Census data on Australian Languages

JANE SIMPSON, DENISE ANGELO, EMMA BROWNE, INGE KRAL, FRANCIS MARKHAM, CARMEL O'SHANNESSY & DANIELLE VENN

Census data on Australian Languages

Jane Simpson, Denise Angelo, Emma Browne, Inge Kral,
Francis Markham, Carmel O'Shannessy, Danielle Venn

Australian National University, Canberra ACT 2612

[jane.simpson@anu.edu.au, denise.angelo@anu.edu.au, emma.browne@anu.edu.au,
francis.markham@anu.edu.au, carmel.o'shannessy@anu.edu.au, danielle.venn@anu.edu.au]

Abstract

In Australia, language ecologies have become complex, as speakers of Australian Indigenous languages (whether traditional or new contact languages) may be multilingual with different levels of proficiency in different languages. Using the Indigenous language spoken by one's ancestors asserts association with a speech community (past or present). People may be using the language for everyday talk, or they may be using it along with other languages, including standard English, the language of wider communication, or they may be re-learning the language. The speakers may live in the heartlands of their traditional country or they may be living in diaspora communities around Australia. Quantitative data on the location and number of speakers of minority languages is needed for convincing governments of where they should direct resources for education, interpreting and service provision. Some data on language use (direct or inferred) is obtainable from the five-yearly *Census of Population and Housing* surveys by the Australian Bureau of Statistics (ABS). The language questions in these studies lead to some unexpected findings that vary in plausibility when compared with case studies of particular communities. A data review of this source is provided, considering the strengths and weaknesses with respect to what has been observed in several case study communities. We propose ways for assessing quantitative results, and combining them with case studies of individual communities to produce more reliable models of language change.

Introduction

In Australia, with more than 300 Indigenous traditional languages, perhaps another 400 varieties, and many local new languages, the language ecologies are complex. The complexity lies along several dimensions:

- *First language*, that is, whether people speak a traditional Indigenous language as a first language, whether they speak a contact Indigenous language as a first language, or whether they speak English as their first language
- *Second language*: whether people are learning a traditional Indigenous language as a second language
- *Proficiency in the language of wider communication,* in this case standard English
- *Location*: do they live in communities where their first language (traditional or new) is the first language of the majority of the community, do they live in a diaspora community, do they live in remote communities, or do they live in cities
- *Age:* whether children speak the language is key to the vitality of the language.

Each of these dimensions has consequences for people and their communities.

- Children who speak a traditional or new Indigenous language may need support for education, both for receiving instruction in their first language and for learning English as a second language. This is especially necessary for children living in communities where English is not spoken much.
- Adults who speak a traditional or new Indigenous language and whose proficiency in English is low may need support for communication and access to services.

- Children and adults who do not speak a traditional or new Indigenous language & who wish to learn their heritage language may need support for doing this.

Children who speak a traditional or new Indigenous language may need support for maintaining this language, and this becomes especially critical for those living in diaspora situations where the language of everyday talk is a language other than their first language.

These consequences involve language rights: people's right to communication and access to government services, children's right to receive an education which gives them access to the dominant language, to literacy and to the wider society, and the right of Indigenous communities to keep and strengthen or revive their heritage Indigenous languages if they wish to do so.

For governments, non-government organizations, advocacy groups and communities to be effective in supporting the language rights of communities across the country, they need good information about the language ecologies. Only then will they be able to plan and provide services, good education and support for maintenance and revival programs.

This paper seeks to provide information about the national surveys, using Australian examples, in order to tease out their strengths and limitations. Communities, researchers and advocates may be able to make changes and additions to questions. A major question is how communities can make use of surveys to present positive views of their language ecologies while at the same time along with recognising where there are needs to be addressed. That is, we want to celebrate the achievements of language revival (e.g. the number of adults and children now putting the revived languages Wiradjuri and Gamilaraay as languages they speak at home in the national ABS census of Australia), while at the same time

avoiding complacency about the vitality of traditional languages still spoken as a first language by children. We also want to celebrate the new languages that have been generated by language contact and shift processes and have become Indigenous community vernaculars (e.g. Lockhart River Creole, Yarrie Lingo, Light Warlpiri).

