

# Word-initial [h]-drop variation in Nmbo

## Change-in-progress in an egalitarian multilingual speech community of Papua New Guinea

Eri Kashima

Australian National University | University of Helsinki

This paper presents a natural speech corpus-based study of word-initial [h]-drop from the Nmbo speech community of southern Papua New Guinea. It is a speech community within a traditional egalitarian multilingual language ecology sustained by a practice of virilocal exogamy, and there is strong intergenerational transmission of local vernacular languages. This study investigates the propensity of word-initial [h]-drop in nouns, based on Nmbo speech data of Kerake tribe people. The results from the Nmbo Sociolinguistic Corpus shows clear age-conditioned variation, with younger speakers showing a higher propensity for [h]-drop. Nmbo speakers residing both within and outside their Nmbo villages of origin appear to be partaking in the innovative [h]-drop. The origin of the [h]-drop appears to be from the village with a more multilingual profile, as would be predicted by the notion of a multilingual *feature pool* (Cheshire, Kerswill, Fox, & Torgersen, 2011, Mufwene 2001).

**Keywords:** Papua New Guinea, Pacific, exogamy, deletion, lenition, phonological variation, network analysis, egalitarian multilingualism, apparent time construct

### 1. Introduction

This study investigates the linguistic and social forces at play in a highly multilingual community in Papua New Guinea. The primary goal of this study is to contribute to our expanding understanding of socio-cultural pressures that contribute to the transmission and diffusion of linguistic innovations. The results of this paper provide empirical evidence from a part of the world which retains a traditional language ecology of non-hierarchical multilingualism between small population vernaculars. There is some evidence that the village with the more

multilingual profile may be the origin of the linguistic innovation under investigation.

The innovation in question is word-initial [h]-drop of nouns occurring in the Nmbo (pronounced /nɛm.bo/) speech community of Southern New Guinea. Nmbo /h/ is a phoneme found in lexical words such as nouns and verbs, and can occur word-initially, intervocally, and word finally. Nmbo /h/ should be an emblematic phoneme in some senses as it is semi-unique within the Nambu branch family to which it belongs. Nmbo /h/ has clear cognate phonemes across its sister languages, having likely derived from Proto-Nambu \*s (Evans, Carroll, & Döhler, 2017); for example /s/ in Nen, and /ɣ/ in Nama (Nmbo /hɛkɛβ/ 'eye', /səkop/ in Nen, /ɣəkəp/ in Nama). The [h] phone occurs in limited contexts in other Nambu branch languages, such as in affirmations and loan words (e.g., in Nen, Evans, 2015a).

There is no social evaluation of [h]-drop by Nmbo speakers, which suggests that the phenomenon has proceeded through the speech community; that is, it is a change in progress. A classic observation made by Labov (1965) is that a "change from below" (i.e., below the level of social awareness) shows no patterns of stylistic variation (p. 110), and that such changes proceed across time by the "incrementation by children, who reproduce and advance their parents' system" (Labov, 2007, p. 379). Change from below can thus be characterised as variation conditioned by age where there is no social meaning associated with variable realisations. In this instance we will see that Labov's predictions are borne out and that the Nmbo [h]-drop appears to have progressed through the Nmbo speech community, conditioned by age.

This study of Nmbo was done in conjunction with a language documentation project. Nmbo is an under-described language, and a sketch grammar has been produced as part of a doctoral dissertation by the author (Kashima, 2020). Two papers on the phonetics and phonology of Nmbo have also been published (Kashima, Williams, Ellison, Schokkin, Escudero, 2016; Kashima, in press), as well as some lexical similarity comparisons across the Nambu branch languages (Rueck, 2011; Tucker, Boevé, Fuller, Gustafsson, & Rueck, 2003). This study is a proof of a concept of doing language documentation concurrently to documenting sociolinguistics (cf. Childs, Good, & Mitchell, 2014), as well as studying variation (Hildebrandt, Jany, & Silva 2017); what Meyerhoff (2017) poetically describes as "writing a linguistic symphony" as opposed to a "sonata of individual variables".

This study also sits at the intersection of small-scale multilingualism research and language variation and change. There is a growing understanding that bi- and multilingual speakers play major roles in the long term trajectory of languages over time (c.f. Evans 2017a, 2018), yet our scientific understanding of bi- and mul-

bilingualism tend to overwhelmingly come from one of two approaches. On the one hand, ethnographic and qualitative studies of multilingual language ecologies shed insight to the variety of such communities across the globe (c.f. Vaughn & Singer, 2018; Lüpke, 2016, 2018; Epps, 2018), but there are few studies that systematically and quantitatively investigate naturalistic speech of speakers from such communities. On the other hand, while variationist work has long investigated the speech of bilingual speakers (c.f. Sankoff, 2001; Torres Cacoullos & Travis, 2018), these studies have tended to be from urbanised societies. This study aims to occupy these gaps.

## 2. The Nmbo speech community

Nmbo (also known as Nambo, glottocode NAMB1293, ISO-639-3 code ncm) is one of the six identified languages of the Nambu branch of the Yam family (also known as the Morehead-Upper Maro languages in Ross, 2005, with Wurm & Hattori, 1981 calling it the highest level of classification the Morehead-Wasur family). It is a Papuan language (i.e., a non-Austronesian language), and is spoken in an area that is geographically central within the Nambu branch dialect chain. Nmbo is a vital vernacular language, though due to the small speaker population (some 700~1000 speakers) it is technically perpetually endangered. Nmbo has two dialects, with the numerically slightly larger Kerake tribe variety being the target of this study. The other variety Namna (also known as Nambo, pronounced /nam.bo/, or Namna) is spoken by the Yarne tribe located immediately to the west of the Kerake.<sup>1</sup> Nmbo is an under-described language without much prior linguistic description. A sketch grammar has been produced by the author for a dissertation (Kashima, 2020), as well as a preliminary acoustic phonetic description of the vowel space (Kashima et. al., 2016). Work is increasingly being conducted in the area of Southern New Guinea (e.g., Evans et al., 2018), with descriptive analyses being published on the sister languages of Nen (e.g., Evans, 2012, 2015b, 2015c, 2017b) and Nama (e.g., Siegel, 2014, 2017).

The speech community of Nmbo is located in Southern New Guinea in an area known as the Morehead administrative district of the South-Fly district of Western Province. Southern New Guinea is a highly linguistically diverse part of

---

1. The Yarne tribe variety of Nambo/Namna is not included in this investigation since this variety has a velar approximant [uɟ] which corresponds with Kerake Nmbo [h]. Yarne Nambo/Namna is treated as another distinct language/variety for the purposes of studying this phonological variable, although in other parts of the linguistic system the two varieties appear to have a very high number of shared features (c.f. Kashima, 2020, pp.85–90)

New Guinea, second only to the Sepik and north coast of New Guinea (Evans et al., 2018, p. 641). In the Morehead area alone there are at least seven structurally distinct languages spoken across sixteen villages (Yam family, Nambu branch: Nmbo, Nen, Nama, Neme, Nā; Tonda branch: Komnzo; Pahoturi River Family: Idi). Languages in the Morehead area are emblems of tribal identity (Kashima, 2020, pp. 49–52), and people will identify themselves as *Kerake tribe* or *Nmbo speaker* relatively interchangeably.

The communities of the Morehead area have historically been small and dispersed. Hitchcock (2010) calculates the population density of the Morehead area to be approximately 0.5 persons/km<sup>2</sup> (p. 75). There is also no evidence of historical super-structures such as a chiefdom. The effects of colonialism are rather minimal, and to this day the presence of the nation state is faint with the absences of roads, hospitals, and commercial operations. The usage of Tok Pisin, the national lingua franca, is negligible, most likely due to this lack of development and infrastructure in the area. English has been taught at local community schools and the district capital Daru since the 1970s, with both men and women aged approximately 35 to 55 having the strongest command of English. People under the age of 30 are not particularly confident in their use of English, and the overwhelming majority of conversations in daily life occur in the local vernaculars. There are two community schools, one of which operates sporadically.

The relative absence of the nation state has meant that traditional *egalitarian multilingualism* (François, 2012, 2011) is still a part of the contemporary linguistic landscape. François used this term to describe the situation in Vanuatu, and the term is also suited for describing the non-hierarchical forms of small-scale vernacular multilingualism in the Morehead area. All languages are socially equal, with no language being viewed as more valuable, or more prestigious than the other. Nmbo speakers express little pressure to align their language usage to that of another language. Despite language being an emblem of tribe, Nmbo speakers tend not to express negative attitudes towards the acquisition of other tribal languages (see Döhler, 2018, pp. 34–36, however, for ambivalence expressed by the nearby Komnzo speakers). As has been noted for New Guinea and other parts of the world (Salisbury, 1962; Sankoff, 1980 for Highlands New Guinea, Lüpke, 2018 for North-western Cameroon), the ability to use multiple languages is often a display of one's social reach and power, and this is also the case in the Nmbo speech community (Kashima, 2020, p. 55). These language ideologies and behaviours are key characteristics of egalitarian multilingualism.

