restrictions of access (Michaels, 1986). In the West, language is often seen as an abstract entity with no particular owner, and this is one of the philosophical underpinnings of research that samples and compares the range of speech in a community. But this concept of "detached research" is at odds with the philosophy of language in some indigenous communities.

The scale of cultural challenges facing indigenous sociolinguistics is brought into particular focus by that methodological mainstay, "the sociolinguistic interview" (Labov 1972b; Tagliamonte 2006). This is a key method for collecting comparable language samples from a substantial number of speakers, but its standard application is very challenging in some indigenous communities. The classic sociolinguistic interview, alone in a quiet place with a single interviewee, may be inappropriate or unrealistic in a vibrant rural village setting, especially where households have large numbers of residents. Likewise, other variationist methods, such as having the interviewer "approximate the vernacular" in their own speech (Tagliamonte 2006: 46), may not be possible or appropriate for outsider researchers in these communities. Finally, some of the staple lines of questioning recommended for sociolinguistic interviews, such as school experiences and "danger of death" (Labov 2013), may be inappropriate in certain communities. For example, school may have been a place of linguistic and cultural oppression; danger of death may be a culturally restricted or uncomfortable topic in some cases.

But even more fundamental than these issues may be the very act of questioning. In Aboriginal Australia, using questions to obtain unknown information is not a normative mode of interaction (Bavin, 1992; Moses & Yallop, 2008; Walsh, 1997). Meeting a question with silence or a monosyllable is quite normal. This severely undermines the whole interview method, because it means that samples of talk cannot be reliably *elicited* at all – or at least not by direct questioning. It seems likely that such communities do have ways of inducing each other to speak about a subject, but these have not been extensively explored by ethnographic or linguistic researchers. Work in this direction could improve our sociolinguistic methodologies in these communities. In the meantime, samples of talk are presented to the linguist when the

speakers see fit. This is another way in which language and information is embedded in quite different ideologies in indigenous communities, and these do not sit easily with standard sociolinguistic methods.

Finally, in some indigenous communities there is socio-economic and political marginalisation that may be starkly at odds with the privileged position of many outsider researchers. Some settings may not have clean water and reliable nutrition; they may not have (or want) the common elements of industrialized societies like electricity, digital communications and cash income. They may be under threat from disease, pollution, invasion or conflict. They may have epidemics of addiction in the wake of colonisation and dispossession. The outsider researcher usually enters the picture as a privileged person who can access more money, resources and civil rights than most local people. In short there may be extreme power imbalances at work. Writing about Western field sites, Milroy (1987, p. 44ff.) recommends that the researcher engaging with a community should avoid association with the school or the church, as these are foci of social power and prestige, which may inhibit access to “the vernacular”. But the outsider researcher entering an indigenous community cannot avoid association with power and prestige; being an educated urbanite *embodies* a certain form of prestige. (This can also be the case in Western urban research, if the field site is one of socio-economic disadvantage.) In these situations, the local community may focus on how the linguist can improve their access to resources and external social networks, and may feel that displaying vernacular, low-status forms of speech is inappropriate in such a relationship.

As mentioned above, we have focused on some “worst case scenarios” in this section, i.e., some of the more difficult challenges that an outsider researcher may face as a cultural outsider. Naturally, there are also many research situations where the cultural gap is narrower. Regardless, we find that engagement with the community presents as much an opportunity as a challenge, and we discuss practical solutions to these challenges in the next section.

*Solution: Building relationships and working closely with cultural insiders*

We find that many of these cultural and linguistic challenges can be met by working closely with cultural insiders – ideally as equal collaborators in the research – and becoming immersed in the language and culture. The relationship between researcher and indigenous community must be honest, personal, and preferably long-term (Good 2012). The foundation of this is open discussion of research goals, as well as negotiability with regards to what those goals should be. Researchers should beware of entering the community with the attitude of an “expert”, as this may be an unwelcome figure in some settings. Far preferable is to approach the community as a learner, not a teacher. When we show the attitude of a learner, cultural insiders are more likely to participate in the research – and more likely to share an insider’s view of language in the community. If at all possible the researcher should attempt to learn the language, not just study it, as there are few better ways of showing respect for the community (Everett, 2001). Be ready to make laughable efforts at speaking an unfamiliar language, perhaps teased about incompetent pronunciations, and to show a degree of vulnerability.

To some extent, this approach overlaps with the field model of many descriptive field linguists. However, the sociocultural challenges are considerably more acute for variationist sociolinguistic researchers. Because of the type of data we are trying

to collect, we need access to a much larger number of speakers in the community, ideally from a wide range of ages and backgrounds. To understand sociolinguistic patterns, we need a deeper integration with the social fabric of the community – more than what is necessary for most traditional structural linguistic goals. As a result, we cannot depend on the traditional field linguist model of interacting with just a few speakers, or limiting our interactions to a narrow relationship of grammatical elicitation. We need extensive community engagement.

Our engagement with the community should be genuine, sustainable and mutually beneficial. In one Sui village in southwestern China (Guizhou province), locals repeatedly expressed a serious need for a road. The second author helped coordinate fundraising overseas to get the road started, and later the local government finished the project. When he returned to the village a year later, they welcomed him enthusiastically, referring to him with a Sui relationship term which means ‘in-law’. Suddenly, the entire village wanted to participate in his sociolinguistic research. Their previous attitude of slight distrust was replaced by a warm welcome. In the Australian town of Wadeye where the first author researched Murrinhpatha, the main request among locals was not for a typical ‘development’ goal, but rather that the author should help them record songs. Alongside his linguistic research, he began videoing song performances and distributing copies on DVDs (cf. Furlan, 2005). This apparently simple contribution generated great enthusiasm in the community, and soon led to an avalanche of recording requests.

