Language contact in North Sulawesi: Preliminary observations

Timothy C BRICKELL

The University of Melbourne

Categorised as a Pidgin Derived Malay (PDM), *Manado Malay* (MM) is spoken throughout northern Sulawesi and on islands to the south of the southern Philippines. After originally functioning as regional *lingua franca*, it is now well established as the first language of up to one million people. This paper examines the language-contact situation between MM and two indigenous languages with a long presence in the region. Despite centuries of continued close contact, an examination of a range of typological features reveals minimal shared features, almost none of which have arisen through borrowing. These results corroborate multiple theories relating to language-contact outcomes, in particular the availability of different structural features for borrowing, the likely direction of any transfer, and the effect of both linguistic and non-linguistic factors on the potential for intense bilingualism.

1. Introduction[1]

Manado Malay (ISO 639-3:XMM) is a non-literary Malay variety. Also termed 'Vehicular' or 'Trade' Malays, these languages all share features which are not inherited directly from Proto-Austronesian (PAN) or Proto-Malayic (PM), but rather from a pidginised Malay ancestor language (Adelaar 2005:204). MM is spoken primarily in the *Sulawesi Utara* and *Gorontalo* provinces of northern Sulawesi.[2] The language and its direct precursor have been in continuous contact with regional indigenous languages since the 16th Century. This paper is a case study of this contact situation and its aims are threefold. Firstly, to identify any borrowed features, and the direction of borrowing, which are the result of contact between MM and two of the five languages of the Minahasan micro-group: *Tondano* (ISO 639-3:TDN) and *Tonsawang* (ISO 639-3:TNW). The second is to situate these results in the context of previously established theories of language-contact outcomes.

---

[2] The first province encompasses the islands of *Sangir* and *Talaud* to the north of North Sulawesi. In addition, MM is also spoken as a second language farther south in parts of Central Sulawesi (Paauw 2008:43).

The third and final aim is to explain the current contact situation with respect to linguistic and socio-historical factors. In addition to answering these questions, it is hoped that this investigation will contribute to the descriptive knowledge of these under-documented languages.

As regards what constitutes 'borrowing' or 'transfer', the definition used follows that of Matras & Sakel (2007), Sankoff (2004), Thomason & Kaufman (1988), and Van Coetsem (1988) in which borrowing is the importation of structural features (phonological, lexical, or morphosyntactic) from external (source) languages into native (recipient) languages. These features are considered as *Matter* (MAT)-type borrowing (Sakel 2007) whereby morphological material and its phonological shape are replicated. As regards conditioning factors, any transfer of structural features is judged to result from contact among languages, explicitly due to 'some level of bilingualism in the history of the relevant speech community'(Matras & Sakel 2007:1).

It will be demonstrated that in the languages under investigation, mutual features which result from borrowing through contact are negligible. In terms of phonology, while the respective phoneme inventories share many items, these are simply segments which are frequently occurring in Austronesian (AN) languages of the Malayo-Polynesian (MP) lineage, rather than the result of any phonological transfer. And while there are two marginally-occurring, non-native phonemes in MM, they are not the result of contact with indigenous languages. With respect to lexicon, MM has borrowed a small amount of vocabulary from TDN, but nothing from TNW. Any borrowed lexical items come from a single open-class category (nouns) and exhibit restricted semantic domains. In the other direction, a greater and more varied amount of MM lexicon is found in the speech of TDN and TNW bilinguals, although in some instances non-native lexical items are in fact single-word code switches. As for shared morphosyntactic features, these are effectively non-existent. While there is some overlap in the distribution of elements within noun phrases and two identical bound morphemes, this is not due to contact. Instead, the indigenous languages have maintained their rich and complex agglutinative morphology in contrast to the paucity of bound elements in MM. The two sets of languages have also preserved their traditionally divergent morphosyntactic alignment systems and strategies for encoding causation.

While somewhat preliminary, these results confirm established generalisations for language contact outcomes. The MM and indigenous speech communities appear to never have achieved long-running, intense bilingualism (see §6), thereby substantiating the notion that this is a fundamental requirement for large-scale borrowing (Thomason 2001:1640; Matras & Sakel 2007:2, 34; Bybee 2015:248). Moreover, levels of bilingualism and the direction of any borrowing are in large part conditioned by non-linguistic factors; the role, status, and domains of use of the languages are all highly relevant (Matras & Sakel 2007:2; Sakel 2007:23). In the event that borrowing does occur, phonological and lexical transfer are the most likely outcomes (Sankoff 2004). With respect to lexicon, words from open-class categories are most easily transferred (Poplack et. al. 1998; Poplack & Meechan 1998:127) and a cross-linguistically common distinction, that of so-called 'core' versus 'cultural' vocabulary, is identifiable. Despite this, ascertaining the exact source language can be problematic, as is differentiating loan words from single-word code switches.

In terms of morphosyntax, bound morphology and syntactic features are both highly re-

sistant to borrowing (Prince 1988; King 2000; Sankoff 2004:648). Moreover, this lack of morphosyntactic transfer in the absence of high levels of phonological or lexical transfer conforms to previously posited 'hierarchies' of borrowing (Whitney 1881; Haugen 1950; Moravcsik 1978; Matras 2007:32–33).

The subsequent sections are arranged in the following fashion: in §2 relevant linguistic, socio-linguistic, and socio-cultural information is provided in §2.1–§2.3. In §2.4 the data corpora are detailed together with some of the major sources. Sections §3–§5 then examine typological characteristics in order to identify and explain any shared features and the source and direction of any borrowing. In §3 phonological systems are investigated while §4 focuses on lexicon. Aspects of morphosyntax are examined in §5; phrase structure in §5.1 and morphology and morphological typology in §5.2. Finally, §6 summarises the findings and identifies the specific factors behind these language contact outcomes.

2. Background information

2.1 The languages of North Sulawesi

Displayed in Figure 1 (Glottolog 2019), the northern tip of the Indonesian island of Sulawesi (with its provincial capital, Manado) is located to the south of Mindanao, to the east of Kalimantan, and to the west of North Maluku.

Figure 1. The island of Sulawesi

All languages in the province are classified into the MP first-order branch of AN. Both the indigenous languages and PDM varieties are situated within the Western Malayo-Polynesian (WMP) subgroup of MP (Blust 2013:31–32). Traditionally, the province exemplifies a complex linguistic environment. Further to any PDMs, languages from three micro-groups are present - *Minahasan*, *Sangiric*, and *Gorontalo-Mondondow* (Blust 2013:82). In addition to Tondano and Tonsawang, the other three languages from the Minahasan micro-group (*Tombulu*, *Tonsea*, and *Tontemboan*), three from the Sangiric micro-group

(*Sangir*, *Bantik*, and *Ratahan*),[3] and a number from the Gorontalo-Mongondow micro-group (*Mongondow*, *Bintauna*, and *Ponosakan*) are spoken in the province. Over and above all these languages are other Malay varieties in the form of standardised and colloquial versions of the official language of Indonesia, *Bahasa Indonesia* (BI).[4] Although it lacks native speakers in North Sulawesi, the standardised version is ostensibly taught in all schools (Steinhauer 2005:69) and is used in state-level government administration, while colloquial varieties are ever present in popular culture and the mass media.

The traditional geographic divisions of a number of the indigenous groups in North Sulawesi are displayed in Figure 2 (Google Maps 2018 based on Henley 1996:xii).
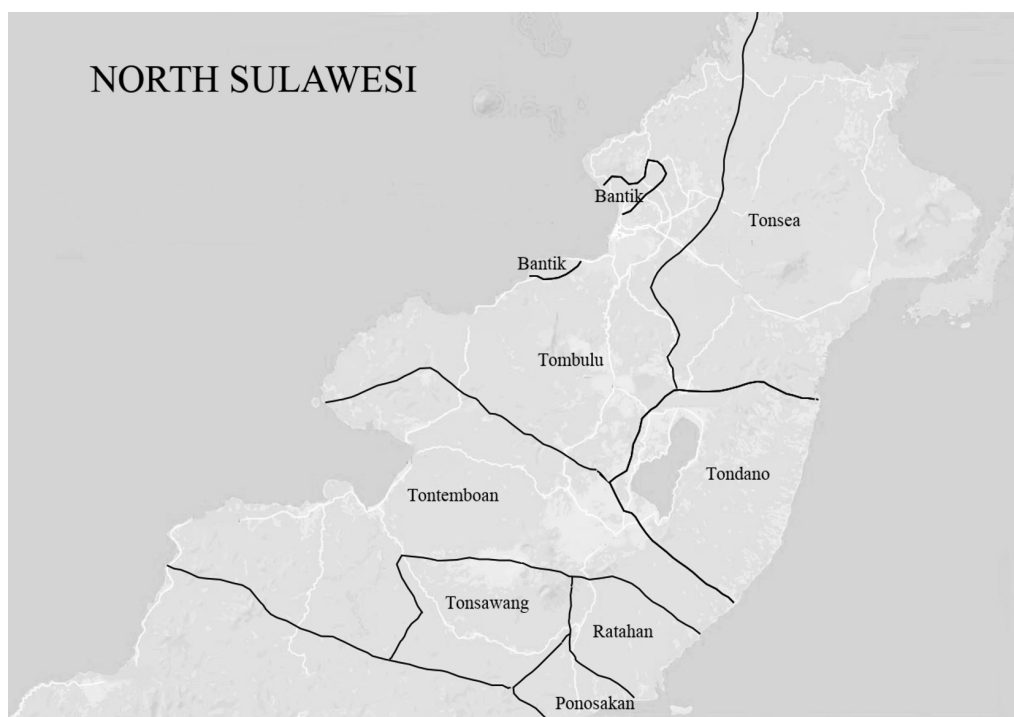
Figure 2. Ethnic and linguistic boundaries in North Sulawesi

---

[3] There are two additional languages in this micro-group which are not mentioned in this paper, *Sangil* and *Talaud*. The first is not included as it is spoken in the Philippines (Mindanao). The second, while spoken in the Talaud islands which are part of the province, is not included as it is not spoken in peninsular areas, unlike Bantik, Ratahan, and Sangir (the latter being spoken in both the Sangir islands and in peninsular areas).