Research questions

The dimensions of language ecology complexity discussed in the previous section lead to the following questions:

1. Where are people speaking Indigenous languages (traditional or new)), and do these languages have officially recognised names?
2. Where are people learning and/or reviving Indigenous languages (most likely traditional languages)?
3. What ages are the people who speak Indigenous languages?
4. What are the Indigenous languages they speak?
5. What level of proficiency have the speakers attained in Indigenous languages and in the language of wider communication (English in Australia)?
6. Are there changes in who speaks/learns what language over time?

Answers to these questions are relevant to future action with respect to maintenance, revival and services, especially when coupled with information about location, remote, rural and regional, and whether people are living in the heartland of their language group or in diaspora settings.

Data-sets that may answer research questions

Information on languages in Australia comes from four main sources: local language surveys, the National Indigenous Languages Surveys, the national Australian Bureau of Statistics (ABS) census, and from other national surveys such as *National Aboriginal and Torres Strait Islander Social Survey* (NATSISS), *Footprints in Time - The Longitudinal Study of Indigenous Children*.

Local language vitality surveys and tailored language surveys are important, but they are not practical in a country with more than 300 traditional languages, perhaps another 400 varieties of these languages, and many local new languages, spoken in many communities across Australia. The cost is very high. For example, the 2011 Yawuru *Knowing our Community* (YKC) household survey was the first Australian Indigenous controlled survey of a town, the remote town of Broome (Western Australia). It involved 20 interviewers working for a 4-month period interviewing more than 900 households,

along with a survey bus, team uniforms, promotional videos, and economic statisticians analysing the results (Taylor et al. 2014).

Australia has had two National Indigenous Languages Surveys (NILS) (Australian Institute of Aboriginal and Torres Strait Islander Studies and Federation of Aboriginal and Torres Strait Islander Languages 2005, Marmion et al. 2014). NILS1 provided some quantitative data on people, their languages and language activities in their region. This was obtained from around 281 people and organisations across the country, about languages spoken. NILS2 focussed on what language activities were taking place, what people thought about them, and what their attitude were towards Australian Indigenous languages. They received responses from around 288 Aboriginal and Torres Strait Islander people and 75 organisations. Because the surveys were online, this restricted who was likely to submit a response. The report focussed on some case studies of individual languages that represented particular language ecologies, rather than, as in NILS1, giving an overall picture of Indigenous languages.

Australia has a national census (the ABS census) conducted every five years. It also holds other more specific national surveys of Indigenous people (such as NATSISS focussing on social factors and LSIC focussing on child development). Of these, only the ABS census is truly comprehensive, attempting to enumerate everyone in Australia. All of the surveys mentioned have questions relevant to Indigenous people and Indigenous languages, but none are set up as language vitality surveys (i.e. no questions such as what language resources are available for the language, and only LSIC has a question on how the language is used in education). Indigenous people have made justified complaints about the failure of national surveys to recognise categories important to Indigenous people (Peter Yu, cited in Taylor et al 2014) or to recognise that some Indigenous practices (such as high mobility) reduce the usefulness of the national surveys as sources of data for social planning.

However, the costs and logistical difficulties of both the local surveys and of NILS are such that governments and other bodies will continue to draw information from large-scale non-local surveys that are not targeted towards information relevant to supporting language rights. As well, because these national surveys have questions related to economic, education and other social indicators, it may be possible to compare figures on use of Indigenous languages with these indicators to answer questions related to Government priorities[1].

The ABS Census and the research questions

The primary data-set is the ABS Census of Population and Housing. The most recent collection was in 2016.

---

[1] These could include whether strengthening (reviving or maintaining) Indigenous languages contribute to economies, to innovation and productivity, and to the well-being of

individuals and communities

The data is available in different formats but we discuss here the counts of persons in a particular category or cross-tabulated category presented as Aggregated tables. Its geographic and demographic coverage is national and aims to be complete. The methods of collections included both online/printed questionnaires and face-to-face interviews for remote Indigenous communities.

The ABS census asks people to state whether they are Indigenous[2]. In 2016 around 650,000 reported themselves (or children) as Indigenous (Aboriginal and/or Torres Strait Islander heritage) on the ABS census. There are major issues in remote areas with comprehension of census questions, and with locating participants, due to the high mobility of Indigenous participants. For example, a comparison between the YKC results (2011) and the 2006 ABS Census by Taylor et al (2014) indicated severe undercounting in the ABS results. Thus, in remote areas, the Indigenous population is likely to be underestimated.