The key to surmounting the challenge of cultural gaps is a collaborative approach with cultural intermediaries, some of whom may become research colleagues. Success of the fieldwork can be facilitated by these cultural insiders (Suwankhong and Liamputtong, 2015) engaging with the sociolinguistic enterprise. As described above, the outsider researcher’s relationship with local intermediaries is usually not characterised by scientific detachment. Rather, it is a multiplex relationship that in many cases involves sharing food, accommodation, and relational goals. This was discussed above as a challenge for gathering broad sociolinguistic samples in an indigenous community, but it can also be turned into a benefit. We may not be able to record dozens of speakers ourselves, but a close relationship with a local may allow us to record a few speakers in many different situations, using diverse speech styles with various interlocutors. Ideally, some local community members will already be engaged in linguistics research themselves. If not, it may be possible to train local people to do field recordings themselves, expanding the scope of the project far beyond what we could do. This also gives agency to cultural insiders and enables them to engage more fully with research goals. Previously unknown sociolinguistic variables of the indigenous language might even be explicitly pointed out to us.

Even more importantly, a relationship-based field project can benefit the community itself. Rather than “helicoptering” into the community, grabbing some data and leaving, the collaborative relationship-based approach means that the researcher becomes part of the community. For example, one of our rural field sites once suffered a massive fire that destroyed many houses. This crisis necessitated urgent communication with the community about the best ways to help as an outsider, and then taking appropriate steps to meet those responsibilities as a participant in community life. Likewise, from time to time outsider researchers may be humbled to find themselves in need as well, becoming the recipients of the community’s care.

Relationship-based fieldwork means that the community has a stake in the project, not just the researcher. As many field linguistics guides point out (e.g.,

Bowern 2015), researchers must be sure that their linguistic goals are in line with community goals. This is especially true for sociolinguistic research, as researchers need even greater community involvement to meet their research goals. As a result, sociolinguists may find themselves unexpectedly involved in language development issues, including dialect standardization, corpus planning, status planning, lexicography, and so on. While field linguists expect this kind of involvement, it may come as a surprise to a sociolinguist more accustomed to variationist work in traditional majority language communities.

Close collaboration with a cultural insider can increase the benefit to the community while also increasing the quantity of data collected and transcribed. Funding agencies now recognise the value of allocating research funds to pay community members as active collaborators in documentation, and cultural insiders may be able to substantially increase the number of recordings and transcriptions made, thanks to increasingly affordable and user-friendly documentation technology (Bird, Hanke, Adams, & Lee, 2014). In situations where direct questioning is not culturally appropriate (as discussed above), cultural insiders will be more attuned to whatever interactional norms are used for eliciting speech. These consultants may be cited in academic works as contributors, but it is even better to take this a step further: genuinely incorporate a local community member into the research as an equal colleague, including collaborative conference presentations and peer-reviewed journal articles. Such collaborative relationships can be very valuable for both sides. One example of effective collaboration between a variationist sociolinguist and a cultural insider is Stanford & Pan's (2013) research on Zhuang of southern China. In Australia, an outstanding recent example of sociolinguistic collaboration with a cultural insider is Kral and colleagues' (2015) study of expressive phonetic affects used in Ngaanyatjarra story-telling practices, in which one co-author, Elizabeth Ellis, is both a scholar and an expert practitioner.

For many field linguists, the most rewarding part of their work is the unaccustomed level of social, emotional and intellectual engagement with people whose lives are very unlike their own. Some indigenous people find just as much fascination in their encounter with the linguist (Sutton, 2009, p. 163ff.). Likewise, sociolinguistic fieldwork can be a particularly stimulating endeavour for both parties, where each is simultaneously learning about the others' culture, as well as gaining new insights into their own.

**3. THE CHALLENGE OF IDENTIFYING SOCIOLINGUISTIC VARIABLES IN AN UNDERDOCUMENTED LANGUAGE.** The outsider researcher may find ways to adapt culturally to indigenous sociolinguistics. But there are still further issues in identifying and analysing variables while at the same time learning the grammar.

Describing the grammatical system of a language – not the fine details of variation and gradience, but the major patterns of phonology, morphology and syntax – is typically the first priority of linguistics research in indigenous communities. Documenting variation is often a peripheral concern, and some field linguists might argue that this is simply inevitable. But Meyerhoff (in press) discusses ways in which variationist approaches can illuminate the basic structural description. Even so, trying to analyse an undocumented or under-documented language is in itself a massive task, and it is challenging to describe sociolinguistic variation without a good description of the features that vary, as well as related features that may be relatively invariant.

Compare this to the greater part of Western sociolinguistic research, in which the language under discussion often has been minutely analysed for decades or centuries. Not only is the grammar known, but the variables themselves are often well known, and have been subject to previous analysis (e.g. English /ɪŋ ~ ən/, rhoticity, t/d-deletion, etc.). Just identifying variables may take a long time in indigenous sociolinguistics, and without a good feel for how the language is used, identifying which variables are socially significant takes even longer.