[4] While MM, standardised BI, and colloquial BI varieties are all Malay languages, it is important to briefly clarify their divergent histories, typologies, status, and use. In contrast to PDMs such as MM, standardised Indonesian is derived from 'Literary' (or 'Classical') Malay varieties which were used in the court of the Riau-Johore Sultanate (based in areas of present-day Sumatra and Peninsular Malaysia) in the 1500s. In this environment a literary Malay variety developed which was subsequently utilised by colonial governments. This variety was influential in the development of the official languages of the modern-day Indonesian and Malaysian nation states (Adelaar & Prentice 1996:674). In terms of typology, varieties of BI are more morphologically agglutinating (especially standarised BI), exhibit verbal diathesis, and contain fewer lexical items of Portuguese and Dutch origin. Unlike PDMs, they have high prestige and expanded domains of use throughout the archipelago.

2.2 Manado Malay

MM belongs to the Malayic subgroup of WMP. It is most closely related to *Ternate Malay* (TM) (ISO 639-3:MAX) from which it is considered to have developed (Stoel 2005:8; Paauw 2008:21). From the mid-to-late 1600s TM (and later MM) was widely utilised by the Dutch colonial administration and the *Nederlandsch Zendeling Genootschap* (NZG) 'Dutch Missionary Society' (Schouten 1998:101) after being transported to the region from Ternate by military and administrative employees of the *Verenigde Oost Indische Compagnie* (VOC) 'United East India Company' (Watuseke & Watuseke-Politton 1981:326) – the so-called *borgo* (from Dutch *burger* 'citizen'). It is also highly probable that contact with TM began in the region prior to any European settlement (Paauw 2008:44), due to the sporadic influence of the powerful Sultanate of Ternate (Henley 1993). It was also present to some degree from the early 1500s when the Portuguese administered the Maluku islands prior to Dutch control. The result of this contact with various powerful groups originating from Ternate is that, in one form or another, a PDM has been in contact with indigenous languages in North Sulawesi for approximately 500 years.

After TM arrived in North Sulawesi, MM developed independently. Currently, it is the language with by far the greatest number of speakers in North Sulawesi and is dominant in almost all domains of use (Mead 2013; Brickell 2018a). Native speakers are estimated to number between 850,000–1,000,000 (Stoel 2005:6; Simons & Fennig 2019). In terms of combined first and second-language speakers, Whisler (2006) provides an unverified figure of up to 3,000,000.

Over the last 200 years the geographical spread of MM has increased rapidly. Historically, native speakers were confined to the larger cities of Manado and Bitung (Paauw 2008:42). As recently as 1925 it was proposed that MM had no native speakers at all (Adriani 1925:142), a situation that, even if true, completely changed in the following generations. MM is now essentially the only native language of a large proportion of the region's population.[5] In contrast, speakers who are bilingual in MM and indigenous languages are usually only found within the older generation.

2.3 Tondano and Tonsawang

TDN and TNW represent two of the five languages of the Minahasan micro-group.[6] At a higher-order level, the Minahasan micro-group belongs to the Philippine subgroup (Zorc 1986; Blust 1991, Blust 2013:82, 740) of WMP.

In terms of speech community locations (Figure 2), TDN speakers historically reside in a more northern part of the province, close to the lake and town of Tondano and in vil-

---

[5] The notion of a monolingual speaker requires some clarification. While not researched in any detail, it is noted in Paauw (2008:44) and Shiohara & Jukes (2018:116) that, due to the constant presence of varieties of Indonesian in various domains, there is currently some convergence between BI and MM. The former publication attests observational evidence as showing that speakers may sometimes employ BI vocabulary and constructions while speaking MM. The latter states that some MM speakers do not have the meta-awareness to know that the language they speak is not BI. In terms of my own personal experience during fieldwork, while I have on occasion observed the first situation, I regard the second as too strong a statement.

[6] The existence of this micro-group and the position of the languages within it is primarily the result of the research of Sneddon (1975, 1978). TDN is one of three languages in the 'Northeast' branch while TNW is a direct descendant of the proto-language (Sneddon 1975:8–9).

lages towards the east coast. In contrast, the traditionally smaller and more isolated TNW community (Sneddon 1978:5; Brickell 2018a) is situated farther south, in and around the town of Tombatu and in approximately a dozen villages in somewhat mountainous terrain far from both coasts.

Published speaker numbers for both languages are outdated and unreliable. Figures of 92,000 for TDN and 20,000 for TNW are found in Simons & Fennig (2019), based on Wurm & Hattori (1981). However, due to the continuing and increasingly-rapid shift of all indigenous languages to MM, and a well-established break in inter-generational transmission (Merrifield & Salea 1996; Mead 2013; Hertz & Lee 2017; Brickell 2018a), actual speaker numbers are almost certainly far lower.

As regards linguistic vitality, this is worryingly low for all Minahasan languages (Merrifield & Salea 1996; Mead 2013), including TDN and TNW (see Hertz & Lee 2017 and Brickell 2015, 2018a for further detail). Domains of use are extremely limited and native speakers are now primarily restricted to the older generation (50–60 years old and above). Fluent speakers in the younger generation are rare. Fieldwork observations attested only a handful of speakers in this demographic, all of whom were inhabitants of remote villages. These younger community members were a rarity in that they had not followed the usual trend of leaving the speech area to pursue education or employment opportunities.

2.4 Data corpora and major sources

All languages spoken in North Sulawesi are under documented. Previous descriptive research on TDN is limited to the phonology and grammar sketch of Sneddon (1975) and the more recent descriptive grammar of Brickell (2015). The TDN data used for this study are taken primarily from the corpus which informed the latter work, part of which is available online via the Multilingual Corpus of Annotated Spoken Texts (Multi-CAST) archive at the University of Bamberg (Brickell 2016a). A small number of examples are taken from other available sources, including Bible translations (Tondano to Indonesian or Manado Malay) and the Tondano-Manado Malay-Indonesian dictionary of Dotulong (2010).

Grammatical descriptions relating to TNW are even rarer. Information on phonology and morphology is included in the reconstruction of Proto-Minahasan undertaken by Sneddon (1975). The most recent documentary work, which provides the data for this study, comprises the archived corpora of Utsumi (2014) and Brickell (2016b) which were deposited in The Pacific and Regional Archive for Digital Sources in Endangered Cultures (PARADISEC) and the Endangered Languages Archive (ELAR) respectively. Descriptions of TNW morphosyntax informed by these corpora are currently limited to the article of Brickell (2018b), the conference presentation of Utsumi (2018), and the recent MA thesis of Hayes (2019). Further to these, a modicum of data are sourced from the Tonsawang-Indonesian dictionary of Kalangi (2012).

MM is somewhat better documented. However, in the absence of annotated and archived primary linguistic data, information and language examples were sourced from the publications of Prentice (1994), Stoel (2005), Paauw (2008), and Dotulong (2010) before being checked with native speakers.

3. Phonology

An overview of the three phoneme inventories, their similarities and differences, and the identification of any problematic marginal phonemes is presented in §3.1. §3.2 then discusses the potential for any shared segments due to language contact.

3.1 Phoneme inventories

The three languages share a large proportion of consonant and vowel segments.[7] All contain voiced and voiceless plosive pairs and nasals at three places of articulation: /p/, /b/, /t/, /d/, /k/, /g/, /m/, /n/, and /ŋ/. Also in common are the fricative, trill, lateral, and approximants /s/, /r/, /l/, /j/, and /w/. Palatal consonants are the primary divergence – only MM has the voiced and voiceless plosives /c/ and /ɟ/ and the nasal /ɲ/. A labiodental fricative /f/ is also only found in MM. Two of the three, MM and TNW, contain the glottal fricative /h/, while only TDN has an unstable velar approximant /ɰ/. Finally, in contrast to marginal occurrences in MM, the glottal plosive /ʔ/ is attested as fully phonemic in all Minahasan languages.

As for vowels, all languages feature the high, high-mid, and low-front vowels /i/, /e/, and /a/, and the high and high-mid back vowels /u/ and /o/. The sole disparity relates to the mid-central schwa; attested for both TDN and TNW in contrast to an uncertain status in MM.

/ʔ/ is accorded full phonemic status in MM by Prentice (1994:413). The analysis of Stoel (2005:11) differs, stating that /ʔ/ is phonemic only in word-final position, while Paauw (2008:81–82) also finds insufficient evidence for full phonemic status. My own observations during fieldwork match the latter's analysis in that /ʔ/ is restricted to a word-internal position (intervocalically) as well as being weakly articulated word finally,[8] for instance in proper nouns (surnames) such as *Paat* ([paʔat]) or *Polii* ([poliʔi]) and in a small number of food or place names such as *tinutuʔanʔ* 'vegetable porridge', *rintek wuʔuk* (r.w.) 'fine hair',[9] *Tumpaʔan*, and *Tanawangkoʔ*.

/ə/ is described by Stoel (2005:12) as present in MM without further clarification. Paauw (2008:95–96, 356) describes it as not infrequent but does not afford it full phonemic status. Prentice (1994:414ff) points out that TM lacks /ə/, an analysis also found in Voorhoeve (1983) and Litamahuputty (2012:15). Prentice (1994) then identifies the environments in which /ə/ has been replaced, retained, and possibly reintroduced in words which are primarily of Malayic and Dutch provenance. This is the same distribution pattern I have observed, for example, /ə/ is found in Dutch-derived words such as *menèr* (*mijnheer*) 'sir', *wastafel* 'sink', and *halte* '(bus) stop'. It is also observed in English-derived words, for instance *mister* 'sir, mister', *filem* 'film', and *telpon/nelpon* 'telephone'. /ə/ also alternates with /a/ in pairs of words which are Malay-derived, such as *sebla* (*sabla*) 'next', *bersi* (*barsi*) 'clean', *perna* (*parna*) 'ever', *kerja* (*karja*) 'work', and *keluar* (*kaluar*) 'go

---

[7] In terms of orthography, the conventions used in the speech community are adhered to and a number of IPA characters are expressed as follows: <ng> is used to represent /ŋ/, <ny> for /ɲ/, <'> for /ʔ/, <gh> for /ɰ/, <y> for /j/, <j> for /ɟ/, <è> for /e/, <e> for /ə/, <èi> for /ej/, and <ou> for /ow/.