The ABS proposes five "language standards" for language variables (Australian Bureau of Statistics 2016):

1. First language spoken
2. Languages spoken at home
3. Main language other than English spoken at home
4. Main language spoken at home, and
5. Proficiency in spoken English.

Reasons given for including these variables include:

> Data relating to [1-4] contributes to understanding proficiency in spoken English, which may be an indicator of ability to participate effectively in Australian society, including accessing government and other services. (Australian Bureau of Statistics 2016)

That is, the listed purpose is not for understanding language maintenance or revival, or well-being associated with this, or Australia's language resources. The only mention of this is given in the description of variable 5, Proficiency in Spoken English:

> The language a person can speak may be an important part of their culture, identity and well-being. However, for people living in Australia a lack of proficiency in spoken English may impact a person's access to employment, education and other government and non-government services. (Australian Bureau of Statistics 2016)

However, in the 2005 review of the *Australian Standard Classification of Languages* (Australian Bureau of Statistics 2005) it was observed that people wanted the ASCL to classify many languages so as to have data on languages with small numbers of speakers, whether they were emerging or declining (not differentiating between indigenous languages and immigrant languages).

The 2016 ABS census only addresses one of the language variables directly, (variable 3:

---

[2] (Q7 in 2016) "Is the person of Aboriginal or Torres Strait Islander origin? For persons of both Aboriginal and Torres

Q. 16 (2016) Does the person speak a language other than English at home?

• Mark one box only.

• If more than one language other than English, write the one that is spoken most often.

Yes, Mandarin

Yes, Italian

Yes, Arabic

Yes, Cantonese

Yes, Greek

Yes, Vietnamese

No, English only

Yes, other (please specify)

This question raises numerous issues which have been discussed before (Angelo and McIntosh 2014, Dixon and Angelo 2014, Kral and Morphy 2006, Morphy 2007, Simpson 2008, Simpson 2013), but which still persist in the 2016 census (albeit in attenuated form). In sum, the issues revolve around the complexity of the language ecologies. First, the question doesn't allow for people speaking more than one language other than English at home. Figure 1 (next page) for example shows figures for child speakers of Warlpiri and Alyawarr, but some people speak both.

Second, the answers are ambiguous between:

• speaks only Language X at home
• speaks Language X and Language Y/English at home equally
• speaks mostly English at home but occasionally uses Language X

In the case of some Warlpiri for example, it is likely that some people will put 'Warlpiri' down because it is their heritage language and they have a passive competence, and use some words in everyday talk.