The term *egalitarian bilingualism* was first used by Haudricourt (1961) in relation to pre-colonial New Caledonia, but François's emphasis on multilingualism is more accurate in characterising the Morehead area. An individual may know anywhere between five and nine languages with varying skill levels of production and comprehension (Kashima, 2020, p.57). The differing levels of production ability result in both reciprocal and receptive forms of multilingualism; *reciprocal multilingualism* is characterised by individuals, or entire communities, speaking the languages of the neighbouring areas (Lüpke, 2016, pp.50–52). *Receptive multilingualism* refers to a context where speakers use their respective languages of choice while speaking to one another, and the interlocutors have receptive comprehension of that language (ten Thije & Zeevaert, 2007, p.1; Singer, 2018, p.102). Nmbo speakers report both types of multilingualism, using local expressions such as *näyāretan* “I hear it” and *nowavtan* “I speak it” as self-assessments.

The dynamics of the Nmbo language ecology are encapsulated in the local practice of virilocal exogamy, where women will marry out of their community of birth and upbringing, and relocate to their husband's community. More specifically, the idealised and often practiced form of marriage in the Morehead area is known in the anthropological literature as *direct sister exchange* (Williams, 1936; Ayres, 1983). In its idealised form, two men from opposing sections<sup>2</sup> will exchange their sisters for marriage. The preference appears to be for marriages to take place across different villages, although this is not explicitly stated unlike the rules of cross-section marriage. This results in a high likelihood of marrying a person who does not speak one's village language as a primary language. Women will, however, continue to speak their own tribal language to their children, and other women from their tribe. They will also speak their husband's village language; either acquired earlier in life from her own multilingual upbringing, or later in life as an adult learner. Children acquire their mother's language early in life and continue to speak it as part of their linguistic repertoire. This is in stark contrast to the exogamous Sui marriages in southern China (Stanford, 2009a, 2009b). In the Sui context, children who acquire their matrilect early in life abandon its use in favour of their patrilect/village-variety in accordance with local language ideologies. Nmbo speakers of both genders report speaking their mother's languages during childhood and into adulthood (Kashima, 2020, p.59). Due to the high lev-

---

2. *Section* refers to a system of lineage and descent tracing, somewhat like the term clan in other parts of the world (e.g., Stanford, 2009a; Mansfield, 2015). The term *section* is used by Williams (1936) in his description of the system in Morehead, and another major ethnography of the area (Ayres, 1983) follows the same terminological use. I follow the convention of these initial works and use the term *section*.

els of individual multilingualism, there is also a high chance that a woman's husband will also speak her village language.

The Nmbo speech community is primarily made up of three villages: from north to south, Govav, Bevdvn, and Arovwe (this study will focus on Govav and Bevdvn villages, with are located closer to one another than is Arovwe). The speech community also includes speakers in adjacent villages where out-marrying Nmbo women reside. Of relevance to this study is the Nen-speaking village of Bimadbn to the east (ISO NQN, glotto-code nenn1238), where Nmbo speaking Kerake tribe women were recorded. The total population of the three Nmbo villages in the years 2014–2017 was roughly 710. Govav and Arovwe are about the same size with approximately 250–300 residents, and Bevdvn is clearly the smallest, with 130 or so residents. If we include the number of Nmbo women who have married out of Nmbo villages into adjacent villages, plus their children and husbands who speak Nmbo regularly, we may have about 1000 speakers of Nmbo in the entire Morehead area.

Despite the low population density, the villages of the Morehead area are often geographically proximate to one another. This results in relatively frequent and prolonged contact between people of different villages. Christmas is often celebrated at a village level with families travelling to visit their in-laws. Families will help their in-laws with subsistence farming work, meaning that a woman who marries out of her Nmbo village maintains contact with Nmbo speakers, and has ample opportunity to physically return to her village of origin. The walking distance between Govav and Bevdvn is about one hour (5.5km), and Govav/Bevdvn to Nen-speaking Bimadbn is roughly a three-hour walk (approx. 14 km).

While the Nmbo villages comprise highly multilingual speakers, not all villages are equally multilingual. I collected data on household couples and their affiliate languages in order to get a rough measure of levels of multilingualism in the relevant village. For example if a household consisted of a Nmbo-speaking husband and a Nen-speaking wife, this was counted as a Nmbo-Nen bilingual household. Figure 1 shows the proportion of households in Bevdvn and Govav, with the latter showing greater multilingualism. Govav can thus be characterised as “more multilingual” than Bevdvn.<sup>3</sup>

---

3. This is not to say that Bevdvn villagers are not multilingual, since Bevdvn villagers have high degrees of multilingualism due to their unique biographies and extended kin. Bevdvn merely has a higher numbers of Nmbo-Nmbo households where the daily language is undoubtedly heavily skewed towards the usage of the village vernacular.

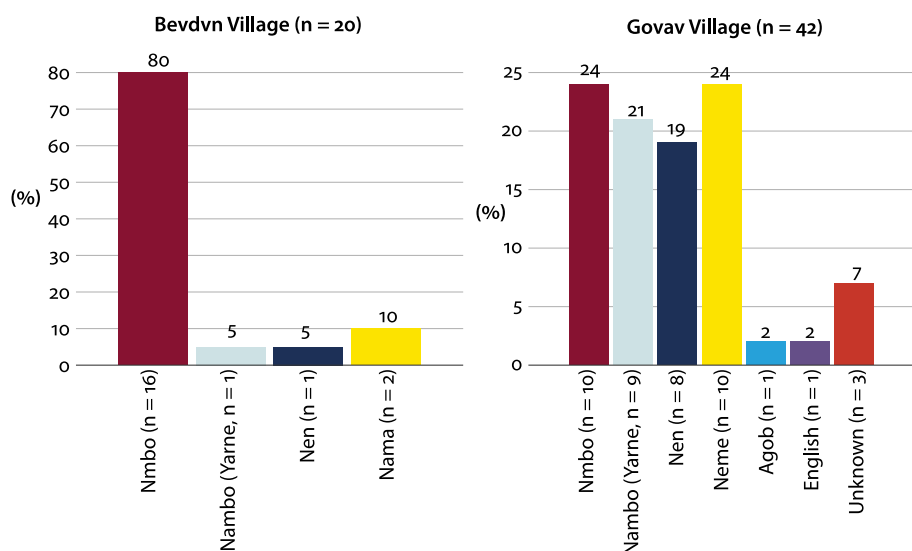


Figure 1. Proportions of household by primary language affiliations of husband and wife

### 3. Data and methodology

The linguistic variable under investigation is word-initial [h] presence or absence in nouns. The envelope of variation has been delimited narrowly for a few reasons. Firstly, nouns were chosen as the dependent because [h] occurs the most frequently in this class. Word-initial position was chosen since word final [h] is difficult to detect, while intervocalic [h] is infrequent in Nmbo speech. For example, intervocalic realisations of /h/ can occur in lexical nouns (e.g., /aha/, ‘dog’), and infinitive verb stems (e.g., on-h=et ‘to gather’ suffixed with the allative {=et} to produce ‘to go gathering’), but these constructions do not occur frequently enough in the corpus for a meaningful quantitative analysis. Given the scarcity of [h] tokens in other contexts, this study has limited the envelope of variation so that more can be said with certainty within its scope.

Nmbo speakers themselves do not remark upon the non-realisation of word-initial [h]. The earliest social evaluation of [h]-drop I heard came from a licensed Nen speaker, Jimmi Nébni, who is a highly capable speaker of Nmbo as is common for a person of the multilingual Morehead area. Jimmi had commented to me in the first year of my fieldwork, “Why are those Kerake dropping their h?”. This anecdotal episode suggests two things. One, that the direction of change is indeed of [h]-drop rather than [h] gain. Second, it shows how non-Nmbo speakers might view [h] as a marker of Nmbo, and might be sensitive to its realisation in a way the Nmbo in-group are possibly not. From my observations in the field

it is not entirely clear whether the Kerake themselves view /h/ as emblematic of their language. As Dorian (2010) notes, speakers in small and dense-knit speech communities are often sensitive to variation in geographical terms (inter-village variation), but less so to local variation in individual terms (intra-village variation, p.33). Dorian's statement applies to the Kerake villages and the Morehead area more generally, where social evaluation and identification often focusses across varieties associated with different tribes or villages, rather than what speakers in one's own tribe or village are doing.

The data were drawn from the Nmbo Sociolinguistic Corpus (NSC), which was designed to sample speakers originating from one of two villages: Bevdvn or Govav. Most of the sampled speakers were living in their village of origin at the time of recording, but some of the women had relocated their residence to a different village. The NSC comprises 15 hours and 56 minutes of continuous speech (i.e., excluding pauses, gaps, and segments with no speech), with 8 hours and 15 minutes worth transcribed. The corpus comprises 39 non-randomly sampled speakers: 9 women and 12 men who originate from Bevdvn (21 total), and 8 women and 10 men who originate from Govav (18 total). The criteria for inclusion of a speaker in the corpus were the following: (1) that they are speakers of Nmbo who were primarily raised in a community of other licensed Nmbo speakers, and (2) have fathers who are also licensed speakers of Nmbo. The speakers were recruited through a locally formed language committee. The years of birth are 1948–1999. The corpus was created based on fieldwork conducted across the years 2014–2017, for a total of eight months. The trips ranged from two weeks to three months in duration.