[8] Paauw (2008:81) states that the glottal also occurs word-initially before vowels. This distribution is not attested in this study, nor in those of Prentice (1994) and Stoel (2005:12).

[9] A euphemism for dog meat.

out'. Words of the same origin may also contain /ə/ in environments where it was previously deleted, resulting in previous word-initial consonant clusters which are now broken up by /ə/: *sekola* (*skola*), 'school' *belum* (*blum*) 'not yet', *terus* (*trus*) 'continue', and *selop* (*slop*) 'sandals'.

3.2 Phonological borrowing

As a generalisation, the basic phonological systems of AN languages are reasonably uniform (Himmelmann 2005:115). Despite the high degree of similarity in the phoneme inventories of the three languages, this is unlikely to be the product of phonological borrowing and instead reflects a shared AN and WMP heritage. Furthermore, most of the segments for each language adhere to expected typology or to reconstructions for Proto-languages; MM reflects those commonly attested to contact varieties of Malay (Paauw 2008:80) while almost all TDN and TNW phonemes match those reconstructed for Proto-Minahasan (Sneddon 1978).

Of the shared phonemes, only those with an uncertain phonemic status, /ə/ and /ʔ/, appear to be candidates for transfer, with both potentially borrowed into MM due to long-term contact with the indigenous Minahasan languages. This is the unsubstantiated analysis posited in Prentice (1994:423) and Paauw (2008:95).

In order to test this hypothesis using previous theories of language contact, the notion of transfer occurring due to either 'adoption' or 'imposition' (Van Coetsem 1988; Haspelmath 2009:50–51) is applicable as a diagnostic. Borrowings in the first category describe native speakers of the recipient language adopting features from a dominant source language into their own, while those in the second involve non-native speakers unintentionally retaining features of their mother tongue when shifting to another language. Using this distinction, the presence of /ʔ/ and /ə/ in MM is explained in one of two ways: either TDN and TNW speakers, who are also non-native speakers of MM, are retaining /ʔ/ and /ə/ as they shift, i.e. imposition, or MM speakers are adopting /ʔ/ and /ə/ due to contact with the dominant indigenous languages, i.e. adoption.

Neither of these options are plausible. Firstly, imposition requires a large number of non-native speakers of MM. But in reality, with the exception of small groups of immigrants,[10] there are essentially no non-native speakers of MM in the province. In the second instance, adoption requires indigenous languages to be somehow dominant. However, the indigenous languages are, and have always been, lower in prestige than MM and in no way dominant.

The alternative scenario must therefore be that the source language is external to North Sulawesi. The fact that /ʔ/ and /ə/ are unambiguously attested in two languages which have had overwhelming official support and prestige post-independence, Indonesian (Ewing

---

[10] Citizens from other parts of the archipelago have settled in North Sulawesi primarily due to the *transmigrasi* 'trans-migration' program which was initiated by the Dutch colonial government and continued by successive Indonesian governments. By 1980 this program had resettled almost one million people to outer regions (World Bank 1988:xviii), many of whom came from Java or Sumatra. However, North Sulawesi received relatively few of these migrants. A report by the World Bank gives a total of 18,817 migrants resettled in the province between 1970–1986 (World Bank 1988:211). By 1986 these people comprised just 0.9 percent of the population.

2005:229) and Javanese (Ogloblin 2005:592–593), provides two potential sources. Under this analysis the presence of /ʔ/ and /ə/ is easily explained as a case of adoption by MM speakers. While speculative, borrowing from one or both of these sources presents a feasible scenario. Varieties of Indonesian have strong influence in domains such as the state-wide education system, government administration, and in most mass media consumed by native MM speakers (Prentice 1994:417; Paauw 2008:44). Javanese also has a frequent presence in the public domain due to the fact it is one of the most widely spoken languages in Indonesia (Ogloblin 2005:591).

4. Lexicon and lexical borrowing

Along with the transfer of phonological features, the borrowing of lexical items is a frequent outcome of language-contact situations (Sankoff 2004:644), so much so as to be considered essentially universal (Tadmor 2009:55). In spite of this, discovering the exact status of these items and their origins is not always easy, and differentiating them from instances of code switching can be problematic (Poplack & Meechan 1998:128; Haspelmath 2009:40).

In contrast to phonological features, there is greater evidence of lexical borrowing between the three languages. In §4.1– §4.2 aspects of lexicon are discussed and examples of non-native vocabulary are identified. Explanations for the presence of non-native lexical items are then provided in §4.3.

4.1 MM lexicon

Throughout their development, TM and MM have had contact with a variety of powerful external groups, both European (Portuguese and Dutch) as well as those closer to North Sulawesi (the Ternate Sultanate and the Java-centric Indonesian nation state). This contact is reflected in much of MM's lexicon. With TM as one of the earliest trade Malays (Paauw 2008:11) based in Ternate during Portuguese rule (1512–1575), Portuguese-derived words as well as those from the non-AN indigenous language of Ternate (ISO 639-3:TFT) are present. Lexicon of Dutch origin is also prevalent due to the widespread use of TM and then MM during the colonial period (1658–1942). Further to this, continuous contact with various Malay-speaking groups has resulted in words with cognates in other Malayic varieties.

Table 1 and Table 2 provide examples of lexical items (both open class and closed class) from these source languages.[11]

---

[11] In modern MM the proportion of lexicon from these different source languages is unknown. To give an idea of the ratio of non-Malay words in TM, in a list of 563 lexical items Voorhoeve (1983:12) attests 92 to Ternate, 34 to Dutch, and 11 to Portuguese.

Table 1. MM lexicon of Dutch and Portuguese provenance

Dutch:	English Gloss:	Portuguese:	English gloss:
broit (bruid)	'bride'	*lènso (lenço)*	'handkerchief'
for (voor)	'so.that'	*frèsko (fresco)*	'fresh'
strom (stroom)	'stream'	*mai (mãe)*	'mother'
lèpèr (lepel)	'spoon'	*soldado (soldado)*	'soldier'
klaar (klaar)	'finish'	*tuturuga (tartaruga)*	'turtle'
flao (flauw)	'faint'	*milu (milho)*	'corn'
wastafel (wastafel)	'sink'	*fastiu (fastio)*	'bored'
opstok (opstoken)	'incite'	*suar (suar)*	'sweat'

Table 2. MM lexicon of Malay and Ternate provenance

Malay:	English Gloss:	Ternate:	English gloss:
makang (makan)	'eat'	*popolulu (polulu)*	'round'
abis (habis)	'finish'	*ngana (ngana)*	'you (2SG)'
tutu (tutup)	'close'	*ngoni (ngon)*	'you (2PL)'
balè (balik)	'return'	*dabu (dabu)*	'condiment'
so (sudah)	'already'	*batobo (matobo)*	'swim'
ampèr (hampir)	'almost'	*gomutu (gomutu)*	'palm leaf fibre'
maso (masuk)	'enter'	*pongo (pongo)*	'deaf'
jatung (jatuh)	'fall'	*goraka (guraka)*	'ginger'

MM also contains words which are derived from none of these major sources. These appear to originate from languages of the larger Minahasan ethnic groups, including TDN. In contrast, there are currently none identified which originate from TNW. Minahasan-derived non-native vocabulary in MM is restricted to open-class words with limited semantic domains; primarily nouns referring to people or place names, food, plants, or animals.

Table 3 is a list of indigenous Minahasan lexicon attested in the speech of MM speakers (mono- and bilingual) together with any corresponding MM items. Rows 1–6 denote cuisine and/or animals, rows 7–11 refer to aspects of pre-Christian culture and spirituality, and rows 12–19 are proper nouns (either place names or exonyms and endonyms for different ethnic groups).

Table 3. Indigenous Minahasan lexicon in MM

Non-native lexicon:	Source:	MM:	English gloss:	Semantic domain:
tinutu'an*	TDN	*bubur manado*	'vegetable porridge'	Food
tinoransak	TDN	-	'pork and spices in bamboo'	Food
binyolos	TDN	-	' (type of) round cake'	Food
manguni*	TDN	*loyot ~ otot*	'owl'	Animal
rintek wu'uk (r.w.*)	TDN	*anjing*	'dog'	Animal/Food
kolombi ~ kelobi'*	TDN	*bia (basaar)*	'snail'	Animal/Food
opo*	TDN	*tuhan*	'ancestor (God)'	People
toar*	TDN	-	proper noun	People
lumimu'ut*	TDN	-	proper noun	People
waruga*	TDN	*kubur*	'tomb'	Place
kawasaran*	TDN	-	'war dance(er)'	People
tondano*	TDN	-	proper noun	People
tombulu*	TDN	-	proper noun	People
tonsèa*	TDN	-	proper noun	People
tontèmboan*	TDN	-	proper noun	People
tonsawang*	TDN	-	proper noun	People
kiniar*	TDN	-	proper noun	Place
ranowangko'*	TDN	-	proper noun	Place
wè'wèlan*	TDN	-	proper noun	Place

Multiple diagnostics confirm the origin of these words, for example many are cognate with those in other languages of the micro-group (as indicated by '*') and a number, for example *manguni* and *rintek*, are reconstructed for Proto-Minahasan and/or Proto-Philippine (Sneddon 1978:157ff). Another distinctive feature is a higher level of morphological complexity than is observed in MM words. Minimally, twelve of the nineteen are deconstructable into their component morphemes, as demonstrated in Table 4.[12]

---

[12] All bound morphemes, which include TAM (e.g. *<in>*) and voice-marking (e.g.*-an*) affixes, are well attested in AN languages. Additionally, in most cases the lexical root is clearly identifiable. One exception is *tinoransak* – in discussions with speakers the root was given as either *ransak* or *toransak*. *Binyolos* is also problematic. While many speakers consider it as indigenous to the Minahasan languages, none could identify a root. Nor is it found in the available dictionary. While it is included as a Minahasan loan into MM, its origins are uncertain, and may in fact lie further afield, when considering Portuguese *bolo(s)* 'cake(s)'.