Identifying sociolinguistic variables is complicated by the unknown terrain of what different dialects are in play. Should a variant be analysed as a socially meaningful expression of “the same language”, or rather as a code-switch into a different language variety? Are our consultants even speaking the same language? Understudied language ecologies (Mühlhäusler, 2002) therefore present a special challenge for identifying sociolinguistic variables and defining the envelope of variation.

*Solution: Tapping into folk linguistic knowledge*

One practical approach to identifying sociolinguistic variables in an understudied language is to write down every instance of variation encountered during language learning or other interactions: note who said it and make a quick guess about possible reasons for the variation. Could it be a variant from a different regional dialect? A different speech style or register? Is it “free variation” or is it conditioned by linguistic context? Is it associated with the social identity of the speaker? Nagy (2009) provides a practical roadmap to organize such linguistic and social information into a detailed “sociogrammar” of the language. At the time when the variable is first encountered, it is likely that correlations will not be clear, but careful notes may gradually reveal patterns. For example, early in the second author’s Sui language learning in southwest China, speakers began telling him that the first singular pronoun differs widely across the Sui region. As he kept observing and asking questions, he eventually found that speakers had a detailed “mental map” of many different clans/locations that used different variants, especially 1st singular. He asked questions like, “I’m learning your language, and I need to learn to handle all kinds of different ways of speaking... How do you say *me*?” He later conducted a survey of first singular production, and found that it matched the community’s knowledge of the locations of these variants very accurately. Of course, this is not always the case, as community members’ perceived notions and self-reported variants may differ drastically from results of production studies. Nonetheless, such an approach can be a very valuable starting point – and it is a great way to build relationships in the community. Moreover, any discrepancies between production results and a community’s “folk linguistic” notions can become a valuable question to study as well (Preston 1989, 1993).

In other cases, the speakers may downplay variation, making it seem like the variants are irrelevant. But it is important to follow up each possible variant because they may be hints of crucial structural or sociolinguistic patterns. The following are some typical “red flags” that a researcher might hear when starting a research project.

- “That’s just how we say it sometimes. It’s no big deal. It means the same thing...”
- “That’s just how some younger people say it...you shouldn’t say it that way...”
- “That’s just how some of those old men say it...it’s old-fashioned...”
- “That’s just how we say it when we’re talking casually, not trying hard...”

- “That’s not the right way to say it, but we sometimes talk that way... but you really shouldn’t learn it that way...”

It can also be useful to directly ask for attitudes, imitations, and other folk views (Preston 1993). Although sociolinguists tend to emphasize speech that is the speakers’ own vernacular (Labov, 1966/1972), rather than imitations of other varieties, the latter can be very useful when exploring a community for the first time. Stereotypical pronunciations may be exaggerated, but they still give clues about linguistic variants and the social meaning attached to them. Niedzielski & Preston’s (2000) work on folk linguistics shows the value of the local non-linguists’ perspectives. Naturally, speakers are more likely to offer words or phrases as variants, rather than phonological or grammatical features (Chambers 1992), and in some cases speakers’ dialectal imitations are mutually contradictory and thus unrevealing. But in other cases a fieldworker can uncover patterns in dialectal stereotypes, especially where multiple speakers give similar responses using different example words.

Be genuinely interested, even if you feel like you have heard it all before. Multiple speakers’ responses can provide valuable confirmation and add new twists and perspectives. The following are some examples of questions probing this kind of information:

- “Tell me about different ways of speaking around here. Do people over there in that village speak the same as you?”
- “I’m trying to learn to speak your language, but I’m confused about different pronunciations/words. I noticed that some people say \_\_\_ but you say \_\_\_. Why is that? Am I hearing it correctly?”
- “Can you imitate the way people in that village speak? Can you teach me to speak like that village? What are some key words I should know?”
- “Do farmers speak differently than people in the town? How does it sound?”
- “Can you imitate how a man/woman speaks? How about a teenager? How about a child?”
- “Do older/younger people speak differently than you? How has the language changed during your lifetime? What do you think about that?”
- “Which dialect do you think I should learn? Why? Are some dialects ‘better’ than others or are they all ‘equal’?” Also ask about codeswitching: When and why do you switch languages/dialects?”

Figure 1 shows how such questions can prompt speakers to begin playfully imitating each other, identifying variants for further investigation. This is a conversation between two middle-aged women living in the same Sui village. Sui is a tonal language (Tai-Kadai family), and numerous dialect differences exist between clans. These two women had married into the village from different clans. The key variants here are a high tone (Tone 6H) versus a low tone (Tone 6L), different first singular variants (*nja* versus *nie*), voicing of nasal onsets (*hn-* versus *n-*), and a very common stereotypical performance speech phrase, *ja-ju*, which is formed with a discourse marker *ja* ‘like this’ placed next to *ju* first singular. All of this information about linguistic variants appears in a short stretch of imitations between these two speakers. The distinctive features of Person 1’s clan dialect are shown in boldface font, and Person 2’s clan dialect features are underlined.

Interviewer: So if you imitate her speaking, what does she do?

**Person 1:** So she [usually] says ja<sup>6H</sup>-ju<sup>2</sup>, so I'd say to her, "ja<sup>6H</sup> nja<sup>2</sup> pai<sup>1</sup> hnu<sup>1</sup> le<sup>2</sup>" [where are you going]...

...