The biographic profiles of the men are relatively straight forward. The men chosen for the NSC are all Kerake men who are from, and reside in a Nmbo village. The women's profiles are a little more complex due to the practice of virilocal exogamy. As explained earlier, seven out of the total 17 women represented in the NSC reside in non-Nmbo speaking villages. For brevity's sake I characterise these women as *daily bilinguals*, who report themselves as daily using both their own Nmbo language, as well as the language of the village. Six of the seven daily bilingual women have married into the Nen speaking village of Bimadbn. The sampled women have varying lengths of residency in their husband's villages, but only one of the seven women has lived in her husband's village for less than twenty years. All these daily bilinguals are treated as a single group for the statistical analyses.

A number of conversational topics are represented in the NSC. Topics and speech events range between naturalistic and semi-structured formats. By naturalistic I mean unprompted monologues without interruptions arising from questions by the interviewer. Semi-structured speech events involve questions and prompts by the interviewer. All speech events involved a local Nmbo



speaker, and the presence of the author. Nmbo interviewers were best-matched to the target speaker by gender, age, and village, with an assumption that matching these features would put the interviewee at ease, such that they will most likely speak freely without monitoring. This is, however, an untested assumption. In a small community where most people know one another, it is possible that some effects unknown to the outside linguist were at play (see Baclawski, 2018, p.79 for the role of community interviewer in Cham, Vietnam; Abtahian, 2018 for Garifuna, Belize).

There are three main reasons why village is taken as a unit for defining the speech community. Firstly, Nmbo speakers commonly refer to themselves on a village basis, such as *Bevdvnmn* ‘originating from Bevdvn’, *Govavnmn är* ‘a person from Govav’, perhaps akin to saying that someone comes from a specific suburb (e.g., Boxhill, Melbourne; Brooklyn, NY). Secondly, social activities such as church gatherings are often conducted and organised on a village-by-village basis. Thirdly, areas of residence are clustered in villages. While not all Kerake reside in the village at all times throughout the year, the relative permanence of housing suggests that much of daily life and socialisation is centred around the physical location of the village. It should be noted, however, that nuclear family and garden hamlet groups are a more likely space where truly unguarded and vernacular speech occurs (Kashima, 2020, p.65). The location of the true vernacular as being outside the village and on family lands is a tendency that has been noted for other Papuan communities (e.g., Stasch, 2009 for the Korowai; overview de Vries, 2012).

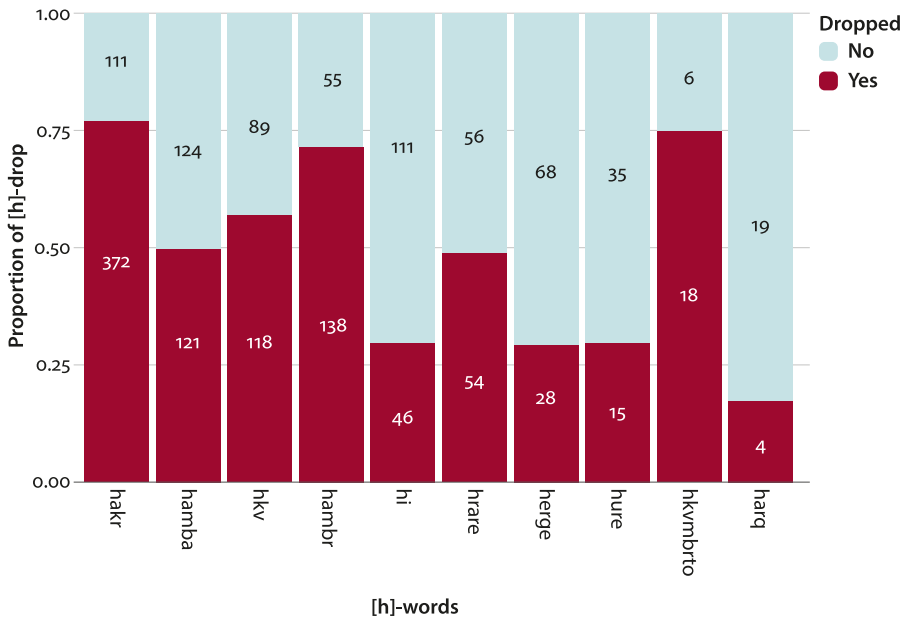
Community members were involved in the transcription and translation of recorded speech. Nmbo was transcribed into a Roman script-based practical orthography, which was then translated into English. These tasks were completed by trained local transcribers, and the author in collaboration with community consultants. Translation and orthographic transcription were conducted in ELAN (MPI, 2018), a specialist software for linguistic transcription and annotation. Transcribers were reimbursed financially for their time and knowledge, and received basic training in the use of laptops as part of the work. Speech segments are analysed on an intonation unit basis, using a modified version of the method developed at the University of California, Santa Barbara (Du Bois, Schuetze-Coburn, Cumming, & Paolino, 1993). Factors such as overlap, vocal noises, precise duration measurement, quality of the speakers’ voices, and non-linguistic sounds, were not annotated. Tokens of [h]-nouns were identified primarily by orthographic transcription and auditory perception and were further verified acoustically in some cases using Praat (Boersma & Weenink, 2016).

Logistic regression models were used to quantitatively test for relationships between predictor variables (linguistic and sociolinguistic variables) and a

numerical outcome (a binary “yes” or “no” for [h]-drop). The logistic regression models were fit using the lme4 package (Bates, Mächler, Bolker, & Walker, 2015) in the R suite (R Team, 2017). All test results are deemed significant when the chance probability of the statistic is less than 0.05.

## 4. Results

The [h]-drop data comprises of 1649 tokens across 39 speakers. 942 of these tokens show [h]-drop (57.1% of the data), while 707 do not (42.9%). Figure 2 shows the top ten most common [h]-nouns in the data, along with the proportions of drop per word.



**Figure 2.** Proportional representation of [h] drop for the ten most frequent words. Ordered from left to right, most frequent to less. 1. *hakr*=boy; 2. *hamba*=village; 3. *hkuv*=eye; 4. *hambr*=leg, foot; 5. *hi*=bush torch; 6. *hrare*=moon; 7. *herge*=life; 8. *hure*=straight, truth, true; 9. *hkuvmbtrto*=prayer; 10. *harq*=thigh

### 4.1 Linguistic conditions

The quantitative results for linguistic factors conditioning variation are weak. Nonetheless, I will discuss some of these factors in brief for three reasons. The

first is that a discussion of linguistic conditions seems pertinent for Nmbo even if the statistical results were not significant, since we know relatively little about the grammar of this under-described language. A record of some quantitative explorations of grammatical factors may be of interest for documentary purposes, not only for variation and change purposes. Second, an independent discussion of the linguistic predictors are helpful in understanding the unified statistical models further ahead, which considers both linguistic and sociolinguistic factors in explaining the patterns of variation. And finally, the relative non-results of the linguistic conditions demonstrate strongly that the propensity for [h]-drop has progressed far enough in the community that linguistic conditions which may have played a role in the past can no longer be detected.

The linguistic predictors explored here concern the environment preceding the [h] nominal; the phonetic-phonological environment, and the word class. Both predictors show weak signs of being significant. The collocation of the preceding word and the form of the [h] nominal were also considered, but these predictors do not appear significant in this data.

Frequency counts indicate that [h]-drop is the most likely outcome for most phonological environments. The proportions of [h]-drop for almost all preceding phonological environments show drop rates greater than 50% (Table 1). The number of phonological natural classes in Nmbo is quite high, with a consonant inventory of 27 phonemes and a vowel inventory of 8 (Kashima, *in press*), which makes collapsing them into principled natural classes a challenge. These were eventually narrowed down to ten levels (see caption of Table 1). Preceding rhotic and obstruents appear to favour [h]-drop the most (rhotics 66.7%,  $n=42/63$ ; voiced stops 64.7%,  $n=33/51$ ; voiceless fricatives 62.1%,  $n=18/29$ ). The two conditions with less than 50% of [h]-dropping are preceding nasals (48.7%,  $n=95/195$ ), and voiced fricatives (38.2%,  $n=13/34$ ).

The Nmbo results of preceding phonetic environment present a comparison point to the few variationist studies that look at the linguistic factors of [h]-drop. Preceding liquids and obstruents are reported as favouring [h]-drop in New Zealand English (Bell & Holmes, 1992). The Nmbo rhotics and voiced stops also show a high proportion of [h]-drop despite their infrequent appearance in the data. On the other hand Bell and Holmes report preceding nasals as favouring [h]-drop in NZE, but they seem to have little effect in the Nmbo data (48.7% dropped,  $n=95/195$ ).