Table 4. Morphological complexity in Minahasan lexicon

Lexical item:	Component morphemes:	Literal gloss:
tinutu'an	*t\<in>utu'-an*	\<PST>stew-LV
tinoransak	*t\<in>oransak*	\<PST>assemble
binyolos	*b\<in>yolos*	\<PST>cake
kawasaran	*kawasal ni=sarian*	follow AN.SG.GEN=dance leader
tondano	*tou N=rano*	person INAN=water
tombulu	*tou N=wulur*	person INAN=mountain
tonsèa	*tou N=sèa*	person INAN=diverge
tontemboan	*tou N=tèmboan*	person INAN=look.out
tonsawang	*tou N=sawang*	person INAN=help
kiniar	*k\<in>iar*	\<PST>excavate
ranowangko'	*rano wangko'*	water big
wè'wèlan	*wè'wèl-an*	tap.palm.branch-LV

4.2 MM lexical borrowing

The words in Table 3 exemplify one of the lexical categories most frequently borrowed via contact, i.e. nouns (Sankoff 2004:644) referring to animals, food and drink, the physical world, or religion and beliefs (Myers-Scotton 2006:226; Tadmor 2009:59–64). While this much is clear, pinpointing the exact reason(s) for the importation of non-native lexicon is rarely straightforward. One method of categorising potential loan words is the 'cultural' versus 'core' vocabulary distinction (Myers-Scotton 2002:41). The former refers to new or foreign concepts for which there are no corresponding indigenous words, while the latter describes duplicate lexicon for concepts already expressed with pre-existing native words. The two categories can be further distinguished by the speed at which borrowings appear in a recipient language. Cultural borrowings may occur abruptly in the lexicon of speakers, either mono- or bilingual. Core borrowings, on the other hand, often occur more gradually; an unverified hypothesis is that they enter the speech of bilinguals as single-word code switching before spreading (Myers-Scotton 2002:41; 2006:254) to monolingual speakers.

These two categories are suitable for analysing non-native lexicon in MM, up to a point. Many of the words in Table 3 which lack corresponding MM counterparts are legitimate candidates for cultural borrowing. They are the result of the daily contact between indigenous and non-indigenous (*borgo*) groups which occurred during the colonial period.[13] Many of the VOC *borgo* employees who imported TM to the region were not native to North Sulawesi and were instead of Asian or Eurasian origin (Schouten 1998:49). For this group, whose descendants would become the first generation of native MM speakers, concepts relating to indigenous flora, fauna, food, culture, and spirituality were unlikely to be familiar. Regardless of the *borgo*'s more dominant socio-economic status, their daily contact with the indigenous population meant that non-native lexicon from common semantic domains had high borrowing potential. Cultural borrowings such as these

---

[13] In contemporary society any historical differences between the two groups have essentially disappeared, as have many distinctions between different indigenous groups. The descendants of the *borgo* are now completely integrated into broader Minahasan society.

were then incorporated into the speech of their descendants, who may or may not also have been speakers of indigenous languages. For modern-day monolingual MM speakers of non-*borgo* origin, these words were acquired directly through MM and/or potentially from native speakers of indigenous languages.

Accounting for the six pairs in Table 3 which have native and non-native counterparts is more problematic, as it requires an answer as to why speakers import lexicon when their native language already has pre-existing vocabulary for the same concept. One potential hypothesis is that these core borrowings have developed, or are in the process of developing, through code switching. To differentiate code switches from loan words, the criteria of 'frequency' and 'variability' (Myers-Scotton 1993:191–204) can be utilised. The first pertains to the idea that if concepts are solely expressed by non-native lexical items then these are loan words. Alternatively, if both non-native and native lexicon are used then single-word code switching is more likely. In addition to the code-switching-to-loan-word scenario for core borrowings, frequently posited, though sometimes vaguely-defined explanations are that these occur due to 'cultural pressure' (Thomason & Kaufman 1988:77) or 'prestige' from dominant linguistic group(s), with the latter undoubtedly highly relevant to language change (Haspelmath 2009:48).

A scenario in which MM speakers have incorporated core borrowings from Minahasan languages due to perceived prestige or cultural pressure is improbable. Another analysis, albeit hypothetical, is that these loans have entered the speech of monolinguals following a period of functioning as single-word code switches in the speech of older bilinguals. However, this explanation is complicated by the fact that most of the six pairs in Table 3 fail the criterion of variability. Under observation, only *tinutu'an* and *bubur manado* and *kolombi ~ kelobi'* and *bia (basaar)* are used interchangeably in unrestricted contexts. In all other pairs, the non-MM word is less frequent and has limited scope. To clarify, rather than broadly defining a genus of bird, *manguni* 'owl' chiefly refers to the culturally-significant animal[14] found on the official logo of administrative districts, as well as in the name of a group which originally functioned as a paramilitary unit.[15] Similarly, *r.w.* 'dog' refers solely to the traditional local dish made from this animal, while *waruga* and *opo* are only relevant in the context of pre-Christian culture and spirituality. In any other contexts, the MM words *loyot ~ otot*, *anjing*, *kubur*, and *tuhan* are normally used.

In part, the difficulty in classifying these potential borrowings stems from the lack of cross-linguistic, predefined one-to-one correspondences in the concepts expressed by nouns. But it also demonstrates the limitations of the core versus cultural distinction, as these words do not fit neatly into either category. The conclusion reached here is that, for bilingual speakers these pairs are better described as single-word code switches. For MM monolinguals, however, the non-MM items in these pairs represent a somewhat indistinct type of borrowing; they are core in the sense that there is MM lexicon for the same or similar concepts, but cultural in the sense that some items represent concepts which are now essentially 'foreign' for many younger monolinguals, i.e. pre-Christian culture and

---

[14] The owl is important in pre-Christian Minahsasan mythology and animism (Renwarin 2006:119, 136, 300).

[15] Known as the *Brigade Manguni Indonesia*, or BMI, this organisation was founded in 2000 in response to the perceived threat of invasion from Muslim extremists during a time of religious-based conflict in several nearby districts of Maluku (Bakker 2016:256).

spirituality.

4.3 TDN and TNW lexicon

TDN and TNW share varying amounts of vocabulary with each other and with the other Minahasan languages,[16] including reflexes of the Proto-Minahasan lexicon reconstructed by Sneddon (1978:120–183). A broad range of non-Minahasan lexicon is also observable and, at first glance, there appears to be a much higher degree of lexical borrowing due to contact than is the case for MM. Tables 5 and 6 present some open-class and closed-class non-native words (and any corresponding indigenous words) which are currently attested in the speech of bilinguals.

Table 5. Non-native lexicon in TDN and TNW (open-class items)

Non-native lexicon:	TDN:	TNW:	English gloss:
klappertaart	-	-	'coconut cake'
krois	-	-	'cross (Christian symbol)'
lèpèr	-	-	'spoon'
lènso	-	-	'handkerchief'
foto	-	-	'photograph'
sopir	-	-	'driver'
motor	-	-	'motorbike'
fidio	-	-	'video'
tèlèfisi	-	-	'television'
kètring	-	-	'cater'
malang	*wengi*	*bengi*	'night'
loyang	*reranoan*	*sambat*	'container (liquid)'
dèsa	*wanua*	*do'ong*	'village'
kabupaten	*walak*	*walak*	'district'
iris	*iwu*	*sosow*	'slice'
kumpul	*erur*	*sesen*	'assemble'
kaluar	*odol*	*luay*	'go.out'
pikir	*ghènang*	*dèkèn*	'think'

---

[16] As the most conservative and genetically-divergent language, TNW shares less vocabulary with the other languages than TDN. Furthermore, the TNW speech community's location adjacent to the non-Minahasan Ponosakan and Mongondow speaking areas has resulted in borrowing from these languages (Sneddon 1978:10).

Table 6. Non-native lexicon in TDN and TNW (closed-class items)

Non-native lexicon:	TDN:	TNW:	English gloss:
jadi	*tuana*	*kèlè (wa')aho*	'thus'
sèrta	*tarèala*	*songad*	'after'
so itu	*ka'a*	*ndè*	'because'
for	*rior*	*saniboho*	'so.that'
dan	*wo*	*bo*	'and'
ato	*ka'apa*	*suma*	'or'
mar	*ta'an*	*sumata'*	'but'
banya	*laker*	*tado'o*	'many'
samua	*waya*	*pahasa*	'all'

In contrast to Table 3, the items in Table 5 and Table 6 encompass a range of lexical categories, including nouns, verbs, coordinating and subordinating conjunctions, and quantifiers. Classifying these items as non-native is unproblematic. They lack the morphological complexity often seen in Minahasan lexicon and a Malay, Dutch, Portuguese or English etymology is attributable for most. Many are also listed as MM entries in dictionaries such as that of Dotulong (2010). However, beyond this point explicitly identifying the source language(s) is less straightforward. Many of these words are also found in varieties of Indonesian and it is difficult to confirm if there has been lexical transfer directly from MM, directly from varieties of Indonesian or other external languages, or both.

4.4 TDN and TNW lexical borrowing

In Table 5, the first ten words are used exclusively and have no corresponding indigenous counterparts. They are categorised as cultural borrowings which are loans of convenience for new(er) or foreign concepts. However, within this group a distinction must be made. The first four, *klappertaart* 'coconut cake', *krois* 'cross', *lèpèr* 'spoon', and *lènso* 'handkerchief', are not found in Indonesian and express concepts which, relatively speaking, are less modern than those of the following six words. All have Dutch or Portuguese origin and all conceivably result from colonial-period contact with Dutch and/or TM speakers. They refer to what were at the time foreign concepts introduced to the indigenous population. In contrast, the following six, all of which are also present in Indonesian, *foto* 'photograph', *sopir* 'driver', *motor* 'motorbike', *fidio* 'video', *tèlèfisi* 'television', and *kètring* 'cater', refer to much more recent technologies. Nonetheless, despite being obvious cultural borrowings, it is not clear if they should be classified as: transfers into indigenous languages via Indonesian and then MM, transfers directly from Indonesian, or, due to the fact that all speakers of indigenous languages are also native speakers of MM, simply transfers from Indonesian into MM.

