A Follow-up Analysis of Listener (Mis)comprehension across Language Varieties in Pentecost, Vanuatu

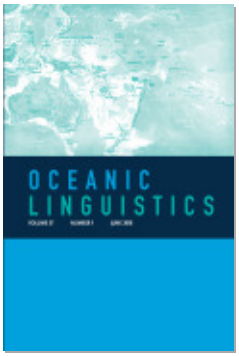
Schneider, Cindy; Gooskens, Charlotte

Published in:
Oceanic Linguistics

DOI:
10.1353/ol.2018.0005

Publication date:
2018

A Follow-up Analysis of Listener (Mis)comprehension across Language Varieties in Pentecost, Vanuatu



Cindy Schneider, Charlotte Gooskens

Oceanic Linguistics, Volume 57, Number 1, June 2018, pp. 144-176 (Article)

Published by University of Hawai'i Press
DOI: https://doi.org/10.1353/ol.2018.0005

A Follow-up Analysis of Listener (Mis)comprehension across Language Varieties in Pentecost, Vanuatu

Cindy Schneider and Charlotte Gooskens

UNIVERSITY OF NEW ENGLAND AND UNIVERSITY OF GRONINGEN

Intelligibility testing in Vanuatu in 2015 investigated how well speakers of three closely related varieties could understand each other. But the research also generated questions about (1) anomalous results; (2) which linguistic factors influenced comprehension; and (3) the relative intelligibility of varieties. This paper interrogates these questions and finds that, first, while most anomalous results are difficult to account for, others are easily explained; the insights gained will help to refine the design of future intelligibility tests. Second, some variables appear to be more important than others in terms of the degree to which they impede intelligibility. Third, test participants' higher comprehension of a relatively distantly related variety over a more closely related variety leads to explanations that draw on both linguistic and social factors. The insights gained in this study contribute to existing research on the same topic in European languages and establish a starting point for similar research on Pacific languages.

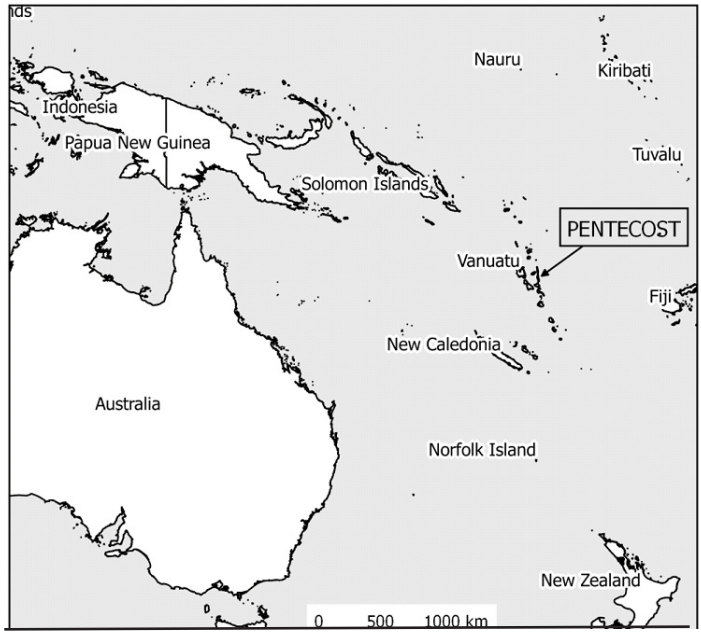
1. INTRODUCTION.[1] There are between 105 and 138 indigenous languages spoken in villages across the Pacific island nation of Vanuatu (François et al. 2015:5–6), and it is normal for speakers of different varieties to live in close proximity to each other. Yet the question of how well people can understand each other's varieties has received little attention in the Vanuatu context. Aside from intelligibility surveys in Efate (Stahl 1994) and north-central Santo (Stahl n.d.), we are aware of no other research in Vanuatu. In 2015, we took initial steps to fill this knowledge gap by devising an intelligibility test that can be done relatively quickly and easily in a largely undeveloped, oral, rural society like Vanuatu (Gooskens and Schneider 2016). In this paper, we examine the data to gain a better understanding of anomalous results: which linguistic factors influenced comprehension, and the relative intelligibility of varieties.

The specific location of our study was on the island of Pentecost in Vanuatu (see map 1). On Pentecost, many language varieties are packed into a relatively small area, which is typical for Melanesia. The island is only 62 kilometers from north to south, and 12 kilome-

1. The authors would like to thank members of the Raga and Apma communities in north and central Pentecost for their participation in our intelligibility test. We would also like to thank two anonymous reviewers for their invaluable comments, and Robert Early for technical assistance. Of course, we are responsible for any errors or omissions.