The rates of [h]-drop for preceding word, and word class, were considered with the aim of trying to detect whether typical collocations might have an effect. Preceding word class is usually not considered a linguistic variable that conditions variation, but it was considered in this study to account for the slight variability in forms of the words preceding the [h]-nouns. For example, demonstratives

**Table 1.** Preceding phonological environment of [h]-words, and the number and percentage of tokens that retain or drop [h] per environment. Nmbo Open Vowels are /æ, α/, Mid Vowels are /e, u/, and Closed Vowels are /i, u/. Semi-vowel [w] is included in the Closed Vowel category

	[h] Retained		[h] Dropped		Total (n)
	n	%	n	%	
Rhotics	21	33.3%	42	66.7%	63
Voiced Stops	18	35.3%	33	64.7%	51
Open Vowels	80	36.4%	140	63.3%	220
Voiceless Fricatives	11	37.9%	18	62.1%	29
IU Breaks	223	39.9%	336	60.1%	559
Voiceless Stops	19	41.3%	27	58.7%	46
Closed Vowels	35	41.7%	49	58.3%	84
Mid Vowels	179	48.6%	189	51.4%	368
Nasals	100	51.3%	95	48.7%	195
Voiced Fricatives	21	61.8%	13	38.2%	34
<b>Total</b>	<b>707</b>		<b>942</b>		<b>1649</b>

preceding nouns often take semantic case markers such as the locative = *no* or the demonstrative ablative *-anma(e)*, which alter the phonetic form immediately preceding the head (e.g., *yna=no hamba* ‘at the village’ vs *ynan-mae hamba* ‘from the village’).

While case-marked words preceding [h]-words were not so frequent in the data, a pattern concerning word classes is identified. Frequency Table 2 shows the ten most common forms preceding [h] nouns. Among these the two possessive pronouns: *tande* (1sg) and *yānde* (3sg) show a slightly disfavoured of [h]-drop with rates of 38.1% ( $n=24/63$ ) and 44.7% ( $n=21/47$ ) respectively. This goes against the trend of the other word classes which show rates of [h]-drop higher than 50%. Given the earlier results of preceding phonological environment, it is unlikely that the similarity in phonological form of the possessive pronouns are giving rise to these results. Furthermore, the phonological form of the word *toge* ‘child’ is very similar to the possessive pronouns, but *toge* shows much higher rates of [h]-drop ( $n=37/52$ , 71.2%). Table 3 is a frequency table showing the rates for preceding word class, which suggests that some patterns are present. Noticeably, possessive forms clearly disavour [h]-drop with a rate well below 50% ( $n=73/183$ , 39.9%).

**Table 2.** Ten most common words that precede [h] words ranked from most frequent to least. Excluding instances where the [h] word is preceded by an IU boundary

Word	[h] Retained		[h] Dropped		Total (n)
	n	%	n	%	
1 <i>yna</i> ‘this’	38	41.8%	53	58.2%	91
2 <i>tande</i> ‘my’	39	61.9%	24	38.1%	63
3 <i>toge</i> ‘small’, ‘child’	15	28.8%	37	71.2%	52
4 <i>yānde</i> ‘his’, ‘hers’	26	55.3%	21	44.7%	47
5 <i>är</i> ‘man’, ‘person’	8	21.1%	30	78.9%	38
6 <i>āmb</i> ‘some’	10	30.30%	23	69.7%	33
7 <i>bā</i> 3ABS, FUT	14	45.2%	17	54.8%	31
8 <i>dena</i> ‘like this’	2	11.1%	16	88.9%	18
9 <i>mrz</i> ‘girl’, ‘daughter’	8	50.0%	8	50.0%	16
10 <i>okei</i> ‘okay’	7	46.7%	8	53.3%	15
11 <i>yao</i> NEG	6	40.0%	9	60.0%	15

A reviewer suggested that the preceding word class results may be an epiphenomenon of prosodic position, since prosodic position is known to have an effect on phonological variation (Cho, 2016; Katz, 2016). If this were the case for the Nmbo data, however, one would expect words that occupy the same prosodic environment to have similar patterns of following [h] variability. Demonstratives are disyllabic words that precede the head (e.g., *yna* /jɛ.na/ ‘this’, *dena* /de.na/ ‘like this’), as are possessive pronouns (e.g., *tande* [tan.de] 1sg, *yānde* /jæn.de/ 3sg). Yet as the numbers in frequency Table 3 show, demonstratives and possessive pronouns have [h]-drop rates in opposite directions; the former shows relatively high rates of [h]-drop ( $n=87/142$ . 61.3%) while the latter shows lower rates ( $n=73/183$ , 39.9%). In addition, the IU (ends of IU units, pauses and breaks) category with its [h]-drop rate of 60% suggests that the effect of prosody may not be so great overall.

We can only speculate as to why preceding possessive forms in particular appear to disfavour [h]-drop.<sup>4</sup> It is possible that this pattern is purely by chance,

4. Like its sister languages Nen (Nambu branch) and Komnzo (Tonda branch), Nmbo has two forms of the possessive suffix. The regular form *-ende* is the form under discussion in this paper, and the alternative form *-nzo* is described as a “close possessive”. The close possessive form is often used to mark kin terms (e.g., *tanzo ama* ‘my real mother’) however its use is non-obligatory and unrelated to the concept of inalienability. Nmbo speakers describe the distinc-

**Table 3.** Preceding word subclass of [h]-words in the NSC. GRM = Grammatical Words (conjunctions, TAM particles, quotatives); MOD = Modifiers (numerals, quantifiers, temporal adverbials, negative particle, proper names, and nouns used attributively); DEM = demonstratives; IU = final IU boundary marker; V = Verbs (inflected); DISC = discourse particles (topic markers, affirmatives; N = Nouns (inflected, nominalised verbs); PRO = Pronouns (personal pronouns excluding the possessive form); POSS = possessives (possessive pronouns)

	[h] Retained		[h] Dropped		Total
	n	%	n	%	n
GRM	9	21.4%	33	78.6%	42
MOD	54	36.0%	96	64.0%	150
DEM	55	38.7%	87	61.3%	142
IU	223	39.9%	336	60.1%	559
V	94	42.5%	127	57.5%	221
DISC	23	43.4%	30	56.6%	53
N	101	45.1%	123	54.9%	224
PRO	38	50.7%	37	49.3%	75
POSS	110	60.1%	73	39.9%	183
<b>Total</b>	<b>707</b>		<b>942</b>		<b>1649</b>

although we will see in the following section that a unified model including sociolinguistic variables still shows preceding word class as significant. What seems to be clear is that the overall lack of strong linguistic conditioning results suggest that linguistic environment is no longer so important as to be a major consideration in differential rates of [h]-drop. Indeed as we will see below, age is the single best predictor in explaining variation in [h]-drop.

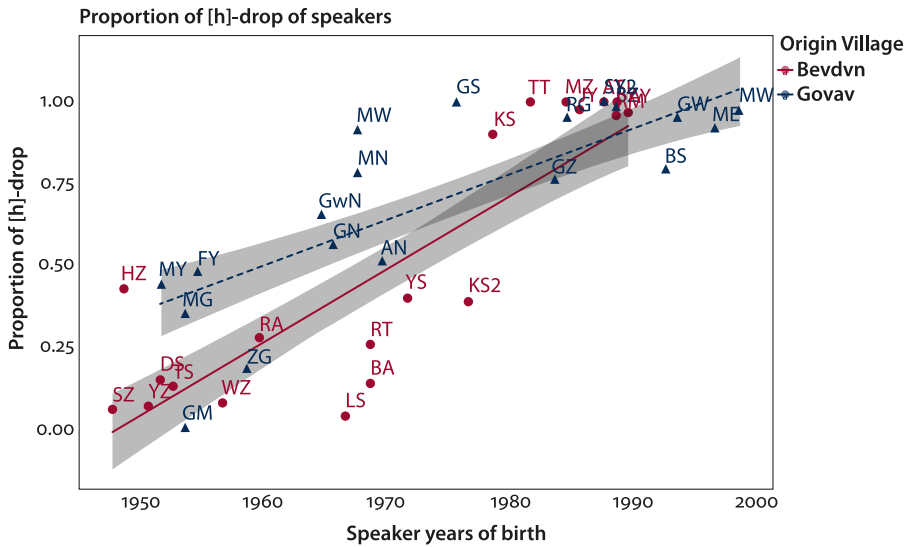
## 4.2 Sociolinguistic and linguistic results

Age is the best predictor in explaining differential rates of [h]-drop in the data. Speaker year of birth (YOB) is used to represent speaker age in this dataset. A fre-

---

tion in usage between the regular and close possessive forms as akin to saying something is “really mine” (Kashima, 2020, p. 124). The close possessive forms can be used to indicate possession of any nominal, and not just kin terms, eg. *tanzo hamba* ‘my real village / the village I originate from’. The speculation is that the use of these possessive form, and the expressions of possession in general, are marked for Nmbo speakers in some way, and that this has consequences for the realisation of possessed [h] initial nominals.

quency plot of proportion of [h]-drop per speaker by YOB shows very clearly an age-related trend. Figure 3 shows older speakers towards the left side of the graph with lower proportion of [h]-drop, with a Pearson’s correlation test returning a strong relationship of .82 between YOB and [h]-drop proportion. In other words, older speakers are more [h]-ful, and produce fewer instances of [h]-drop. On the other hand younger speakers towards the right of the graph show higher proportions of [h]-less words. Visual inspection of the graph suggests that the group showing the most variability are those born between 1965 and 1973. All speakers born after 1973 show a proportion of h-drop greater than .5 (i.e., 50% of the time).