The remaining non-native lexical items in Table 5 and Table 6 have attested indigenous counterparts, thereby raising the possibility of core borrowing or code switching. In this case the criteria of frequency and variability function well as diagnostics, as there is no attested variation in semantics or contexts of use. Each word in these pairs is used interchangeably to refer to the same lexical or grammatical concept. These are therefore best described as single-word code switches.

Although appropriate for classifying these items as code switches, the criterion of vari-

ability does not take into account the frequency of use for the items in each pair. To provide a snapshot of the relative frequencies of the items in one pair, Table 7 presents a comparative count of one subordinating conjunction ('but'), *mar* (MM) and *ta'an* (TDN) and *mar* and *sumata'* (TNW), in seven recordings.[17]

Table 7. Comparative frequency of native and non-native pairs

Recording:	Speaker (age):	Native:	Non-native:
TNW_03_SM_JP_26062016	SM (42 years)	*sumata'* (0)	*mar* (4)
TNW_03_SM_JP_26062016	JP (60 years)	*sumata'* (2)	*mar* (0)
TNW_08_JP_07072016	JP (60 years)	*sumata'* (4)	*mar* (1)
TNW_06_OK_VA_03072016	OK (40 years)	*sumata'* (0)	*mar* (13)
TNW_06_OK_VA_03072016	VA (28 years)	*sumata'* (0)	*mar* (4)
TNW_10_RP_26062016	RP (58 years)	*sumata'* (7)	*mar* (0)
TDN_11_AW_HL_27082011	AW (68 years)	*ta'an* (3)	*mar* (0)
TDN_11_AW_HL_27082011	HL (72 years)	*ta'an* (6)	*mar* (0)
TDN_28_OL_KK_25092012	OL (71 years)	*ta'an* (7)	*mar* (0)
TDN_28_OL_KK_25092012	KK (45 years)	*ta'an* (15)	*mar* (0)
MAPALUS	AW (68 years)	*ta'an* (5)	*mar* (0)

While admittedly limited in scope, this comparison displays a clear correlation between age demographic and frequency of use, that is, older and more fluent speakers are more likely to use native lexical items. The exception to this generalisation is one middle-aged speaker KK (row 10). The less frequent use of indigenous lexicon by comparatively younger bilinguals is best viewed in the context of the low linguistic vitality of the indigenous languages. What is more, it is possible that even older and more fluent bilinguals may progressively cease using their minority language in favour of solely using the dominant majority language, a scenario previously posited by Myers-Scotton (2002:51).

5. Morphosyntax: morphology and NP structure

In this section morphosyntactic aspects of the languages are examined, beginning with NP structure in §5.1. Morphological typology and the potential for morphological borrowing are then discussed in §5.2.

5.1 NP structure

In terms of constituent order, previous literature lacks agreement on the probability that this level of structure may change through contact. Thomason (2001:1640) proposes that it is one of the most easily observable results of language contact, while Myers-Scotton (2002:202–203) states that the modification of clause-level constituents is possible. In contrast to these viewpoints, the results of the cross-linguistic study of Matras (2007:60) fail to find any frequent contact-induced change in word order.

The examination of constituent order in the three North Sulawesi languages focuses on

---

[17] These comprise four recordings in the Tonsawang language (two monologues and two dialogues totaling 57 minutes) and three Tondano recordings (one monologue and two dialogues totaling 32 minutes).

NPs. To begin, NP structure for MM is outlined by Prentice (1994:424–429) and Shiohara & Jukes (2018:116) as:

<div align="center">

(article) HEAD (modifier)

</div>

This default distribution, whereby articles precede the head and any modifiers follow it, is demonstrated by (1a)-(1d).

(1)   a.   *tu=   ruma itu*
           ART= house that

           'That house.'   (Shiohara & Jukes 2018:117)

    b.   *ni=   pulo ini*
           ART= island this

           'This island.'   (Shiohara & Jukes 2018:117)

    c.   *oto banya*
           car many

           'Many cars.'   (Paauw 2008:324)

    d.   *torang dua*
           1PL   two

           'We two (two of us).'   (Paauw 2008:581)

Paauw (2008:324) attests the same basic NP structure. However, both Paauw (2008:324) and Shiohara & Jukes (2018:116) provide the caveat that the distribution of quantifiers and numerals is flexible. These may also precede the head, as exemplified by (2a)-(2c).

(2)   a.   *samua ni=   pulo*
           all   ART= island

           'All the islands.'   (ELICITED)

    b.   *banya buku*
           many   book

           'Many books.'   (Paauw 2008:325)

    c.   *tiga bulan*
           three month

           'Three months.'   (Paauw 2008:368)

To compare the MM structure with that of the two Minahasan languages, the description of Sneddon (1975:114) asserts that in TDN all modifiers with the exception of phrase markers follow the head. Watupongoh et. al. (1992:50–60) instead attest phrase markers preceding the head, quantifiers and numerals either preceding or following the head, and demonstratives always occurring post-head. The latter structure is also outlined in

Brickell (2015:266–268, 312) and is found in the current data corpus. It is summarised as:

<div align="center">(modifier) (phrase marker) HEAD (modifier) (DEM)</div>

The position of phrase markers, quantifiers, and numerals as pre-head is demonstrated by examples (3a)-(3c). In (4a)-(4c) quantifiers, numerals, and demonstratives follow the head.

(3)   a.   *siokiku*
           si=oki'=ku
           AN.SG=small=1SG.GEN

           'My child.'   WATULANEY_01092011_00:05:42

      b.   *laker   nuka*
           laker   N=uka
           many   INAN=coconut.shell

           'Many coconut shells.'   KINIAR02_14052013_00:09:52

      c.   *telu   tou*
           telu   tow
           three   person

           'Three people.'   KINIAR02_14052013_00:10:18

(4)   a.   *ntimpa*   *ye'i*
           N=timpa'   ye'i
           INAN=palm.sugar.sap   DEM1

           'This palm sugar sap.'   KINIAR02_14052013_00:01:01

      b.   *pupuk   waya ye'i*
           pupuk   waya ye'i
           fertilizer   all   DEM1

           'All this fertilizer.'   TDN_10_JL_FM_24082011_00:24:58

      c.   *koko'   rua*
           ko'ko'   rua
           chicken   two

           'Two chickens.'   TDN_14_DK_NK_03092011_00:02:36

With regard to TNW, the current analysis of NP structure shows a pattern which is identical to that of TDN, with one exception; demonstratives precede the head rather than follow it, that is:

<div align="center">(DEM) (modifier) (phrase marker) HEAD (modifier)</div>

Thus, numerals and quantifiers are able to either precede or follow the head, while demonstratives and phrase markers always precede it, as observed in (5a)-(5c) and (6a)-(6c).

(5) a. *wa'aho manguni*
    wa'aho manguni
    DEM2 owl

    'That owl.'   TNW_02_JP_26062016_00:08:31

    b. *telu sèng*
       telu sèng
       three roofing.iron

       'Three (sheets of) roofing iron.'   TNW_16_JP_MT_05102016_00:05:58

    c. *pahasa walè*
       pahasa walè
       all house

       'All the houses.'   TNW_37_YK_07072016_00:02:56

(6) a. *nia'i ndaran*
    N=ia'i   N=dalan
    INAN=DEM1 INAN=road

    'This road.'   TNW_10_RP_28092016_00:07:51

    b. *latang esa*
       latang esa
       clothes one

       'One (pair of) clothes.'   TNW_08_JP_07072016_00:05:42

    c. *kita pahasa*
       kita pahasa
       1PL.IN all

       'All of us.'   TNW_42_BP_25022017_00:01:34

Examples (1)-(6) demonstrate that the three languages exhibit similar restrictions and flexibility in the distribution of quantifiers, numerals, and articles/phrase markers. Additionally, demonstratives show similar distribution in two of the languages, MM and TDN. This structural similarity hints at a potential for contact-induced change. In order to test this, the patterns attested in larger-scale studies of contact situations offer an effective diagnostic. Specifically, in a wide-ranging examination of structural borrowing, Matras & Sakel (2007) find that the modification of elements within possessive constructions is the most common contact-induced change in constituent order[18] (Matras 2007:43–44, 60).

To potentially confirm structural borrowing at the NP level, we should be able to observe

---

[18] Because any change in the distribution of possessor and possessed does not affect the verb and therefore leaves predicate structure unaffected (Matras 2007:60).

similar distribution patterns for the possessor and possessed elements. In MM, the order of elements is for the possessor to be followed by a linker and then the possessed – a structure frequently observed in PDMs (Adelaar 2005:213). This structure (Table 8) has the possessor, whether an independent or bound pronominal or an NP, followed by *pè*, a shortened form of *punya* 'have, own' which functions as a linker between the two elements.

Table 8. Possession in MM

	Possessor (pronoun) + *pè*	+ possessed:	English gloss:
1SG	*kita ~ ta= pè*	*ruma*	'my house'
2SG	*ngana pè*	*ruma*	'your house'
3SG	*dia ~ dè= pè*	*ruma*	'his/her house'
1PL	*torang ~ tong= pè*	*ruma*	'our house'
2PL	*ngoni pè*	*ruma*	'your house'
3PL	*dorang ~ dong= pè*	*ruma*	'their house'

	Possessor (NP) + *pè*	+ possessed:	English gloss:
	Bapak Jan pè	*ruma*	'Mr Jan's house'

The distribution of elements observed in TDN and TNW is outlined in Tables 9 and 10. In contrast to MM, in both Minahasan languages the element expressing the possessed entity always precedes the possessor pronoun or NP. In TDN, all possessive pronouns are realised as enclitics, while possessors expressed lexically must host one of the two proclitic phrase markers *ni=* (SG) or *nè=* (PL).