Third, the question assumes people have the same idea of what 'speak' means. In the case of heritage learners of language X, they may strongly identify with their heritage, and so may say they speak language X at home because they are learning it, or because they use some words of language X at home. They may use English for everyday talk, but using language X marks their association with their ancestral language, and shows their aspiration to use it more generally. In the 2016 ABS census 432 people put down Wiradjuri, a language of Central NSW. To the best of our knowledge there are no L1 speakers of Wiradjuri, but there has been a widespread and effective campaign for promoting the learning of Wiradjuri, which includes the establishment of the North West Wiradjuri Language and Culture Nest (2013) based at Dubbo Primary School and with links to at least 6 small towns elsewhere in NSW. It is likely that these 432 L2
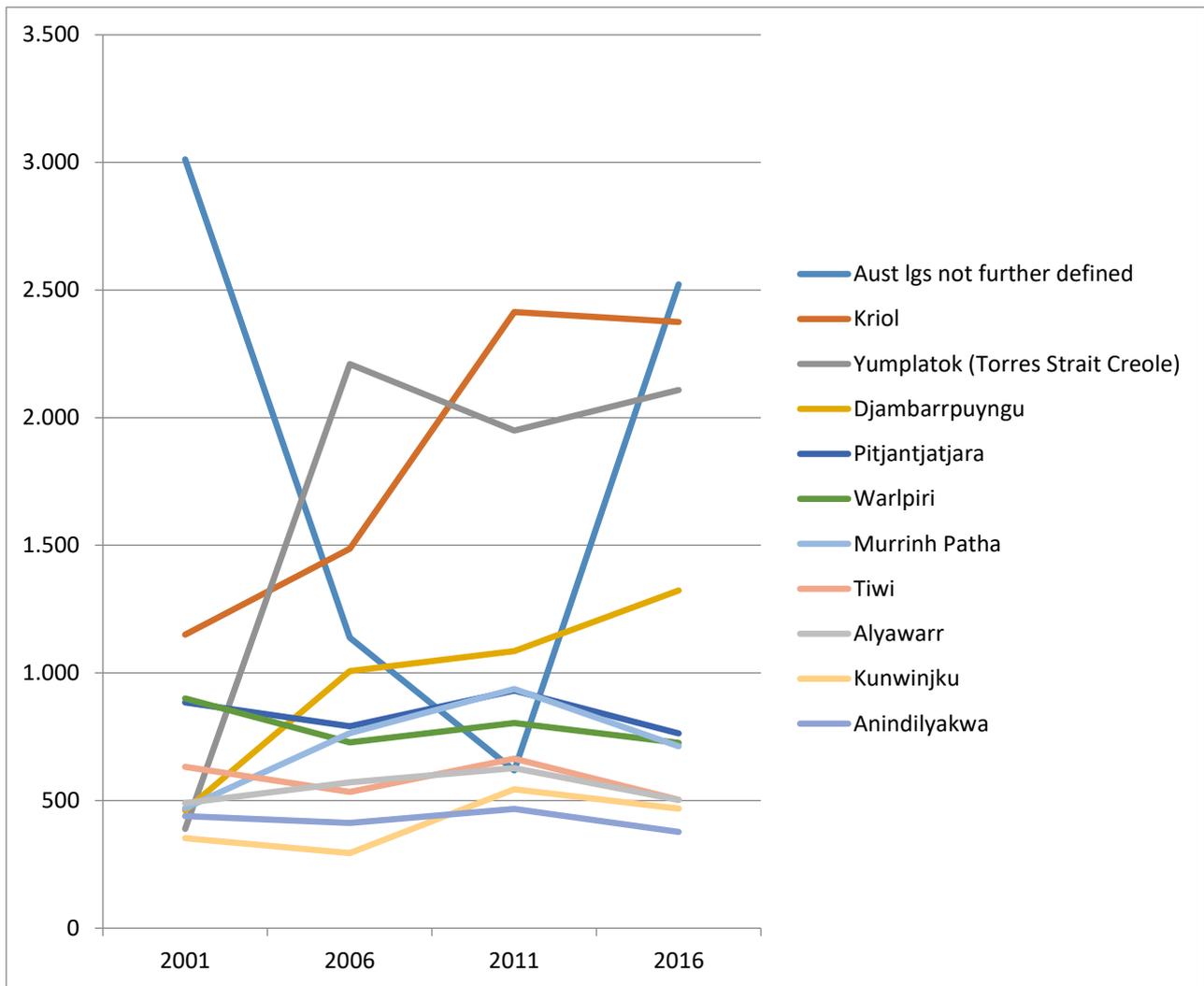
Strait Islander origin, mark both 'Yes' boxes"

Figure 1: Children reported as speaking one of the 10 languages with the most speakers, or "Australian Indigenous languages 'not further defined', ABS, 2001–2016 Census of Population and Housing, Customised Data Report

learners of Wiradjuri use Wiradjuri words and phrases at home, but that they most often use another language (probably English) for everyday talk.

Fourth, the question assumes that people who speak the same way share the same name for their language or way of talking, or indeed even have a name for their language. This is certainly not always the case in Australia even for traditional languages. Cases in point are the Yolŋu languages of Arnhem Land. In the 2001 census the majority language was labelled as Dhuwal-Dhuwala with 1,361 speakers. In 2006 the decision was made to recognise Djambarrpuyngu instead, and in subsequent census iterations Djambarrpuyngu appeared as the majority language and Dhuwal-Dhuwala disappeared until the 2016 census when a new category Dhuwal (not further defined) appears with 77 speakers. For new languages the situation is even more difficult. Many people speak creoles or mixed languages which are far removed from English, but for which they do not have a name, or have a name which incorporates 'English', such as Camp English. Speakers of these new languages are severely undercounted, because they often report themselves as speaking a

kind of English/creole/pidgin/broken and/or their responses are misinterpreted (Angelo and McIntosh 2014).