**Figure 3.** Proportion of [h]-drop of speakers, by year of birth. Speaker organised by village of origin (VoO)

Unified models with linguistic and social predictors were run. Drager and Hay (2012) suggest using random intercepts to account for the effects of individual speaker, but such models failed to converge for the Nmbo data. While speaker as a random effect explains much of the variability in many datasets generally (e.g., Meyerhoff & Walker, 2007), a model without random effects is still beneficial. Such a model provides a holistic account that considers as many variables as possible.

A logistic regression was modelled with the dependent variable set as a binary “yes/no” for [h]-drop, and the independent variables of preceding word class, speaker village of origin and year of birth. A version of the model was run with speaker sex included as a variable along with Village of Origin and YOB, but the

results were barely significant for sex and weakened the results of the other variables. This model will be discussed further on in the context of speaker gender more broadly. The two social factors were tested as interactions. The large range in coefficient values and standard error indicates the fragility of the model (Table 4), but the results mostly confirm what we could infer from the frequency results. Age is the single best predictor for the rates of [h]-drop, and preceding possessive forms disfavour [h]-drop. The results of the Village of Origin (VoO) have a highly significant  $p$  value but given the values of the standard error we should be cautious to read too much at this stage ( $SE=23.4$ ,  $p<.001$ ). The interaction between VoO and YOB returns a  $p$  value below .05, but yet again, the SE is quite high ( $SE=118.9$ ,  $p<.001$ ).

**Table 4.** Logistic regression model fit summary on speakers aged over 45. [h]-drop YES vs NO, fixed effects of Preceding Word Class, and VoO. Speaker and Word as random intercepts. IU=final IU boundary marker; V=Verbs (inflected); DISC=discourse particles (topic markers, affirmatives; GRM=Grammatical Words (conjunctions, TAM particles, quotatives); MOD=Modifiers (numerals, quantifiers, temporal adverbials, negative particle, proper names, and nouns used attributively); N=Nouns (inflected, nominalised verbs); PRO=Pronouns (personal pronouns excluding the possessive form); POSS=possessives (possessive pronouns). Observations=977, 21 speakers, 20 words. Significance codes:  $p<.000$ ='\*\*\*' –  $.005$ ='\*\*',  $.01$ – $.05$ ='\*'

Fixed Effects	Coefficient	St. Error	Z-Value	p
Intercept	0.029	19.11	-15.24	$2 \times 10^{-16}$ ***
Linguistic Predictors				
POSS	-0.7	0.29	-2.45	0.0145 *
Social Predictors				
YOB	0.15	$9.69 \times 10^{-3}$	15.26	$2 \times 10^{-16}$ ***
Village of Origin: Govav	120.7	23.4	5.152	$2.58 \times 10^{-7}$ ***
Village*YOB	-0.06	118.9	-5.111	$3.2 \times 10^{-7}$ ***

In order to investigate whether there are conditioning factors other than age, a logistic regression was run on a subset of the sample. Tokens stemming from those born before 1973 were investigated for this end (speaker  $n=20$ , token  $n=977$ ). The dependent variable was set as a binary “yes/no” for [h]-drop, while the independent variables were speaker VoO, and speaker gender. This time speaker and word (i.e., lexical item of [h]-word) were set as random intercepts and the model converged. For the speakers born before 1973 (i.e., aged over 45), we find that preceding possessives disfavour [h]-drop ( $SE=.4$ ,  $p<.03$ ). VoO is also significant



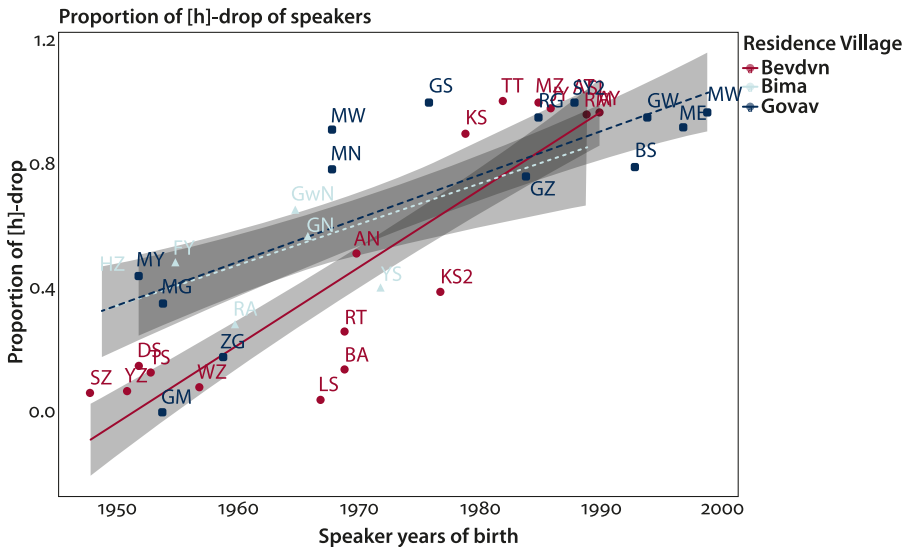
( $SE=.6, p<.03$ ), with speakers originating from Govav favouring [h]-drop (summary Table 5). Visual inspection of the frequency graph supports the statistical output of VoO as significant. A linear regression fitted for speaker VoO shows the proportion of [h]-drop beginning at higher rates on average by Govav speakers, with an increase in [h]-drop frequency over time. The rate of [h]-drop by Bevdvn villagers begins with lower frequencies than Govav villagers, until those born after 1979 show rates comparable to their Govav peers.

**Table 5.** Logistic regression model fit summary on speakers aged over 45. [h]-drop YES vs NO, fixed effects of Preceding Word Class, and VoO. Speaker and Word as random intercepts. IU=final IU boundary marker; V=Verbs (inflected); DISC=discourse particles (topic markers, affirmatives; GRM=Grammatical Words (conjunctions, TAM particles, quotatives); MOD=Modifiers (numerals, quantifiers, temporal adverbials, negative particle, proper names, and nouns used attributively); N=Nouns (inflected, nominalised verbs); PRO=Pronouns (personal pronouns excluding the possessive form); POSS=possessives (possessive pronouns). Observations=977, 21 speakers, 20 words. Significance codes:  $p<.000$  = ‘\*\*\*’,  $.005$  = ‘\*\*’,  $.01-.05$  = ‘\*’

Random Effect	Variance	St. Dev.		
Speaker	1.67	1.29		
Word	1.20	1.09		
Fixed Effects	Coefficient	St. Error	Z-Value	p
Intercept	-1.69	0.60	-2.80	0.005**
POSS	-0.90	0.40	-2.28	0.022*
V	-0.08	0.39	-0.20	0.839
GRM	0.56	0.80	0.70	0.484
IU	-0.20	0.35	-0.57	0.563
PRO	-0.19	0.52	-0.36	0.720
DISC	-0.26	0.57	-0.46	0.644
MOD	-0.93	0.48	-1.94	0.053
N	0.04	0.39	0.11	0.913
Village: Govav	1.32	0.60	2.19	0.028*

For speakers born after 1973 (speaker  $n=18$ , token  $n=672$ ) neither the word class nor village were significant ( $p>.1$ ). These results suggest strongly that the linguistic conditioning factors have indeed fallen away for younger speakers, and we have evidence of a change-in-progress phenomenon.

Village of residence (VoR) was modelled to investigate whether daily bilingual women would pattern differently as a group. Statistical modelling could not be conducted due to low token counts, but upon visual inspection (Figure 4) we can see the regression line for Govav residents and Bimadbn residents look very similar to each other. We cannot say much more with certainty, but it appears that the Bimadbn residing women who originate from Govav are patterning much like the other sampled Govav residents. The ethnographic observations of out-marrying women maintaining interaction with their origin village and community form a partial basis for this analysis.



**Figure 4.** Proportion of [h]-drop of speakers, by year of birth. Speakers organised by village of residence (VoR)

We do not find speaker gender effecting [h]-drop variability in this data set. Various statistical modelling was attempted, but these either failed to converge or had non-significant results. A version of the holistic model was run with speaker sex included as an additional variable along with VoO and YOB, but the results were barely significant and weakened the results of the other variables. A naive model with gender as the sole independent variable, and speaker and word as random effects, also did not find gender to be significant. The implications of these will be discussed in the following section.