Table 9. Possession in TDN

	Possessed:	+ possessor (pronoun):	English gloss:
1SG	*walè*	*=ku*	'my house'
2SG	*walè*	*=mu*	'your house'
3SG	*walè*	*=na*	'his/her house'
1PL.IN	*walè*	*=ta*	'our house'
1PL.EX	*walè*	*=mèy*	'our house'
2PL	*walè*	*=miu*	'your house'
3PL	*walè*	*=nèa*	'their house'

	Possessed:	*ni=/nè=* + possessor (NP):	English gloss:
	walè	*ni=Bapak Jan*	'Mr Jan's house'

As shown in Table 10, the TNW structure is almost identical. The sole difference is that only singular possessors are realised with an enclitic. Plural possessors are encoded with an independent pronoun preceded by the phrase marker *i=*, as are possessors expressed with an NP.

The order of possessor and possessed clearly diverges in the two categories of languages. The MM structure matches that of TM (Litamahuputty 2012:59) and of the non-AN Ternate

Table 10. Possession in TNW

	Possessed:	+ possessor (pronoun):	English gloss:
1SG	*walè*	=*ku*	'my house'
2SG	*walè*	=*mu*	'your house'
3SG	*walè*	=*na*	'his/her house'
1PL.IN	*walè*	*i=kita*	'our house'
1PL.EX	*walè*	*i=kamitahula*	'our house'
2PL	*walè*	*i=kamotahula*	'your house'
3PL	*walè*	*i=latahula*	'their house'

	Possessed entity:	*i=* + possessor (NP):	English gloss:
	walè	*i=Bapak Jan*	'Mr Jan's house'

language (Watuseke 1991:233), with the latter the likely source of this structure in the two creoles. The order of elements in the two Minahasan languages matches that frequently attested to AN languages of the same genetic and geographic provenance (Himmelmann 2005:142).

When considered together with the absence of high levels of phonological and lexical borrowing, this divergent possessive structure negates any analysis of shared NP constituent order resulting from contact. Any corresponding flexibility of NP constituents is in all probability coincidental. It is accounted for in the Minahasan languages as a feature of many languages classified as the 'Philippine-type' (Himmelmann 2005:142). With respect to MM, the flexibility of numerals and quantifiers is best explained as an inheritance from TM.

5.2 Morphology and morphological type

In spite of occasional arguments to the contrary (e.g. Gardani 2012; Seifart 2017), the transfer of bound morphology, especially inflectional morphemes, is often considered as one of the least probable outcomes of language contact (Sankoff 2004:648, Myers-Scotton 2006:231). When taking into account the historically dissimilar morphosyntax of MM and the indigenous Minahasan languages, there are minimal expectations of any transfer at this level.

Bound morphology in MM is limited, a not unexpected typological feature in pidgin and creole languages which are often considered to lack inflectional morphology (Meakins 2019:69). The inventory of MM bound morphemes comprises the clitic pronouns in Table 8, the definite articles *tu=* and *ni=*, and the prefixes *ba-*, *ta-*, *baku-*, *paN-*, and *ka(h)-* (Prentice 1994:431–34; Stoel 2005:18; Paauw 2008:238–241).[19] Of these affixes, arguably only *ba-* and *ta-* are productive. What is more, all prefixes are optional (Stoel 2005:18) and MM words are frequently monomorphemic.

The morphological typology of MM is exemplified by (7a)-(7c) in which clausal con-

---

[19] This prefix inventory is a much reduced version of the original Malayic morphology. Cognates of these forms are observed in many PDMs (Adelaar 2005:216)

stituents are overwhelmingly monomorphemic and require no bound morphology. Note also the periphrastic causative construction in (7c) as expressed by the auxiliary verb *bèk-ing* adjacent to the main verb *bodo*.

(7) a. *dia da   saki puru   tadi malang.*
         dia da   saki puru   tadi malang
         3SG ASP sick stomach last night

         'He had a stomach ache last night'.   ([Stoel 2005](#):139)

     b. *kita dengar kwa' dorang somo   kaweng.*
         kita dengar kwa' dorang so=mo   kaweng
         1SG hear   PART 3PL   ASP=ASP marry

         'But I heard they were about to get married '.   ([Stoel 2005](#):38)

     c. *dia so   bèking bodo pa torang.*
         dia so   bèking bodo pa torang
         3SG ASP AUX   stupid at 1PL

         'He fooled us (i.e. made us stupid)'.   ([Stoel 2005](#):35)

In contrast to this paucity of bound morphology, both TDN and TNW contain large inventories of affixes and clitics, many of which are highly productive or obligatory.[20] These include prefixes, infixes, suffixes, circumfixes, proclitics, and enclitics. They encode a variety of functions at both phrasal and clausal level, ranging from phrase marking and nominalisation, to voice and TAM marking, in addition to pragmatic-related functions.

Examples (8a)-(8c) (TDN) and (9a)-(9c) (TNW) demonstrate this morphological complexity. Voice-marking affixes which occur on verbs, and which encode either ACTOR voice (AV) or one of three UNDERGOER voices (PV, LV, or CV), are obligatory. Additional verbal morphology is observed in the use of the affix *pa-* in TDN and *pah-* in TNW to encode multiple constructions, including causatives as shown by (8c) and (9c).

(8) a. *siso'o   rèèn mesesèrola   sèwalina.*
         si=so'o   rè'èn ma-Ce-sèro=la   sè=walina
         3SG.NOM=not.want   PART AV.DYN-IRR-search=DIR.PROX AN.PL=other

         'He does not want to look for others (women) then'.
                                                    TDN_28_OL_KK_25092012_00:03:53

     b. *pesesèroana   ensekolana.*
         pa-Ce-sèro-an=na   N=sekola=na
         DYN-IRR-search-LV=3SG.GEN INAN=school=3SG.GEN

         'He will seek his education'.   TDN_31_KK_17102012_00:05:37

---

[20] For further information on TDN and TNW morphology see [Sneddon](#) (1975, 1978) and [Brickell](#) (2015, 2016a, 2016b).

c. *sia   papaloongkula*   *wia   siwewènè.*
   nisia pa-pa-loo'-en=ku=la   wia   si=wewènè
   3SG   CAUS-DYN-see-PV=1SG.GEN=DIR.PROX PREP   AN.SG=woman

   'I showed (i.e introduced) him to the woman '.

                                                            TDN_31_KK_17102012_00:06:40

(9) a. *mahinjomohosia   latang.*
       mah-injo=moho=sia   latang
       AV.DYN-take=DIR=3SG   clothes

       'He takes the clothes'.   TNW_08_JP_07072016_00:05:44

   b. *bati' pinahinjoaiku.*
      bati' p<in>ah-injo-an=ku
      only DYN.PST-take-LV=1SG.GEN

      'I only took (that story) '.   TNW_22_JP_16102016_00:08:21

   c. *iwawatianamoho   aho   sopoina.*
      i-pah-pah-tia'=na=moho   wa'aho   sopoy=na
      CV-CAUS-DYN-fall=3SG.GEN=DIR   DEM2   cloth.container=3SG.GEN

      'He pushes down his cloth (betel nut) container '.

                                                            TNW_02_JP_26062016_00:14:59

Despite this strong divergence in morphological typology, the three languages do in fact
share two bound morphemes: *ta=* and *ka(h)-*. The latter is found in all three languages,
while *ta=* is attested in both MM and TDN, as demonstrated by (10a)-(10b).

(10) a. *tada   manyanyi.*
        ta=da   manyanyi
        1SG=ASP   sing

        'I was singing'.   (Stoel 2005:51)

    b. *tasumèrèt   pesawat.*
       ta=s<um>erèt   pesawat
       1PL.IN.NOM=<AV>ride   machine

       'We board an airplane'.   TDN_14_DK_NK_03092011_00:05:35

The fact that *ta=* is currently attested in MM and TDN, but not in TM (Paauw 2008:169;
Litamahuputty 2012:142), raises the possibility of morphological transfer from TDN to
MM. When contemplating this prospect, examining the corresponding independent forms,
kita in MM and *nikita ~ kita* in TDN, provides some clarification. *Kita* is ubiquitous
in AN languages and is reconstructed for both PAN and PMP (Ross 2006:36, 51) with
identical person and number values to those in TDN, i.e. 1st person plural inclusive. This
contrasts with *kita* in MM which encodes 1st person singular referents. These latter person
and number values are pertinent, as they match those for *kita* in the colloquial variety
of Indonesian spoken in Jakarta (Adelaar 2005:213; Ewing 2005:246). With Jakartan

Indonesian so prevalent in popular culture, this could be viewed as another instance of borrowing from this source. Nonetheless, this possibility is disregarded due to the fact that many PDMs, including TM (Litamahuputty 2012:142), exhibit *kita* for 1st person singular (Adelaar 2005:213; Paauw 2008:169). This fact has been taken as an indication that a singular **kita* pronoun was the original form in the ancestor creole (Adelaar 2005:213). Rather than an analysis of borrowing, a more suitable theory is therefore that *kita* was present in MM and other PDMs prior to any ongoing contact with Indonesian.

The point here is that the shared *kita* pronoun can be explained as a homophonous form representing two lexical items with separate origins, as reflected in the differing person and number values. The same is true of the bound counterparts. Explaining shared *ta=* via a borrowing scenario would involve morphological transfer (from a less dominant language) of core vocabulary in the absence of substantial phonological or lexical transfer. A more reasonable hypothesis is that this similarity represents the frequent cross-linguistic phenomenon of independent person markers developing cliticised forms (Siewierska 2004:262). The presence of *ta=* in both MM and TDN therefore reflects two independent but identical diachronic developments - both independent pronouns developed shortened forms which maintained their specific person and number values.

The second shared morpheme, the prefix *ka(h)-*, derives ordinal numerals in all three languages, as demonstrated by (11a)-(11c).