Person 2: My place says it like this, "ja<sup>6H</sup> nja<sup>2</sup> pai<sup>1</sup> ndjong<sup>3</sup> hnu<sup>1</sup> le<sup>2</sup>"

...but her place says "nai<sup>6H</sup> nie<sup>2</sup> pai<sup>1</sup> ndjong<sup>3</sup> hnu<sup>1</sup> le<sup>2</sup>"

**Person 1:** <interrupts to correct her> "**nai<sup>6L</sup> nie<sup>2</sup>**... **nai<sup>6L</sup> nie<sup>2</sup>** pai<sup>1</sup> nu<sup>1</sup> le<sup>2</sup>"

Person 2: <laughs, changes to Person 1's tone> **nai<sup>6L</sup>** ... **nai<sup>6L</sup>** ...

**Person 1:** "**nai<sup>6L</sup> nie<sup>2</sup>** pai<sup>1</sup> nu<sup>1</sup> le<sup>2</sup>" Like this.

Person 2: <laughs> Sometimes we're a bit wrong.

FIGURE 1: Two Sui women imitating each other's dialect features for the phrase *Where are you going*. The conversation was monolingual Sui, and it is translated into English here. **Bold**=the dialect variants of **Person 1**. Underline=the dialect variants of Person 2. Superscripts represent tones (see Stanford 2009: 479-80).

**4. THE SOCIOLINGUISTIC ANALYSIS CHALLENGE: LIMITED DATA.** We now turn to the third challenge. Just as grammatical analysis is limited at an early stage of documentation, quantitative analysis may also be limited by data volume. As discussed above, social norms at the field site and the slow initial process of uncovering variation may prohibit data collection from a large sample of speakers, so statistical methods may be challenging to apply.

Collaboration and engagement with locals can increase the quantity of sociolinguistic data, but in many cases it will still fall short of the token counts typically used for urban sociolinguistics. This is not a fatal flaw, but it may limit the potential for nuanced factor analyses. For example, Harlow and colleagues' (2009) study of changing stop realisation in Maori focuses on a comparison between aspiration rates in older and newer data, whereas an urban study might be able to extract more detailed findings on age-grading and its interaction with other social factors (e.g. English: Wagner, 2012). Nonetheless, this finding for Maori is still a highly worthwhile advance, since the variable and its social significance are otherwise underdocumented. It can also be worth reporting sociolinguistic variation even where data quantity is not sufficient for statistical analysis – for example, if variation has been observed among a handful of speakers (e.g. Yolngu Matha: Bower, 2008).

In some situations, such as Jamsu Reynolds' (2012) work on Amdo Tibetan in Qinghai province (China), a solid sample of 80–100 speakers may be possible. But in other cases, a sociolinguistic sample of such sizes may be out of the question (cf. Tagliamonte, 2006, p. 185). There may not even be 50–100 speakers in existence for a given language variety. Moreover, if we are able to record just 10 or 20 speakers, then once these are divided into clan, age and gender categories (or whatever is appropriate in the social setting), we may have so few speakers in each demographic cell that social characteristics and individual characteristics cannot be distinguished.

In cases of severe language endangerment, we may even have recorded all the living speakers (Evans, 2001). Even in cases where we can record more speakers, transcription of naturalistic speech is far more difficult when the researcher does not speak the language well. Tagliamonte (2006, p. 54) estimates four hours of transcription work for every hour of recorded data (4:1). But for the first author's research in Murrinhpatha, the ratio is closer to 10:1. In a recent workshop on language documentation technology, a ratio of 50:1 was proposed (Cavar, 2016). In summary, the requirement to find quantified sociolinguistic patterns is severely challenged by the difficulty of obtaining large enough data samples in these situations.

### *Solutions to the challenge of limited data*

There are a number of ways to study sociolinguistic variation where there is insufficient naturalistic data. An example of a sociolinguistic investigation hampered by a limited volume of data is Mansfield's (2015) study of variable suffix ordering in Murrinhpatha, a morphologically complex Aboriginal language of northern Australia. Murrinhpatha is spoken by some 3000 people, and is still acquired by children as a first language. Mansfield found that three classes of suffix morphemes, TENSE, NUMBER and IMPERFECTIVE, can be produced in variable orders, contrary to the rigid order attested in earlier grammatical descriptions (e.g. Street, 1987). For example, any of the orders (1–3) are possible, without any change in meaning:

- (1) parde-lili-dha-nime-pardi  
be.3PAUC.PST-walk-PST-PAUC.M-BE.IMPF<sup>3</sup>  
{-TENSE-NUM-IMPF}  
'they were walking'
- (2) parde-lili-nime-dha-pardi  
be.3PC.PST-walk-PAUC.M-PST-BE.IMPF  
{-NUM-TENSE-IMPF}
- (3) parde-lili-nime-pardi-dha  
be.3PC.PST-walk-PAUC.M-BE.IMPF-PST  
{-NUM-IMPF-TENSE} (Mansfield 2015)

Mansfield suspected that suffix order is correlated with age of speakers, with younger speakers moving away from the -TENSE-NUM-IMPF sequence attested in earlier grammars. However the variable only appears in verbs that have multiple suffixes. After some 12 months, tokens had only been collected from 10 speakers, and the diversity of syntactic contexts made comparability problematic. This initial data set was somewhat suggestive of a correlation between suffix sequence and speaker age, but was insufficient as evidence. To investigate the association of the variable with age, more tokens from more speakers were needed. But as described above, initiating recording sessions with new speakers – especially those from different clans and age groups than the speakers with whom Mansfield was already well acquainted – was a very slow process.