## 5. Analysis and discussion

Statistical modelling shows that age is the strongest predictor of who will drop [h], with younger speakers being near categorical [h]-droppers. The apparent time construct (Labov, 1972, c.f. Cukor-Avila & Bailey, 2013) suggests there is a change in progress, given the reliability of the construct where there are strong s-curve distributions (Blythe & Croft, 2012; Pope, Meyerhoff, & Ladd, 2007). Evidence suggests that change has progressed quite far in the speech community. The results of the statistical modelling which tested for the effects of linguistic conditioning were different for speakers over the age of 45, and under 45. Older speakers show some evidence of still being affected by linguistic conditions (in this case preceding word class) while younger speakers do not show this pattern. Since speakers under the age of 45 show no linguistic conditioning, nor any social conditioning affecting [h]-drop, it seems that [h]-lessness has diffused quite broadly through the community.

There is some evidence that the origin of [h]-drop began in the village of Govav and then disseminated to Bevdvn. Statistical modelling shows that the older Govav speakers are the ones with higher propensity of [h]-drop compared to their peers from Bevdvn. The rate of [h]-drop is consistently slightly higher for Govav villagers, until Bevdvn villagers born after 1973 catch up with similar proportions. It may be pure chance that Govav village was the origin of the [h]-drop, since there is no reason to assume that a process such as [h]-drop requires some inciting conditions other than the natural processes of articulatory laxity. But the characteristic of Govav as the more multilingual village compared to Bevdvn makes it the likely origin community for the innovation for two reasons. Firstly, languages spoken in highly multilingual communities have been suggested as having a large *feature pool* (Mufwene, 2001) from which innovations can arise. Cheshire et al. (2011) suggest for multi-ethnic London that high linguistic diversity leads to a heterogeneous feature pool from which new variants are introduced into the speech community. Similar findings come from non-hierarchical multilingual and multiethnic contexts of India (Satyanath, 2018; Suokhrie, 2016). The feature pool, however, seems to require some direct linguistic input into the pool from which speakers then choose features to appropriate. A direct analogy for the Nmbo case would be that some language that does not have word-initial [h] for cognate Nmbo words affects the Nmbo forms. There are no specific Nambu branch languages in the areas immediately adjacent to Govav that have altogether lost word-initial phones that correspond to Proto-Nambu \*s. A weaker version of the feature pool analysis would be that the mere presence of multiple inputs is conducive to phonetic innovations compared to a community which does not have diverse inputs.

Govav's location within the village network also makes it the candidate as the origin of [h]-drop. As a multi-link node in the network of Morehead Area villages, Govav is ideally placed to spread innovations rapidly if one arises. Govav has the highest network connectivity with other villages through marriage, which leads Rueck (2011) to suggest that Govav "is the best place to introduce innovations if one wants them to spread" throughout the network rapidly (p. 89). When we look at the links between Govav and Bevdvn specifically, there are five women who have married from Govav to Bevdvn, while there are only two women who have married from Bevdvn into Govav. The greater multiplexity of ties from Govav to Bevdvn might predict a more frequent and/or larger flow of people going from Govav to Bevdvn for short term visits. This is then a case of diffusion and change, where the younger speakers acquired [h]-dropped lexical nouns as part of their input, either through Govav villagers visiting Bevdvn, or by Bevdvn children following their mothers back to Govav and receiving input there.

Gender effects were not found to affect [h]-drop variability in this data. This result is of interest given that gender often manifests in variation and change phenomena across various speech communities ranging from urbanised Sweden (Nordberg & Sundgren, 1998) to regional dialect of K'iche' (Mayan; Guatemala, Romero, 2009). The lack of gender variation in Nmbo [h]-drop re-emphasises the findings that *gender* reflects complex and highly localised interactions of various social dimensions in any given speech community (Meyerhoff, 1999, pp. 233–234, Stanford 2009a).

The Nmbo [h]-drop results concerning gender can thus be interpreted in a few ways. One is that linguistic change can diffuse through the community via non-gendered connections. Another interpretation is that there were perhaps clearer gender patterns in the past, but this is no longer visible in the current data. The implication of this latter interpretation is that diffusion of [h]-drop happened quite rapidly through both genders. The first interpretation is perhaps more likely given that we can still detect village-based difference. If gender were meaningful in the past, we may expect the data to show some trace of this as we can see with the village of origin results.

The lack of a gender distinction also suggests that the daily-bilingual women are patterning very closely with their peers who continue to reside in Nmbo villages. The non-result of village of residence also supports this picture. The similarity in patterning can again be understood as due to the high degree of maintained interaction between individuals in the speech community. The notion of *linguistic loyalty*, and more specifically, the notion of *communities of descent*, may also be relevant. Stanford (2009b) describes the strict adherence of out-marrying Sui women to their clan-lects as a manifestation of a local ideology that states how one should speak the variety that is prescribed to members of particular clans.

The communities of the Morehead Area also have a stated ideology where one should speak their tribal/village variety. Stanford and Pan (2013) have described this adherence of linguistic loyalties to one's descent or origin group under the label *the sociolinguistics of exogamy*, which is found in communities that practice some form of exogamy (e.g., Suokhrrie, 2016 for Angami speakers, Nagaland, India; Campbell & Grondona, 2010 for Misión La Paz in Argentina). The label is apt for this Nmbo situation also, given the majority post-marital residency pattern for Nmbo women is that of section exogamy which often results in village and linguistic exogamy.

The degree to which egalitarian multilingualism plays a role in this particular phenomenon is still an open question. But the non-negative attitudes towards speaking one's tribal language, may be contributing to the daily bilingual women's ability to follow certain linguistic patterns of their tribal peers even after postmarital relocation. The daily bilingual women who continue to speak Nmbo first and foremost identify themselves as Nmbo speakers, and express no negative attitudes towards the continued use of their tribal language (Kashima, 2020, pp.55–61). Their husbands, and their postmarital residence village, also accept the continued use of their wives' language. As Potowski (2013) puts it when describing the forces of language maintenance in general, "positive attitudes are not enough to guarantee language maintenance, [but] negative attitudes [are known to] lead to rapid shift" (p.323). We may infer from the absence of negative attitudes towards multilingualism in the Morehead area that egalitarian multilingual ideals at least do not lead people to abandon the use of their tribal language when moving away from their village of origin.

## 6. Conclusion

The primary goal of this study is to contribute to our growing understanding of the variety of socio-cultural pressures involved in the transmission and diffusion of innovations in linguistic change phenomena. The Nmbo [h]-drop has progressed in the speech community such that we can no longer detect any clear linguistic conditioning effects, though there is some evidence that preceding word class had an effect in the past. Social conditioning other than age also proves difficult to capture. Nonetheless, if one invokes the apparent time construct to analyse the Nmbo data, it appears that the larger village with the more multilingual profile is the origin of the [h]-drop. Gender is non-significant in this data. Correspondingly we find no evidence that the daily bilingual Nmbo women are ahead or lagging in their [h]-drop rates; they are patterning very closely to the rest of the speech community that continue to reside in Nmbo speaking villages. This study

provides a rare glimpse into the patterns of a change-in-progress from an understudied language context and hopes to be one of the many future works to empirically document how language change phenomena progress in a complex language ecology.

## Funding

This study was conducted under the auspices of Professor Nicholas Evans's Australian Research Council Laureate Project: The Wellsprings of Linguistic Diversity (Grant No. FL130100111). In the same vein I would like to thank the ARC Centre of Excellence for the Dynamics of Language for the additional support I have received.

## Acknowledgements

My sincerest thank you goes to the Kerake people of the South-Fly District of Western Province, Papua New Guinea, with a special thanks to transcribers and translators Mr. Bunai Aniba, Mr. Lawrence Mike, Mr. Kawas Säme, Mr. Gima Zoga, Mr. Richard Gima, Ms. Rhouda Timoti, Ms. Jacklyn Yavs, Ms. Rachel Midawa, Ms. Towal Taga, Mr. Lazarus Smako, Mr. Tano Sam, Mr. Mwigm Nuwira, Ms. Anna Dawi, Mr. Toka Arura, Mr. Wayni Mege, Ms. Ella Awasi, Mr. Kata Suga, Mr. Karuwa Wagra, Mr. Weko Sovam, Mr. Zoga Gima, and Ms. Yumer Yäkimo. Additional thanks goes to the two anonymous reviewers whose comments have made this paper much stronger, and Dr. Yoko Yonezawa for her suggestions to the Japanese version of the abstract.