(11) a. *katelu,   ni'tu   lelaya'an.*
          ka-telu   ni'tu   Ce-laya'-an
          ORD-three DEM2 NMLZ-dance.happily-LV

          'The third (phase), is that joyful dancing '.

                                                    TDN_31_OL_KK_17102012_00:03:29

      b. *kasiow   ngando   mapera.*
          kah-siow   nga-ando mah-pela'
          ORD-nine   LNK-day   STAT-dry

          'On the ninth day (the land) dried out'.   TNW_02_JP_26062016_00:16:10

      c. *so   katiga   kali tubèl   grèja   babunyi.*
          so   ka-tiga   kali tu=bèl   gerèja ba-bunyi
          ASP ORD-three kali ART=bell church INTRANS-sound

          'The church bell has already rung for the third time '.

                                                    (Dotulong 2010:47)

Ordinal numerals formed with *ka(h)-* are present in all Minahasan languages (Sneddon 1978:104). In the event of a borrowing scenario, which would again need to overcome the issue of minimal borrowing at other structural levels, the most plausible direction would be from MM into TDN and TNW. Evidence against this is found in the **ka-* prefix which is reconstructed with the same function for Proto-Minahasan (Sneddon 1978:104). What is more, *ka(h)-* in TDN and TNW in all probability has its origin in the PAN prefix **sika-* 'marker of ordinal numerals' (Blust 2013:372), reflexes of which are widespread among contemporary AN languages (Blust 2013:381).

The alternative borrowing scenario, in which MM speakers have adopted *ka(h)-* from indigenous languages, is equally unlikely as it requires that a dominant language is the recipient of morphological transfer. This then leaves two alternatives: either *ka(h)-* was already present in TM when it was transported to North Sulawesi or it has been borrowed from another source. With regard to the first option, while *ka-* (Paauw 2008:372) or *ke-* (Litamahuputty 2012:67) is also attested in contemporary TM with an identical function, a lack of documentary and historical record makes it impossible to prove how long this morpheme has been present. An alternative hypothesis provided by Paauw (2008:193–194) posits the existence of this ordinal marker in multiple PDMs as a recent innovation due to contact with Indonesian. Considering the likely phonological and lexical transfers from Indonesian, an analysis of MM having adopted *ka(h)-* in this way is highly credible.

6. North Sulawesi contact outcomes explained

Despite hundreds of years of contact between TDN and TNW and two closely related varieties of a Trade Malay, this preliminary investigation concludes that with the exception of lexical transfer, none of the small measure of shared features are the result of language contact. Instead, they are accounted for as pre-existing features, independent innovations, or the result of contact with the omnipresent varieties of Indonesian.

The lack of linguistic transfer, except at the cross-linguistically frequent level of the lexicon, indicates a contact scenario in which the prime requirement for borrowing, stable and ongoing bilingualism (Thomason & Kaufman 1988; Thomason 2001:1640; Bybee 2015:248), appears never to have occurred, neither pre- nor post-colonisation. Quantifying the exact level and duration of any bilingualism is difficult, primarily because the stage at which MM first gained native speakers is uncertain. Based on the available information, it appears that there have been at most three generations of bilinguals, with the first generation appearing sometime in the late 19th or early 20th Century. Of more certainty is that the current generation will be the last.

A long-running and fundamental barrier to bilingualism has been the historical asymmetry in status, institutional support, and agency between the two sets of languages and their speakers. From the first period of sustained contact, the indigenous population were almost certainly forced to gain some degree of fluency in a PDM variety in order to communicate with the *borgo*, who had higher-status positions within the colonial administration and who were already speakers of this language (Schouten 1998:49, 55, 101). For the native Minahasans, increasing their fluency was a necessity for any economic, social, or political advancement. In contrast, none of these requirements or pressures were experienced by the *borgo*, most of whom presumably learned little of the indigenous languages beyond common vocabulary describing basic day-to-day concepts.

This preference for MM rather than indigenous languages continued and became entrenched during more than 200 years of concerted efforts by the VOC, the NZG, and the Dutch state to create a unified territorial, ethnic, and cultural Minahasan unit. Inherent to these policies was the use of the PDM in regional administration as well as in religious and educational domains by missionaries.[21] These policies were so successful that, in the

---

[21] It has been noted that the missionaries originally used a different variety of Malay, *Zendelingen-Maleisch* 'Missionary Malay' (de Clercq 1871:403 cited in Stoel 2005:8). However, this variety was soon supplanted

modern era, it is now MM which is used as a marker of Minahasan cultural and ethnic identity, rather than indigenous languages (Jacobson 2002:41; Brickell 2018a). MM now dominates the linguistic ecology of North Sulawesi at the expense of the linguistic vitality of the indigenous languages (Merrifield & Salea 1996; Mead 2013; Brickell 2015, 2018a; Hertz & Lee 2017). This situation has produced a break in inter-generational transmission which prevents the emergence of further generations of bilingual speakers.

In addition to these extra-linguistic factors, divergent typologies are an obstacle to structural borrowing at the morphosyntactic level. MM is morphologically isolating, a feature which presumably characterised TM from its inception. This contrasts greatly to the complex PAN-derived verbal morphology and morphosyntactic alignment systems found in TDN and TNW. On the assumption that linguistic change through contact is far more prevalent in typologically similar languages (Whitney 1881:15; Thomason 2006:345), this disparity further lowers the potential for structural transfer.

The language-contact outcomes identified in North Sulawesi match various established theories. Socio-historical and linguistic factors have reduced the potential for bilingualism and have decreased the capacity for transfer. When lexical borrowing does take place, the expected pattern occurs in which the dominant language is the source of most transferred material. Furthermore, in the absence of high levels of phonological or lexical borrowing the potential for transfer at the morphosyntactic level is remote.

---

by MM (Henley 1996:86).

Abbreviations

1	first person	IRR	irrealis
2	second person	LNK	linker
3	third person	LV	locative voice
AN	animate	NMLZ	nominaliser/nominalisation
ART	article	NOM	nominative
ASP	aspectual marker	NP	noun phrase
AUX	auxiliary	ORD	ordinal
AV	actor voice	PREP	preposition
CAUS	causative	PART	particle
CV	conveyance voice	PROX	proximate
DEM1	proximate demonstrative	PART	particle
DEM2	medial demonstrative	PL	plural
DIR	directional particle	PP	prepositional phrase
DYN	DYNAMIC verbal affix	PST	past
EX	exclusive	PV	patient voice
GEN	genitive	REL	relativizer
INAN	inanimate	SG	singular
IN	inclusive	STAT	STATIVE verbal affix
INTRANS	intransitive		

References

Adelaar, Alexander. 2005. The Austronesian languages of Asia and Madagascar: A historical perspective. In Alexander Adelaar & Nikolaus P. Himmelmann (eds.), *The Austronesian languages of Asia and Madagascar.* London: Routledge. 1–42.

Adelaar, Alexander & David Prentice. 1996. Malay: its history, role and spread. In Stephen Wurm, Peter Mühlhäusler, & Darrel Tryon (eds.), *Atlas of languages of intercultural communication in the Pacific, Asia, and Americas.* Berlin: Mouton de Gruyter. 673–693.

Adriani, Nikolaus. 1925. De Minahasische talen [The Minahasan languages]. *Bijdragen tot de Taal-, Land- en Volkenkunde* 81. 134–164.

Bakker, Laurens. 2016. Organised violence and the state: Evolving vigilantism in Indonesia. *Bijdragen tot de Taal-, Land- en Volkenkunde* 172(2/3). 249–277

Blust, Robert. 1991. The Greater Central Philippines hypothesis. *Oceanic Linguistics* 30(2). 73–129.

Blust, Robert. 2013. *The Austronesian languages*. Ebook. Revised edition. Canberra: Asia Pacific Linguistics. https://tinyurl.com/y6pe9v39 (accessed 2019-05-17).

Brickell, Timothy C. 2015. *A grammatical description of the Tondano (Toundano) language*. Ph.D thesis, La Trobe University.

Brickell, Timothy C. 2016a. Tondano. In Geoffrey Haig & Stefan Schnell (eds.), *MultiCAST (Multilingual Corpus of Annotated Spoken Texts).* Bamberg: University of Bamberg archive. https://multicast.aspra.uni-bamberg.de (accessed 2019-10-09).

Brickell, Timothy C. 2016b. *Tonsawang: a collaborative multimedia project documenting an endangered language of North Sulawesi.* London: SOAS, Endangered Languages Archive. https://elar.soas.ac.uk/Collection/ MPI1035088 (accessed 2019-10-08).

Brickell, Timothy C. 2018a. Tonsawang (Toundanow), North Sulawesi, Indonesia – Language contexts. *Language Documentation and Description* 16. 55–85.

Brickell, Timothy C. 2018b. Reduplication in Tondano and Tonsawang. *NUSA: Linguistic studies of languages in and around Indonesia* 65. 81–107.

Bybee, Joan. 2015. *Language change.* Cambridge: Cambridge University Press.

de Clercq, F S A. 1871. Eene bijdrage tot de kennis van het Maleisch der Minahasa [A contribution to the knowledge of the Malay of Minahasa]. *Tijdschrift voor Nederlandsch-Indie* 2. 403–425.

Dotulong, Boeng. 2010. *Kamus Malayu Manado-Indonesia-Toundano.* Jakarta: Percetakan Gramedia.

Ewing, Michael C. 2005. Colloquial Indonesian. In Alexander Adelaar & Nikolaus P. Himmelmann (eds.), *The Austronesian languages of Asia and Madagascar.* London: Routledge. 227–258.

Gardani, Francesco. 2012. Plural across inflection and derivation, fusion and agglutination. In Lars Johanson & Martine Robbeets (eds.), *Copies versus cognates in bound morphology.* Leiden: Brill. 71–97.

Glottolog. 2019. Spoken L1 language: Manado Malay. In Hammarström, Harald, Forkel, Robert & Martin Haspelmath (eds.), *Glottolog 4.1.* Jena: Max Planck Institute for the Science of Human History. https://glottolog.org/resource/languoid/id/mala1481 (accessed 2019-24-03).