<sup>3</sup> IMPF = imperfective; M = masculine; NUM = number; PAUC = paucal; PST = past.

Mansfield accelerated the analysis of Murrinhpatha suffix ordering by turning to elicited speech as a way to raise the token count. This was made possible by the fact that the variation was not limited to spontaneous speech, but rather occurred in all speech styles, with little evidence of self-correction to a “standard” variant. Mansfield was thus able to collect more tokens of the variable by targeted questioning, as opposed to waiting for the variable context to turn up in naturalistic recordings. In the first instance, the token count was expanded using a version of the “rapid anonymous survey” (Labov 1966/1972, see also Matsuda 1993, Borowsky & Horvath, 1997). The rapid anonymous survey technique is uncommon in modern sociolinguistics, despite its fame as a part of Labov’s (1966/1972) work in New York City. Mansfield developed a rapid data collection method to obtain further tokens of the somewhat rare Murrinhpatha suffix ordering variable. In day-to-day encounters with Murrinhpatha speakers, Mansfield began emphasizing his own difficulty in producing verbs with past tense and paucal number suffixes, provoking a context for suffix variation. This prompted native speakers to provide clear tokens of the missing verbs for Mansfield, who noted the variants produced and the approximate speaker ages in a notebook after the Labovian method. The data thus gathered cannot be taken as a representation of vernacular Murrinhpatha, but it did produce a collection of 101 tokens from 25 speakers, in which the hypothesized correlation of suffix ordering and speaker age was supported. Speakers in their teens and early twenties showed a strong preference for the -NUMBER-TENSE sequence (which is unattested in earlier grammatical descriptions, e.g. Street 1987), while speakers in their mid-twenties and above were more likely to use the -TENSE-NUMBER sequence (Mansfield 2015). The rapid and anonymous data thus shed light on an age-related pattern, while leaving further questions regarding gender, clan heritage and linguistic factors unanswered.

To further expand the collection of Murrinhpatha variable suffix tokens, Mansfield turned to picture stimuli. Picture stimuli can be very effective in situations where a language does not have its own writing system or where the writing system is not widely used. Speakers were asked to describe events that happened “yesterday” or will happen “tomorrow” in a series of a dozen pictures, depicting different events all featuring dual or paucal participants. Again these do not attempt to represent the vernacular, but in comparison to rapid and anonymous data, they induce each speaker to produce the variable with a wider and more controlled range of verb lexemes and morphosyntactic features. Finally, more tokens of the variable from an older generation of speakers were collected by trawling archives of speech recorded in previous decades. Together these data sources are being utilized for further quantitative analysis of linguistic and sociolinguistic factors influencing Murrinhpatha suffix ordering. In summary, the ideal dataset of comparable, naturalistic samples from a large number of speakers is not available for this variable. However, elicited speech samples can be used to provide an adequate number of tokens and alternation of contexts. This initially allowed an age correlation to be established, and prospectively offers hope for establishing further correlations in ongoing research.

Where statistical significance is problematic (and is acknowledged as such), hypothesised patterns may still be worth reporting. Without reporting such initial results for a given language, the state of published knowledge about variation would be one of complete ignorance. We need to report such preliminary findings so that more robust sociolinguistic analyses of indigenous language ecologies can be developed. In addition, as noted in section 2, cultural insiders may be linguists themselves, bridging the divide to document their own languages, and making larger field projects

more attainable. Cultural insiders have special access to ethnographic and qualitative information, probing individual speakers' sociolinguistic motivations and personal backgrounds, as well as folk linguistic perspectives (Preston 1993, Niedzielski & Preston 2000). Collecting direct quotations on these matters can help us understand the social meaning and usage of variables, as well as communicate those local viewpoints to academic audiences (recall Figure 1, for example). Regardless of the size of the data set, a few representative direct quotations showing folk perspectives can be a valuable window into speakers' attitudes and usage. In other words, this can serve as an insightful supplement to the quantitative results. For this reason, field researchers may want to include numerous metalinguistic discussions or question-and-answer sessions where language/dialect attitudes can be recorded.

**5. CONCLUSION.** In this article we have argued that documenting sociolinguistic variation in indigenous languages poses methodological challenges, but can be deeply rewarding for community members, outsider researchers, and perhaps especially for community members who are sociolinguistic researchers themselves. The linguistic subfields of language documentation and variationist sociolinguistics have until recently been largely independent, but there is now an emerging interest in the intersection of these pursuits.

For outsider researchers working in these communities, the challenges of sociolinguistic field research in indigenous communities begin with the large cultural gap that often exists between researchers and the speech community, and often a lack of fluency in a shared language. The outsider researcher is learning how to fit in with local norms and expectations, how to interpret interactions with local contacts, and at the same time attempting to conduct research. The second challenge, identifying sociolinguistic variables, is difficult in this setting, but even the earliest observations should be noted carefully, and the researcher should be alert to any hints that locals provide in metalinguistic comments. Third, the quantity of speech data recorded may be limited by community size and accessibility to a Western researcher.