## References

- Abtahian, Maya (2018). *Style and language maintenance*. Paper presented at Sociolinguistic Symposium 22, Long Colloquium: *Fresh insights on traditional variationist methods in non-English contexts*, Auckland, New Zealand, June 28, 2018.
- Ayres, Mary C. (1983). *This side, that side: Locality and exogamous group definition in Morehead area, southwestern Papua*. Unpublished doctoral dissertation, University of Chicago.
- Baclawski, Kenneth Jr. (2018). Diglossia and change from below in Eastern Cham. *Asia-Pacific Language Variation*, 4(1), 73–102. <https://doi.org/10.1075/aplv.17003.bac>
- Bates, Douglas, Mächler, Martin, Bolker, Ben, & Walker, Steve (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bell, Allan, & Holmes, Janet (1992). H-droppin': Two sociolinguistic variables in New Zealand English. *Australian Journal of Linguistics*, 12(2), 223–248. <https://doi.org/10.1080/07268609208599478>
- Blythe, Richard A., & Croft, William (2012). S-curves and the mechanisms of propagation in language change. *Language*, 88(2), 269–304. <https://doi.org/10.1353/lan.2012.0027>

- Boersma, Paul, & Weenink, David (2016). *Praat: Doing phonetics by computer* [Software program]. Amsterdam: University of Amsterdam.
- Campbell, Lyle, & Grondona, Verónica (2010). Who speaks what to whom? Multilingualism and language choice in Misión La Paz. *Language in Society*, 39(5), 617–646.  
<https://doi.org/10.1017/S0047404510000631>
- Cheshire, Jenny, Kerswill, Paul, Fox, Sue, & Torgersen, Eivind (2011). Contact, the feature pool and the speech community: The emergence of Multicultural London English. *Journal of Sociolinguistics*, 15(2), 151–196. <https://doi.org/10.1111/j.1467-9841.2011.00478.x>
- Childs, Tucker, Good, Jeff, & Mitchell, Alice (2014). Beyond the ancestral code: Towards a model for sociolinguistic language documentation. *Language Documentation and Conservation*, 8, 168–191.
- Cho, Taehong (2016). Prosodic boundary strengthening in the phonetics-prosody interface. *Language and Linguistics Compass*, 10(3), 120–141. <https://doi.org/10.1111/lnc3.12178>
- Cukor-Avila, Patricia, & Bailey, Guy (2013). Real time and apparent time. In Jack K. Chambers & Natalie Schilling-Estes (Eds.), *The handbook of variation and change* (pp. 239–262). Hoboken, NJ: John Wiley and Sons. <https://doi.org/10.1002/9781118335598.ch11>
- De Vries, Lourens (2012). Speaking of clans: Language in Awyu-Ndumut communities of Indonesian West Papua. *International Journal of the Sociology of Language*, 2012(214), 5–26.
- Döhler, Christian (2018). *A Grammar of Komnzo*. Berlin: Language Science Press.
- Dorian, Nancy C. (2010). *Investigating variation: The effects of social organization and social setting*. Oxford: Oxford University Press.  
<https://doi.org/10.1093/acprof:oso/9780195385939.001.0001>
- Drager, Katie, & Hay, Jennifer (2012). Exploiting random intercepts: Two case studies in sociophonetics. *Language Variation and Change*, 24(1), 59–78.  
<https://doi.org/10.1017/S0954394512000014>
- Du Bois, John W., Schuetze-Coburn, Stephan, Cumming, Susanna, & Paolino, Danae (1993). Outline of discourse transcription. In Jane Edwards & Martin Lampert (Eds.), *Talking data: Transcription and coding in discourse* (pp. 45–89). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Epps, Patience (2018). Contrasting linguistic ecologies: Indigenous and colonially mediated language contact in northwest Amazonia. *Language and Communication*, 62, 156–169.  
<https://doi.org/10.1016/j.langcom.2018.04.010>
- Evans, Nicholas (2012). Even more diverse than we had thought: The multiplicity of Trans-Fly languages. *Language Documentation and Conservation Special Publication No. 5*, 109–149.
- Evans, Nicholas (2015a). Restricted phonemes and interaction in Nen. *Workshop on the Languages of Melanesia*, Kioloa Australia.
- Evans, Nicholas (2015b). Inflection in Nen. In Matthew Baerman (Ed.), *Oxford handbook of inflection* (pp. 543–575). Oxford: Oxford University Press.
- Evans, Nicholas (2015c). Valency in Nen. In Andrej Malchukov & Bernard Comrie (Eds.), *Valency classes in the world's languages: Case studies from New Guinea, Australia, and the Americas, and theoretical outlook* (pp. 1069–1116). Berlin: De Gruyter Mouton.
- Evans, Nicholas (2017a). Did language evolve in multilingual settings? *Biology and Philosophy*, 32(6), 905–933. <https://doi.org/10.1007/s10539-018-9609-3>

- Evans, Nicholas (2017b). Quantification in Nen. In Denis Paperno & Edward Keenan (Eds.), *Handbook of quantifiers in natural language, volume 2* (pp. 571–607). New York, NY: Springer. [https://doi.org/10.1007/978-3-319-44330-0\\_11](https://doi.org/10.1007/978-3-319-44330-0_11)
- Evans, Nicholas (2018). The dynamics of language diversity. In Rajend Mesthrie & David Bradley (Eds.), *Plenary and focus lectures from the 20th International Congress of Linguists* (pp. 11–41). Cape Town, South Africa: UCT Press.
- Evans, Nicholas, Arka, Wayan I., Carroll, Matthew J., Choi, Yun Jung, Döhler, Christian, Gast, Volker, Kashima, Eri, Mittag, Emile, Olsson, Bruno, Quinn, Kyla, Schokkin, Dineke, Tama, Phillip, van Tongeren, Charlotte, & Siegel, Jeff (2018). The languages of southern New Guinea. In Bill Palmer (Ed.), *The languages and linguistics of New Guinea: A comprehensive guide* (pp. 640–774). Berlin: De Gruyter Mouton.
- Evans, Nicholas, Carroll, Matthew J., & Döhler, Christian (2017). Proto-Yam phonology languages of Southern New Guinea. In *The Languages of Melanesia Workshop*. Manokwari, Indonesia.
- François, Alexandre (2011). Social ecology and language history in the northern Vanuatu linkage: A tale of divergence and convergence. *Journal of Historical Linguistics*, 1(2), 175–246. <https://doi.org/10.1075/jhl.1.2.03fra>
- François, Alexandre (2012). The dynamics of linguistic diversity: Egalitarian multilingualism and power imbalance among northern Vanuatu languages. *International Journal of the Sociology of Language*, 214, 85–110.
- Haudricourt, André G. (1961). Richesse en phonèmes et richesse en locuteurs [Number of phonemes and number of speakers]. *L'Homme*, 1, 5–10.
- Hildebrandt, Kristine A., Jany, Carmen, & Silva, Wilson (Eds.). (2017). Documenting variation in endangered languages. *Language Documentation and Conservation Special Publication*, 13, 1–5.
- Hitchcock, Garrick (2010). Mound-and-ditch taro gardens of the Bensbach or Torassi River area, southwest Papua New Guinea. *The Artefact*, 33, 70–90.
- Kashima, Eri. (in press). The phonetics of Nmbo (Nembo) with some comments on its phonology (Yam Family; Morehead District). *Language Documentation and Conservation Special Publication*.
- Kashima, Eri (2020). *Language in my mouth: Linguistic variation in the Nmbo speech community of Southern New Guinea*. Unpublished doctoral dissertation, Australian National University.
- Kashima, Eri, Williams, Daniel, Ellison, T. Mark, Schokkin, Dineke, & Escudero, Paola (2016). Uncovering the acoustic vowel space of a previously undescribed language: The vowels of Nambo. *Journal of the Acoustical Society of America*, 139(6). <https://doi.org/10.1121/1.4954395>
- Katz, Jonah (2016). Lenition, perception and neutralisation. *Phonology*, 33(1), 43–85. <https://doi.org/10.1017/S0952675716000038>
- Labov, William (1965). On the mechanisms of linguistic change. *Georgetown Monographs on Language and Linguistics*, 18, 91–114.
- Labov, William (1972). The social motivations of sound change. In *Sociolinguistic patterns* (pp. 1–42). Philadelphia, PA: University of Pennsylvania Press.
- Labov, William (2007). Transmission and diffusion. *Language*, 83(2), 344–387. <https://doi.org/10.1353/lan.2007.0082>
- Lüpke, Friederike (2016). Uncovering small-scale multilingualism. *Critical Multilingualism Studies*, 4(2), 35–74.