Google Maps. 2018. North Sulawesi, scale undetermined. Google Maps 2018. TinyURL: https://tinyurl.com/y54nrgn7 (accessed 2019-18-03).

Haugen, Einar. 1950. The analysis of linguistic borrowing. *Language* 26 (2). 210–231.

Haspelmath, Martin. 2009. Lexical borrowing: Concepts and issues. In Martin Haspelmath & Uri Tadmor (eds.), *Loanwords in the world's languages: A comparative handbook.* Berlin: Mouton de Gruyter. 35–54.

Hayes, Rebekah. 2019. *A morphosyntactic description of verbal constructions in Tonsawang, an endangered language of northern Sulawesi.* MA thesis, SOAS, University of London.

Henley, David. 1993. A superabundance of centers; Ternate and the conquest for North Sulawesi. *Cakalele; Maluku Resarch Journal* 4. 39–60.

Henley, David. 1996. *Nationalism and regionalism in colonial Indonesia: The case of Minahasa.* Leiden: KITLV Press.

Hertz, Regina & Sandra Lee. 2017. *Assessment of the vitality of the Tondano Language (Sulawesi, Indonesia).* SIL electronic survey report. https://tinyurl.com/ycxhlrgw (accessed 2019-16-11)

Himmelmann, Nikolaus. 2005. The Austronesian languages of Asia and Madagascar: Typological characteristics. In Alexander Adelaar & Nikolaus P. Himmelmann (eds.), *The Austronesian languages of Asia and Madagascar.* London: Routledge. 110–181.

Jacobson, Michael. 2002. On the question of contemporary identity in Minahasa, North Sulawesi province, Indonesia. *Asian Anthropology* 1. 31–58.

Kalangi, Patrice. 2012. *Kamus Tonsawang/Toundanow Indonesia.* Tompaso: Yayasan Institut Seni Budaya Sulawesi Utara.

King, Ruth. 2000. *The lexical basis of grammatical borrowing: A Prince Edward Island French case study.* Amsterdam: John Benjamins.

Litamahuputty, Betty. 2012. *Ternate Malay: Grammar and texts.* Ph.D thesis, Leiden University.

Matras, Yaron & Jeanette Sakel. 2007. Introduction. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective.* Berlin: Mouton de Gruyter. 1–13.

Matras, Yaron. 2007. The borrowability of structural categories. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross linguistic perspective.* Berlin: Mouton de Gruyter. 31–73.

Mead, David. 2013. *An initial assessment of the vitality of the indigenous languages in North Sulawesi.* Unpublished manuscript.

Meakins, Felicity. 2019. Morphology. In Jereon Darquennes, J. Salmons, & W. Vandenbusshe (eds.), *Language contact: An international handbook* Berlin: De Gruyter Mouton. 63–76.

Merrifield, Scott & Martinus Salea. 1996. *North Sulawesi language survey.* Summer Institute of Linguistics Publications in Sociolinguistics. Dallas: SIL International.

Moravcsik, Edith. 1978. Universals of language contact. In Joseph Greenberg (ed.), *Universals of human language: Vol. 1 Method and theory* Stanford: Stanford University Press. 95–122.

Myers-Scotton, Carol. 1993. *Duelling languages: Grammatical structure in code switching.* Oxford: Clarendon.

Myers-Scotton, Carol. 2002. *Contact linguistics: Bilingual encounters and grammatical outcomes.* Oxford: Oxford University Press.

Myers-Scotton, Carol. 2006. *Multiple voices: An introduction to bilingualism.* Malden: Blackwell.

Ogloblin, Alexander K. 2005. Javanese. In Alexander Adelaar & Nikolaus P. Himmelmann (eds.), *The Austronesian languages of Asia and Madagascar.* London: Routledge. 590–624.

Paauw, Scott. 2008. *The Malay contact varieties of Eastern Indonesia: A typological comparison.* Ph.D thesis, State University of New York.

Poplack, Shana & Marjory Meechan. 1998. How languages fit together in code-mixing. *International Journal of Bilingualism* 2(2). 127–138.

Poplack, Shana & Sankoff, David & Christopher Miller. 1988. The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics* 26. 47–104.

Prentice, David. 1994. Manado Malay: Product and agent of language change. In Tom Dutton & Darrel T. Tryon (eds.), *Language contact and change in the Austronesian world*. Berlin: Mouton de Gruyter. 411–441.

Prince, Ellen. 1988. On pragmatic change: the borrowing of discourse functions. *Journal of pragmatics* 12(5/6). 505–518.

Renwarin, Paul. 2006. *Matuari and Tona'as: The cultural dynamics of the Tombulu in Minahasa*. Ph.D thesis, Leiden University.

Ross, Malcolm. 2006. Reconstructing the case-marking and personal pronoun systems of Proto Austronesian. In Henry Y. Chang, Lillian M. Huang, & Dah-an Ho (eds.), *Streams converging into an ocean: festschrift in honor of Professor Paul Jen-kuei Li on his 70th birthday*. Taipei: Academia Sinica. 521–563.

Sakel, Jeanette. 2007. Types of loan: Matter and pattern. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*. Berlin: Mouton de Gruyter. 15–30.

Sankoff, Gillian. 2004. Linguistic outcomes of language contact. In Peter Trudgill, J. Chambers, & N. Schilling-Estes (eds.), *The handbook of language variation and change.* Oxford: Blackwell. 638–668.

Schouten, Mieke. 1998. *Leadership and social mobility in a Southeast Asian society: Minahasa 1677–1983*. Leiden: KITLV Press.

Seifart, Frank. 2017. Patterns of affix borrowing in a sample of 100 languages. *Journal of historical linguistics*. 7(3). 389–431.

Shiohara, Asako & Anthony Jukes. 2018. Two definite markers in Manado Malay. In Sonja Riesberg, Asako Shiohara & Atsuko Utsumi (eds.), *Perspectives on information structure in Austronesian languages* (Studies in Diversity Linguistics 21). Berlin: Language Science Press. 115–135.

Siewierska, Anna. 2004. *Person*. Cambridge: Cambridge University Press.

Simons, Gary F. & Charles D. Fennig (eds.) 2019. *Ethnologue: Languages of the world, twenty-first edition*. Online version: https://ethnologue.com (accessed 2019-01-08). Dallas: SIL International.

Sneddon, James. 1975. *Tondano phonology and grammar*. Canberra: Pacific Linguistics.

Sneddon, James. 1978. *Proto-Minahasan: phonology, morphology, and wordlist*. Canberra: Pacific Linguistics.

Steinhauer, Hein. 2005. Colonial history and language policy in Insular Southeast Asia and Madagascar. In Alexander Adelaar & Nikolaus P. Himmelmann (eds.), *The Austronesian languages of Asia and Madagascar.* London: Routledge. 65–86.

Stoel, Ruben. 2005. *Focus in Manado Malay: Grammar, particles, and intonation*. Leiden: CNWS Publications.

Tadmor, Uri. 2009. Loanwords in the world's languages: Findings and results. In Martin Haspelmath & Uri Tadmor (eds.), *Loanwords in the world's languages: A comparative handbook*. Berlin: Mouton de Gruyter. 55–75.

Thomason, Sarah & Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.

Thomason, Sarah. 2001. Contact-induced typological change. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher, & Wulfgang Raible (eds.), *Language typology and language universals: An international handbook*. Berlin: Walter de Gruyter. 1640–1647.

Thomason, Sarah. 2006. Language contact and language change. In Keith Brown (ed.), *Encyclopedia of languages & linguistics*. Boston: Elsevier. 338–368.

Utsumi, Atsuko. 2014. *Recordings of Tonsawang (*AU1*)*. Digital collection managed by PARADESIC. https://catalog.paradisec.org.au/collections/AU1 (accessed 2019-08-03)

Utsumi, Atsuko. 2018. The Tonsawang language's basic morphology and syntactic features. *Paper presented at The Fourteenth International Conference of Austronesian Linguistics (14-ICAL). July 17-20.* Antananarivo: Universitè di Antananarivo.

Van Coetsem, Frans. 1988. *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.

Voorhoeve, C. L. 1983. Some observations on North Moluccan Malay. In James T. Collins (ed.), *Studies in Malay dialects Part II*, NUSA 17. Jakarta: Universitas Atma Jaya, 1–13.

Watupongoh, G. Y. J., Najoan, J. A. K., Kodong, F., Lotulong, D. & S. K. Kanuang. 1992. *Struktur bahasa Tondano*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa, Departemen Pendidikan dan Kebudayaan

Watuseke, F. S. 1991. The Ternate language. In Tom Dutton (ed.), *Papers in Papuan Linguistics* 1. Canberra: Pacific Linguistics, 223–244.

Watuseke, F. S. & W. B. Watuseke-Politton. 1981. Het Minahasa- of Manado-Maleis [It's Minahasa or Manado Malay]. *Bijdragen tot de Taal-, Land- en Volkenkunde* 137(2). 324–346.

Whisler, Ronald. 2006. *The historical development of Manado Malay*. Universitas Kristen Indonesia Tomohon (UKIT) journal. Tomohon, North Sulawesi: Universitas Kristen Indonesia Tomohon.

Whitney, William. 1881. On mixture in language. *Transactions of the American Philological Society* 12. 5–26.

World Bank.1988. *Indonesia–The transmigration program in perspective (English)*. A World Bank Country study. https://tinyurl.com/uzyd376 (accessed 2019-11-23). Washington DC: The World Bank.

Wurm, Stephen & Shiro Hattori (eds.) 1981. *Language Atlas of the Pacific*. Australian Academy of the Humanities and The Japan Academy. Canberra: Pacific Linguistics.

Zorc, David. 1986. The genetic relationships of Philippine languages. In Paul Gerahty, Lois Carrington, & Stephen A. Wurm (eds.), *Tracking the travelers* (FOCAL II: Papers from The fourth International Conference of Austronesian Linguistics). Canberra: Pacific Linguistics, 147–173.