A related problem comes with the denotation of names. For example, 'Kriol' is used as a name for a creole that is spoken in northern Australia, especially in the Kimberley and the Katherine area. Yumplatok/'Torres Strait Creole' is used as a name for a similar creole that is spoken in the Torres Strait and nearby diaspora areas in Queensland. Not all creoles which do have names are in the Census list of names (e.g. Yarrie Lingo for a creole used at Yarrabah). But the similarity of these names to each other and to the language type 'creole' have led to enormous confusion, e.g. the 2016 Census reports a large number of Kriol speakers (172) living in Yarrabah near Cairns on Cape York. At Napranum on Cape York, Mauritian Creole (33), Kriol (23) and English (601) appear in unlikely numbers. These rather odd-looking numbers reflect name confusions and indicate the unreliability of the figures for all new languages. Thus in Figure 1 we cannot assume accuracy or reliability regarding children reported as speaking Torres Strait Creole, or Kriol, let alone assume much difference over time between the rate of increase in

children reported as speaking Torres Strait Creole, and the rate of increase in children reported as speaking Kriol.

Fifth, when people and census-takers have trouble finding names for the ways they speak and linking them with names recognised in the census list of languages, they are assigned a residual code "Not further defined' ('nfd') codes linked to a larger group, which may be as large as Australian Indigenous Languages (given in Figure 1) or it may be much smaller. The nfd group may be geographic (e.g. Torres Strait Island Languages, nfd) or linguistic (e.g. Dhuwal, nfd). But there is a limited set of regional codes, and the regional code Australian Creoles nfd is no longer used.

The ABS have been refining their approach. For example a major review of languages classification (Australian Bureau of Statistics 2005) led to some changes in language names, in the addition of language names and consequently a sharp decrease in the nfd class between the 2001 census and the 2006 census. A further 48 Australian Indigenous languages were added to the classification in 2011, and a number were reclassified, which led to a small decrease in the nfd category. In 2016 one Australian Indigenous language was added and a number were reclassified. Figure 1 shows the decline in some nfd categories. Why there was such a large increase in the nfd in 2016 from 2011 is as yet unknown.

Recognising these concerns about the reliability of the Census data, there are nonetheless useful observations that can be made. The first relates to trends over time with respect to children. The number of children reported as speaking one of eight traditional languages with the most speakers has declined for all but Djambarrpuyngu (and the undercounting in the homelands of Djambarrpuyngu speakers has become notorious since the 2006 census (Kral and Morphy 2006, Morphy 2007)). This is a worrying trend[3,] which indicates an area for checking. The checking can also be seen in the context of the proportion of children to adults speaking an Australian Indigenous language. Table 1 shows this for the same eight languages along with two new languages. The higher proportion for the new languages, for Djambarrpuyngu and for Murrinh Patha reflect the anecdotal observations of linguists that children speak these languages as their first languages. Communities where the proportion drops below .5 could be investigated to see if the lower proportion is due to the children switching to other languages, and the nature of these new languages could be checked on, how they should be named, reported on etc.

	2001	2006	2011	2016
Kriol	0.64	0.62	0.56	0.50
Yumplatok (Torres Strait Creole)	0.52	0.62	0.61	0.54

	2001	2006	2011	2016
Djambarrpuyngu	0.51	0.58	0.59	0.45
Pitjantjatjara	0.44	0.44	0.38	0.33
Warlpiri	0.45	0.42	0.47	0.47
Murrinh Patha	0.70	0.72	0.64	0.57
Tiwi	0.46	0.46	0.43	0.33
Alyawarr	0.58	0.52	0.51	0.48
Kunwinjku	0.60	0.48	0.50	0.38
Anindilyakwa	0.51	0.49	0.46	0.34

Table 1: Proportion of children to non-children reported as speaking an Indigenous language. ABS, 2001–2016 Census of Population and Housing, Customised Data Report

A second area where Census data indicates areas to investigate is with respect to the location of speakers. The 2016 ABS census shows that more than half of the 2276 people who reported themselves as speaking Warlpiri were located outside of heartland Warlpiri country in the Central Desert (Table 2). Many were located in country adjacent to heartland Warlpiri country (Barkly, Western Desert) or in the nearest large town (Alice Springs, Katherine). Many of these people are likely to be part of the orbiting diaspora, spending time in the heartland as well as where they were recorded. But more than 150 people are living further afield in places from which it would be expensive to travel to Warlpiri heartland. Being scattered in small groups (6 in Perth, 9 in Sydney) means that, apart from phone, their opportunities to use Warlpiri in everyday talk is reduced. The small numbers mean that, while governments may give diaspora communities support for accessing services in English, they are unlikely to assist them with language maintenance or with accessing services in the diaspora language (for example in old age homes).