We have argued that these challenges can be mitigated by (1) embracing a more engaged type of research relationship: one that is multiplex, personal and collaborative. Since researchers are often outsiders in the speech communities we seek to research, we must engage cultural insiders as collaborators. These mutually beneficial research relationships can have a long-lasting positive impact on both the community and the outside researcher. (2) We also suggest several modified sociolinguistic field methods, including an emphasis on folk linguistic approaches to help uncover variation during the early stages of a new research program in a lesser-studied language. (3) In lesser-studied indigenous communities, variationist researchers may also need to revise our expectations of data set sizes, especially during an initial study. Even so, with some creative approaches, it is possible to expand our data sets and "go deep" with our analysis of the existing data.

The intersection of sociolinguistics and indigenous language documentation is at an exciting point of florescence. Researchers working at this interface have the chance to pioneer new forms of sociolinguistic methodology. This is especially important in light of the *Principle of Sociolinguistic Distance* that we suggested in section 1. As researchers explore socioculturally and linguistically diverse field situations, we are all likely to have more opportunities to test prior assumptions and uncover new perspectives and research questions. Bridging the gap between variationist and documentary traditions can be challenging, but it is also a valuable opportunity

for new methods and new knowledge. Moreover, as outsider researchers in these communities, we look forward to the ever-increasing role of community insiders as partners and leaders in this process.

## REFERENCES

- Bavin, Edith L. 1992. The acquisition of Warlpiri. In Dan Slobin (ed.), *The crosslinguistic study of language acquisition*. Hillsdale: Lawrence Erlbaum, 309–372.
- Bird, Steven., Florian R. Hanke, Oliver Adams & Haejoong Lee. 2014. Aikuma: A mobile app for collaborative language documentation. *Workshop on the Use of Computational Methods in the Study of Endangered Languages*, 1–5.
- Borowsky, Toni, & Barbara Horvath. 1997. L-vocalization in Australian English. In Frank Hinskens, Roeland van Hout, & W. Leo Wetzels (eds.), *Variation, change and phonological theory* 101–124. Amsterdam: John Benjamins Publishing Company.
- Bowern, Claire. 2008. Seven speakers, eight varieties: Variation in an Arnhem Land clan language. Presented at the NWA V 37, Rice University, Houston.
- Bowern, Claire. 2015. *Linguistic fieldwork: A practical guide* (Second edition). Basingstoke: Palgrave Macmillan.
- Bradley, John. 1998. Yanyuwa: “Men speak one way, women speak another.” In Jennifer Coates (ed.), *Language and gender: A reader*. 13–20, London: Wiley.
- Bybee, Joan L. 2006. From usage to grammar: the mind’s response to repetition. *Language* 82. 711–733.
- Cavar, Damir. 2016. Bridging between language documentation, computational and theoretical linguistics, and natural language processing. Presented at the Language Documentation Tools Summit, University of Melbourne.
- Chambers, Jack K. 1992. Dialect acquisition. *Language* 68(4). 673–705.
- Chelliah, Shobhana L. & Willem J. de Reuse. 2010. *Handbook of descriptive linguistic fieldwork*. Dordrecht: Springer.
- Clarke, Sandra. 2009. Sociolinguistic stratification and new dialect formation in a Canadian aboriginal community: Not so different after all? In James Stanford & Dennis Preston (eds.), *Variation in indigenous minority languages*. 109–128. Amsterdam: Benjamins.
- Di Paolo, Marianna, & Malcah Yaeger-Dror. 2011. *Sociophonetics: A student’s guide*. New York: Routledge.
- Dorian, Nancy. 1981. *Language death: the life cycle of a Scottish Gaelic dialect*. Philadelphia: University of Pennsylvania Press.
- Dryer, Matthew S., & Martin Haspelmath (eds.). 2005. *The world atlas of language structures*. Oxford: Oxford University Press.
- Eckert, Penelope. 1989. *Jocks and burnouts: Social categories and identity in the high school*. New York: Teacher College Press.
- Enfield, Nick J. 2002. Introduction. In Nick J. Enfield (ed.), *Ethnosyntax: Explorations in culture and grammar*. 1–30. Oxford: Oxford University Press.
- Evans, Nicholas. 2001. The last speaker is dead – long live the last speaker! In Paul Newman & Martha Ratliff (eds.), *Linguistic fieldwork*. Cambridge: Cambridge University Press, 250–281.
- Evans, Nicholas. 2009. *Dying words: Endangered languages and what they have to tell us*. Chichester: Wiley-Blackwell.
- Everett, Daniel L. 2001. Monolingual field research. In Paul Newman & Martha Ratliff (eds.), *Linguistic fieldwork*. 166–88. Cambridge: Cambridge University Press.
- Feagin, Crawford. 2002. Entering the community: Fieldwork. In J. K. Chambers, Peter Trudgill, & Natalie Schilling-Estes (eds.), *Handbook of language variation and change*. 20–39. Oxford: Oxford University Press.
- Foley, Lawrence. 1980. Sociolinguistic patterns. In Lawrence Foley (ed.), *Phonological variation in Western Cherokee*. 158–213. New York: Garland.