- Lüpke, Friederike (2018). Multiple choice: Language use and cultural practice in rural Casamance between convergence and divergence. In Jacqueline Knörr & Wilson Trajano Filho (Eds.), *Creolization and pidginisation in contexts of post-colonial diversity: Language, culture, and identity* (pp. 181–208). Leiden: Brill. [https://doi.org/10.1163/9789004363397\\_011](https://doi.org/10.1163/9789004363397_011)
- Mansfield, John B. (2015). Consonant lenition as a sociophonetic variable in Murrinh Patha (Australia). *Language Variation and Change*, 27(2), 203–225.
- MPI (2018). *ELAN (Version 5.2)*. The Language Archive. Nijmegen: Max Planck Institute for Psycholinguistics.
- Meyerhoff, Miriam (1999). Sorry in the Pacific: Defining communities, defining practice. *Language and Society*, 28(2), 225–238. <https://doi.org/10.1017/S0047404599002055>
- Meyerhoff, Miriam, & Walker, James A. (2007). The persistence of variation in individual grammars: Copula absence in ‘urban sojourners’ and their stay-at-home peers, Bequia (St Vincent and the Grenadines). *Journal of Sociolinguistics*, 11(3), 346–366. <https://doi.org/10.1111/j.1467-9841.2007.00327.x>
- Meyerhoff, Miriam (2017). Writing a linguistic symphony: Analyzing variation while doing language documentation. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, 62(4), 525–549. <https://doi.org/10.1017/cnj.2017.28>
- Mufwene, Salikoko S. (2001). *The Ecology of language evolution*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511612862>
- Nordberg, Bengt, & Sundgren, Eva (1998). *On observing real-time language change: A Swedish case study*. Uppsala: Uppsala Universitet.
- Pope, Jennifer, Meyerhoff, Miriam, & Ladd, Robert (2007). Forty years of language change on Martha’s Vineyard. *Language*, 83(3), 615–627. <https://doi.org/10.1353/lan.2007.0117>
- Potowski, Paul (2013). Language maintenance and shift. In Robert Bayley, Richard Cameron, & Ceil Lucas (Eds.), *The Oxford handbook of sociolinguistics* (pp. 321–339). Oxford: Oxford University Press.
- Romero, Sergio (2009). Phonological markedness, regional identity, and sex in Mayan: The fricativization of intervocalic /l/ in K’iche’. In James N. Stanford & Dennis R. Preston (Eds.), *Variation in indigenous minority languages* (pp. 281–297). Amsterdam: John Benjamins. <https://doi.org/10.1075/impact.25.14rom>
- Ross, Malcolm (2005). Pronouns as a preliminary diagnostic for grouping Papuan languages. In Andrew Pawley, Robert Attenborough, Jack Golson, & Robert Hide (Eds.), *Papuan pasts: Cultural, linguistic and biological histories of Papuan-speaking people* (pp. 15–66). Canberra: Pacific Linguistics.
- R Team (2017). *R: A language and environment for statistical computing*. R foundation for statistical computing. Vienna, Austria. Retrieved from <https://www.r-project.org/>
- Rueck, Michael J. (2011). *Social network analysis applied to language planning in the Morehead District, Papua New Guinea (Electronic Survey Report 2011, 37)*. Dallas, TX: SIL International.
- Salisbury, Richard F. (1962). Notes on bilingualism and linguistic change in New Guinea. *Anthropological Linguistics*, 4(7), 1–13.
- Sankoff, Gillian (2001). Linguistic outcomes of contact. In Peter Trudgill, Jack K. Chambers, & N. Schilling-Estes (Eds.), *Handbook of sociolinguistics* (pp. 638–668). Oxford: Basil Blackwell.

- Sankoff, Gillian (1980). Multilingualism in Papua New Guinea. In *The social life of language* (pp. 95–132). Philadelphia, PA: University of Pennsylvania Press.  
<https://doi.org/10.9783/9781512809589-008>
- Satyanath, Shobha (2018). Language variation and change in a small but diverse city in India. In Dick Smakman & Patrick Heinrich (Eds.), *Urban sociolinguistics: The city as a linguistic process and experience* (95–112). London: Routledge.
- Siegel, Jeff (2014). The morphology of tense and aspect in Nama, a Papuan language of southern New Guinea. *Open Linguistics*, 1(1), 211–231.
- Siegel, Jeff (2017). Transitive and intransitive verbs in Nama, a Papuan language of Southern New Guinea. *Oceanic Linguistics*, 56(1), 123–142. <https://doi.org/10.1353/ol.2017.0005>
- Singer, Ruth (2018). A small speech community with many small languages: The role of receptive multilingualism in supporting linguistic diversity at Waruwi Community (Australia). *Language and Communication*, 62, 102–118.  
<https://doi.org/10.1016/j.langcom.2018.05.002>
- Stasch, Rupert (2009). *Society of others: Kinship and mourning in a West Papuan place*. Berkeley, CA: University of California Press.
- Stanford, James N. (2009a). Clan as a sociolinguistic variable: Three approaches to Sui clans. In James N. Stanford & Dennis R. Preston (Eds.), *Variation in indigenous minority languages* (pp. 463–484). Amsterdam: John Benjamins.  
<https://doi.org/10.1075/impact.25.23sta>
- Stanford, James N. (2009b). “Eating the food of our place”: Sociolinguistic loyalties in multidialectal Sui villages. *Language in Society*, 38(3), 287–309.  
<https://doi.org/10.1017/S0047404509090502>
- Stanford, James N., & Yanhong Pan (2013). The sociolinguistics of exogamy: Dialect acquisition in a Zhuang village. *Journal of Sociolinguistics*, 17(5), 573–607.  
<https://doi.org/10.1111/josl.12052>
- Suokhrie, Kelhouvinuo (2016). Clans and clanlectal contact: Variation and change in Angami. *Asia-Pacific Language Variation*, 2(2), 188–214. <https://doi.org/10.1075/aplv.2.2.04suo>
- Thije, Jan D. ten, & Zeevaert, Ludgar (2007). *Receptive multilingualism: Linguistic analyses, language policies and didactic concepts*. Amsterdam: John Benjamins.  
<https://doi.org/10.1075/hsm.6>
- Torres Cacoullos, Rena, & Travis, Catherine E. (2018). *Bilingualism in the community: Code-switching and grammars in contact*. Cambridge: Cambridge University Press.  
<https://doi.org/10.1017/9781108235259>
- Tucker, Sara, Boevé, Marco, Fuller, Elizabeth, Gustafsson, Catharina, & Rueck, Michael (2003). *Nambu Subfamily survey report (Tech. Rep.)*. Ukarumpa: Summer Institute of Linguistics.
- Vaughan, Jill, & Singer, Ruth (2018). Indigenous multilingualisms past and present. *Language and Communication*, 62, 83–90. <https://doi.org/10.1016/j.langcom.2018.06.003>
- Williams, Francis E. (1936). *Papuans of the Trans-Fly*. Oxford: Clarendon Press.
- Wurm, Stephen A., & Hattori, Shirō (1981). *Language atlas of the Pacific area, Part 1 and 2*. Pacific Linguistics, Series C, 66 and 67. Canberra, Australia: ANU.

Appendix A. Speakers from NSC, token count for [h]-drop

Origin village: Bevdvn					
Speaker code	Year of birth	Gender	Residence	Tokens (n)	[h]-drop ratio
ymbbME	1997	M		13	0.92
ymggBS	1993	M		68	0.79
yfbbAY	1990	F		38	0.97
ymmbSL	1989	M		28	1
yfabRM	1989	F		45	0.96
yfbbAT	1988	F		35	1
yfbbJY	1986	F		42	0.98
ymbbMZ	1985	M		16	1
ymbbTT	1982	M		38	1
mmbbKS	1979	M		49	0.9
mmbbKS2	1977	M		41	0.39
mfbmYS	1972	F	Nen village	57	0.4
mfaaRT	1969	F		46	0.26
mmaaBA	1969	M		57	0.14
mmbbLS	1967	M		26	0.04
sfamRA	1960	F	Nen village	63	0.28
sfbaWZ	1957	F		25	0.08
smbbTS	1953	M		46	0.13
smbbYZ	1951	M		75	0.07
sfbfHZ	1949	F	Nen village	14	0.43
smbbSZ	1948	M		47	0.06
Govav					
Speaker code	Year of birth		Residence	Tokens	
yfggMW	1999	F		35	0.97
ymggGW	1994	M		61	0.95
yfgmBZ	1989	F	Nen village	53	0.98
ymggSY2	1988	M		6	1
mmggRG	1985	M		58	0.95

mmggGZ	1984	M		33	0.76
mmggGS	1976	M		13	1
mfgbAN	1970	F		59	0.51
mfggMW	1968	F		45	0.91
mmggMN	1968	M		59	0.78
mfgmGN	1966	F	Nen village	57	0.56
mfgmGwN	1965	F	Nen village	46	0.65
smggZG	1959	M		45	0.18
sfgmFY	1955	F	Nen village	58	0.48
smggMG	1954	M		37	0.35
smggGM	1954	M		28	0
smggMY	1952	M		39	0.44
sfggDS	1952	F		48	0.15
39 speakers				1649	

## Abstract (Japanese)

本稿では南パプアニューギニアの言語変異調査の結果を発表する。族外婚によって維持され、近隣言語の多言語話者で構成されるケラケ部族の言語、ナンボ語自然音声コーパスをもとに、名詞語頭音 [h] の無音化を分析した。年齢によるバリエーションが最も重要な予測変数であり、若年層話者が高確率で [h] を無音化するという結果が見られた。夫方居住婚制で族外婚を果たしたケラケ女性のバリエーションも言語同体と同様である。 *Feature pool* (Cheshire, Kerswill, Fox, & Torgersen, 2011, Mufwene 2001) の概念に基づき [h] 無音化の原点は最も多言語性が高いケラケ集落と見られる。

## Address for correspondence

Eri Kashima  
University of Helsinki  
Unioninkatu 40  
Helsinki, 00100  
Finland  
eri.kashima@helsinki.fi