Location	Number	Distance fr. heartland
Melbourne	6	far
Perth	6	far
APY lands	7	far
Yarrabah	8	far
Sydney	9	far
Halls Creek	21	close
Santa Teresa	27	far
Adelaide	37	far
Western Desert	54	close

---

[3] However, in the Northern Territory where all but one of the traditional languages is spoken, the number of births to Indigenous parents has declined slightly (1688 in 2001, 1374 in 2016). Source: Aboriginal and Torres Strait Islander births

and confinements, summary, by state ABS stat.data.abs.gov.au/index.aspx?DatasetCode=ATSI_BIRTHS_SUMM, data extracted 2018-08-05

Location	Number	Distance fr. heartland
Katherine	55	medium
Darwin	67	far
Alice Springs	181	medium
Barkly	254	close
Kalkarindji	474	close
Central Desert	1,042	heartland
Australia	*2,276*	
Far= remote diaspora	*167*	

Table 2: Top Locations of people reporting they speak Warlpiri. ABS, 2016 Census of Population and Housing, Customised Data Report

Conclusion

Government will continue to use large-scale national datasets for forming policy on Indigenous language maintenance. It behoves linguists and language advocates to continue to test the reliability of this data against what is happening in individual communities, and to point out discrepancies with the aim of improving future survey collection. Despite the unreliability of certain types of data, age-coding and location of speakers can reveal trends relevant to language vitality and language maintenance, which can then assist members of speech communities in deciding where to focus their efforts.

Acknowledgements

References

Angelo, Denise, and McIntosh, Sophie. 2014. Anomalous data about Aboriginal and Torres Strait Islander language ecologies in Queensland. In Elke Stracke (Ed.) *Intersections: Applied linguistics as a meeting place.* 270–293. Newcastle on Tyne: Cambridge Scholars Publishing.

Australian Bureau of Statistics. 2005. Australian Standard Classification of Languages (ASCL), Second Edition Catalogue number 1267.0–2005 Released 2005-07-14

Australian Bureau of Statistics. 2016. Language Standards, Catalogue number 1200.0.55.005–2016 Released 2016-08-03

Australian Institute of Aboriginal and Torres Strait Islander Studies, and Federation of Aboriginal and Torres Strait Islander Languages. 2005. National Indigenous Languages Survey Report 2005: Report submitted to the Department of Communications, Information Technology and the Arts 255. Canberra: DCITA.

Dixon, Sally, and Angelo, Denise. 2014. Dodgy data, language invisibility and the implications for social inclusion: A critical analysis of indigenous student language data in Queensland schools. *Australian Review of Applied Linguistics* 37, 213–233.

Kral, Inge, and Morphy, Frances. 2006. Language. In Boyd Hunter (Ed.) *Assessing the evidence on Indigenous socioeconomic outcomes*, 279–289. Canberra: Centre for Aboriginal Economic Policy Research, The Australian National University.

Marmion, Douglas Edric, Obata, Kazuko, and Troy, Jakelin. 2014. Community, identity, wellbeing: the report of the Second National Indigenous Languages Survey. Canberra: Australian Institute of Aboriginal and Torres Strait Islander Studies.

Morphy, Frances ed. 2007. *Agency, contingency and census process: Observations of the 2006 Indigenous Enumeration Strategy in remote Aboriginal Australia*. CAEPR Research Monograph 28. Canberra: Centre for Aboriginal Economic Policy Research, The Australian National University.

Simpson, Jane. 2008. Language landscapes of children in remote Australia. In Robert Amery and Joshua Nash (Eds.), *Warra wiltaniappendi: Strengthening Languages: Proceedings of the Inaugural Indigenous Languages Conference (ILC), 24–27 September 2007* 51–60. Adelaide: University of Adelaide.

Simpson, Jane. 2013. What's done and what's said: Language attitudes, public language activities and everyday talk in the Northern Territory of Australia In *Journal of Multilingual and Multicultural Development: Special issue on attitudes towards endangered languages*, eds. Julia Sallabank and Peter Austin, 383–398.

Taylor, John, Bruce Doran, Maria Parriman, and Eunice Yu. 2014. Statistics for Community Governance: The Yawuru Indigenous Population Survey, Western Australia. *The International Indigenous Policy Journal* 5 (2). ir.lib.uwo.ca/iipj/vol5/iss2/2/ (accessed 2018-08-10)