- Furlan, Alberto. 2005. *Songs of continuity and change: The reproduction of Aboriginal culture through traditional and popular music*. PhD thesis, University of Sydney.
- Goldberg, Adele. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
- Good, Jeff. 2012. 'Community' collaboration in Africa: Experience from Northwest Cameroon. In Peter Austin & Stuart McGill (eds.), *Language documentation and description, vol. 11*. 28–58. London: SOAS.
- Haeri, Niloofar. 1996. "Why do women do this?" Sex and gender differences in speech. In Gregory R. Guy, Crawford Feagin, Deborah Schiffrin, & John Baugh (eds.), *Towards a social science of language: Papers in honor of William Labov* (Volume 2). 101–114. Amsterdam: John Benjamins Publishing Company.
- Harlow, Ray, Peter Keegan, Jeanette King, Margaret Maclagan & Catherine Watson. 2009. The changing sound of the Maori language. In James N. Stanford & Dennis R. Preston (eds.), *Variation in indigenous minority languages*. 129–52. Amsterdam: John Benjamins.
- Hildebrandt, Kristine. 2003. *Manange tone: Scenarios of retention and loss in two communities*. PhD dissertation, University of California, Santa Barbara.
- Hildebrandt, Kristine. 2005. A phonetic analysis of Manange segmental and suprasegmental properties. *Linguistics of the Tibeto-Burman area* 28. 1–36.
- Hildebrandt, Kristine, Carmen Jany & Wilson Silva. 2016. Symposium on documenting variation in endangered languages. Linguistic Society of America Annual Meeting, Washington, DC. January 7.
- Hill, Jane. 1996. Languages on the land: Toward an anthropological dialectology. In David Skomp distinguished lectures in dialectology. 4–40. Bloomington: Indiana University Press.
- Hill, Jane & Kennet Hill. 1986. *Speaking Mexicano: Dynamics of syncretic language in central Mexico*. Tucson: Arizona University Press.
- Hopper, Paul. 1987. Emergent Grammar. *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society* 13. 139–157.
- Jamsu Reynolds, Jermay. 2012. *Language variation and change in an Amdo Tibetan village: Gender, education, and resistance*. PhD dissertation, Georgetown University.
- Jones, Caroline and Felicity Meakins. 2013. Variation in voice onset time in stops in Gurindji Kriol: Picture naming and conversational speech. *Australian Journal of Linguistics* 33(2). 196–220.
- Kral, Inge, Jennifer Green, Elizabeth Marrkilyi Ellis, Jane Simpson, & Hywel Stoakes. 2015. Desert clicks, taps and sound symbolism: Vocal style in Ngaanyatjarra narratives. Presented at the Australian Linguistics Society, Melbourne.
- Krauss, Michael. 1992. The world's languages in crisis. *Language*, 68(1). 4–10.
- Labov, William. 1966/1972. The social stratification of (r) in New York City department stores. Chapter 2 in William Labov *Sociolinguistic patterns*. 43–54. Philadelphia: University of Pennsylvania Press.
- Labov, William. 1972a. *Language in the inner city: Studies in the Black English vernacular*. Oxford: Basil Blackwell.
- Labov, William. 1972b. Some principles of linguistic methodology. *Language in Society* 1. 97–120.
- Labov, William. 1994. *Principles of linguistic change. Vol. 1: Internal factors*. Malden, Massachusetts: Wiley Blackwell.
- Labov, William. 2001. *Principles of linguistic change, Vol. 2: social factors*. Malden, Massachusetts: Wiley Blackwell.

- Labov, William. 2010. *Principles of linguistic change. Vol. 3: cognitive and cultural factors*. Malden, Massachusetts: Wiley Blackwell.
- Labov, William. 2013. *The language of life and death: The transformation of experience in oral narrative*. Cambridge: Cambridge University Press.
- Mahood, Kim. 2012. Kartiya are like Toyotas. *Griffith Review*, 36. 43–59.
- Mallinson, Christine, Becky Childs & Gerard Van Herk (eds.). 2013. *Data collection in sociolinguistics: Methods and applications*. London/New York: Routledge.
- Mansfield, John B. 2014. *Polysynthetic sociolinguistics: The language and culture of Murrinh Patha youth*. PhD dissertation. Australian National University, Canberra.
- Mansfield, John B. 2015. Morphotactic variation, prosodic domains and the changing structure of the Murrinhpatha verb. *Asia-Pacific Language Variation*, 1(2). 162–188.
- Matsuda, Kenjiro. 1993. Dissecting analogical leveling quantitatively: The case of the innovative potential suffix in Tokyo Japanese. *Language Variation and Change*, 5(1). 1–34.
- Meyerhoff, Miriam. In press. Writing a linguistic symphony: Analysing variation while doing language documentation. *Canadian Journal of Linguistics*.
- Meyerhoff, Miriam. 2015a. *Doing sociolinguistics: A practical guide to data collection and analysis*. New York: Routledge.
- Meyerhoff, Miriam. 2015b. Turning variation on its head: Analysing subject prefixes in Nkep (Vanuatu) for language documentation. *Asia-Pacific Language Variation* 1(1). 79–109.
- Meyerhoff, Miriam. 2009. Animacy in Bislama? Using quantitative methods to evaluate transfer of a substrate feature. In James Stanford & Dennis Preston (eds.), *Variation in indigenous minority languages*. 369–396. Amsterdam: Benjamins.
- Michaels, Eric. 1986. *The Aboriginal invention of television in central Australia, 1982–1986*. Canberra: Australian Institute of Aboriginal Studies.
- Miller, Wick. 1965. *Acoma grammar and texts*. Berkley: University of California Press.
- Milroy, Leslie. 1987. *Language and social networks* (Vol. 2). Oxford: Blackwell.
- Moses, Karin and Colin Yallop. 2008. Questions about questions. In Jane Simpson and Gillian Wigglesworth (eds.), *Children's language and multilingualism*. London: Continuum, 30–55.
- Mühlhäusler, Peter. 2002. Why one cannot preserve languages (but can preserve language ecologies). In David Bradley & Maya Bradley (eds.), *Language endangerment and language maintenance*. 34–39. London: Routledge Curzon.
- Nagy, Naomi. 2009. The challenges of less commonly studied languages: Writing a sociogram of Faetar. In James Stanford & Dennis Preston (eds.), *Variation in indigenous minority languages*. 397–417. Amsterdam: Benjamins.
- Nagy, Naomi & Miriam Meyerhoff. 2008. The social lives of language. In Miriam Meyerhoff & Naomi Nagy (eds.), *Social lives in language – Sociolinguistics and multilingual speech communities: Celebrating the work of Gillian Sankoff*. 1–17. Amsterdam: Benjamins.
- Nettle, Daniel, & Suzanne Romaine. 2000. *Vanishing voices: The extinction of the world's languages*. Oxford: Oxford University Press.
- Niedzielski, Nancy & Dennis Preston. 2000. *Folk linguistics*. Berlin/New York: Mouton de Gruyter.
- O'Shannessy, Carmel. 2009. Language variation and change in a North Australian indigenous community. In James Stanford & Dennis Preston (eds.), *Variation in indigenous minority languages*. 419–440. Amsterdam: Benjamins.

- Pierrehumbert, Janet B. 2001. Exemplar dynamics: word frequency, lenition and contrast. In *Frequency and the emergence of linguistic structure*. 135-37. Amsterdam: John Benjamins.
- Preston, Dennis. 1989. *Perceptual dialectology: Nonlinguists' views of areal linguistics*. Dordrecht: Foris.
- Preston, Dennis. 1993. The uses of folk linguistics. *International Journal of Applied Linguistics* 3(2). 181-259.
- Sakel, Jeanette & Daniel L. Everett. 2012. *Linguistic fieldwork: A student guide*. Cambridge: Cambridge University Press.
- Sankoff, Gillian. 1980. *The Social life of language*. Philadelphia: University of Pennsylvania Press.
- Smakman, Dick. 2015. The Westernising mechanisms in sociolinguistics. In Dick Smakman & Patrick Heinrich (eds.), *Globalising sociolinguistics*. 16-36. London: Routledge.
- Smith, Linda Tuhiwai. 1999. *Decolonizing methodologies: Research and indigenous peoples*. London: Zed Books.
- Smith, Ian, & Steve Johnson. 1986. Sociolinguistic patterns in an unstratified society: The patrillects of Nghanhcara. *Journal of the Atlantic Provinces Linguistic Association* 8. 29-43.
- Stanford, James. 2007. *Dialect contact and identity: A case study of exogamous Sui clans*. PhD dissertation, Michigan State University.
- Stanford, James. 2009. Clan as a sociolinguistic variable. In James N. Stanford & Dennis R. Preston (eds.), *Variation in indigenous minority languages*. 463-484. Amsterdam: John Benjamins.
- Stanford, James. 2016. A call for more diverse sources of data: Variationist approaches in non-English contexts. *Journal of Sociolinguistics* 20(4). 525-41.
- Stanford, James & Yanhong Pan. 2013. The sociolinguistics of exogamy: Dialect acquisition in a Zhuang village. *Journal of Sociolinguistics* 17(5). 573-607.
- Stanford, James & Dennis Preston (eds.) 2009. *Variation in indigenous minority languages*. Amsterdam: Benjamins.
- Stewart, Jesse. 2015. Production and perception of stop consonants in Spanish, Quichua, and Media Lengua. PhD thesis, University of Manitoba.
- Street, Chester. 1987. *An introduction to the language and culture of the Murrinh-Patha*. Darwin: Summer Institute of Linguistics.
- Suokhrie, Kelhouvinuo. 2016. Clans and clanlectal contact: Variation and change in Angami. *Asia-Pacific Language Variation*, 2(2). 188-214.
- Sutton, Peter. 1978. *Wik: Aboriginal society, territory, and language at Cape Keerweer, Cape York Peninsula, Australia*. PhD dissertation, University of Queensland.
- Sutton, Peter. 2009. *The politics of suffering: Indigenous Australia and the end of the liberal consensus*. Melbourne: Melbourne University Publishing.
- Suwankhong, Dusanee, & Pranee Liamputtong. 2015. Cultural insiders and research fieldwork: Case examples from cross-cultural research with Thai people. *International Journal of Qualitative Methods*, 14(5). 1-7.
- Tagliamonte, Sali A. 2006. *Analysing sociolinguistic variation*. Cambridge: Cambridge University Press.
- Thomas, Erik. 2011. *Sociophonetics: An introduction*. Basingstoke: Palgrave Macmillan.
- Tomasello, Michael 2003. *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

- Trugill, Peter. 1972. Sex, covert prestige and linguistic change in urban British English. *Language in Society* 1. 179–95.
- Trudgill, Peter. 2011. *Sociolinguistic typology: social determinants of linguistic complexity*. Oxford: Oxford University Press.
- Wagner, Suzanne Evans. 2012. Real-time evidence for age grad(ing) in late adolescence. *Language Variation and Change*, 24(2). 179–202.
- Walsh, Michael. 1997. Cross cultural communication problems in Aboriginal Australia. Darwin: North Australia Research Unit.

john.mansfield@unimelb.edu.au

james.n.stanford@dartmouth.edu