Oceanic Linguistics, Volume 57, no. 1 (June 2018)

MAP 1. VANUATU IN THE PACIFIC†



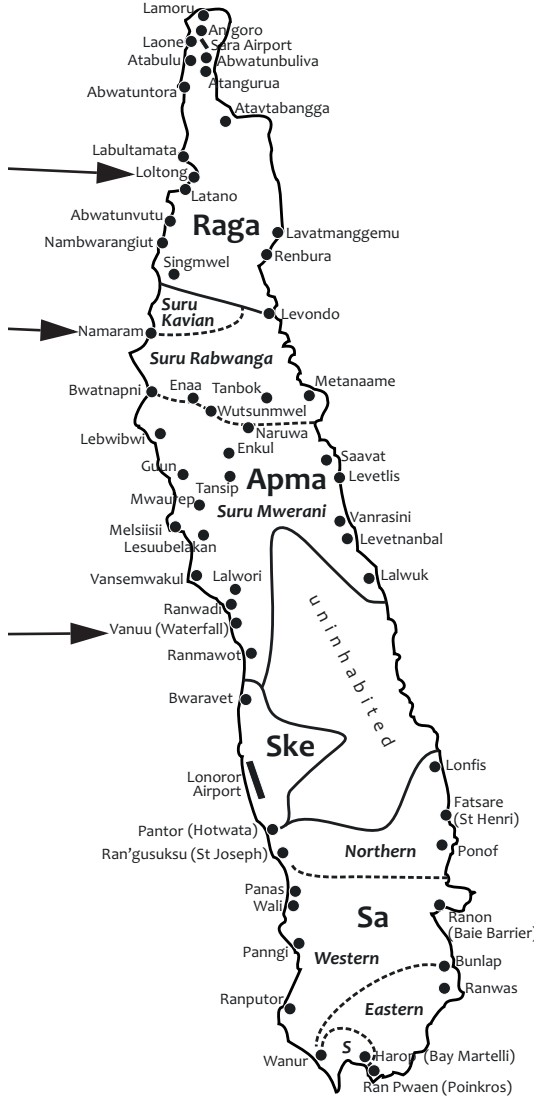
† Source: Jennifer Donlan, April 5, 2017.

ters across (Gray 2012:8), yet it is home to four language varieties and a number of smaller dialects. Going from north to south, these languages are Raga, Apma, Ske, and Sa. The present study focuses on Raga as well as two of Apma's three dialects, Suru Kavian and Suru Mwerani. Specifically, we will analyze how well Suru Mwerani speakers understood Raga and Suru Kavian. Map 2 shows all of the languages and dialects currently spoken on Pentecost. The three varieties discussed in this study are marked on map 2, and the villages where most of the testing was conducted are indicated there by arrows.

Raga (with about 6,500 speakers) is spoken in North Pentecost, and Apma (with about 7,800 speakers) is spoken in Central Pentecost (Lynch and Crowley 2001; the current population is likely higher due to natural increase since 2001). Although Raga is spoken across north Pentecost, it shows no evidence of regional variation (Marie-France Duhamel, pers. comm., August 28, 2017). Raga shares 60 percent cognacy with Apma, according to Gray (2012:14). This is based on his sample of 247 words (Andrew Gray, pers. comm., May 30, 2012). According to Tryon (1976:106), the two languages share 52 percent cognacy based on a sample of 242 words. Our own small testing sample of 80 words showed a cognacy rate of 54 percent between Raga and Apma.

Apma is purported to have three dialects: Suru Rabwanga, Suru Mwerani, and Suru Kavian. Suru Rabwanga and Suru Mwerani share 99 percent cognacy. Although the lexicon and pronunciation of these two varieties differ from each other in minor ways, speakers of these two varieties anecdotally report that they have no problems understanding one another. The anecdotal reports are supported by data that demonstrate that

MAP 2. PRESENT-DAY LANGUAGES OF PENTECOST ISLAND†



† Source: Andrew Gray, June 15, 2016. Purported language boundaries are drawn in solid lines; purported dialect boundaries in dotted lines. Smaller hamlets are not represented. The majority of testing was conducted in the Raga-speaking village of Loltong, the Suru Kavian-speaking village of Namaram, and the Suru Mwerani-speaking village of Waterfall, each marked by arrows on the map.

Suru Rabwanga speakers understand Suru Mwerani at levels above 90 percent (see Gooskens and Schneider 2016:295). Due to time constraints, we did not test how well

Suru Mwerani speakers understood Suru Rabwanga, but it is highly likely that the results would be similar. In contrast to the 99 percent cognacy shared between Suru Rabwanga and Suru Mwerani, Suru Kavian only has 90 percent cognacy with the other two dialects, according to Gray (2012:14). Results from our own word sample reveal a slightly lower cognacy rate of 87.5 percent. Suru Kavian is also highly endangered, with only 250 speakers. Adding to its endangerment is the strong perception in the wider Apma-speaking community that Suru Kavian is linguistically distinctive from the other two varieties (see Schneider 2017).

2. METHODOLOGY. During our 2015 field trip, we aimed to test a large number of participants of different ages from different places in north and central Pentecost. Knowing that it is time consuming to employ methods that have been used for intelligibility testing in oral societies, such as the RTT method (see Hickerson, Turner, and Hickerson 1952; Pierce 1952; Voegelin and Harris 1951; Casad 1974; Nahhas 2006), we decided to adapt tests that have recently been used for intelligibility testing in Europe (Vanhove 2014; Gooskens 2013; Gooskens and van Heuven 2017). In doing so, we had to take a number of circumstances into consideration. For example, we wanted participants of all ages to take the test, and we did not want to exclude those who could not read or write. We had a limited amount of time to collect our data, so we wanted short, efficient tests that could easily be carried out in the field and test a large number of participants in a short time.

We opted for word intelligibility tests rather than testing intelligibility of whole texts. An advantage of testing isolated words is that the influence from context on the understanding of a word can be excluded. This allows us to draw conclusions about the role of individual word characteristics for intelligibility. For example, we wanted to examine intelligibility across cognates to see whether the degree of similarity between the test words and the corresponding words in the native variety of the participant could predict intelligibility of individual words. By analyzing noncognates separately, we could draw conclusions about the role of exposure, since noncognates would only be understood by participants who have heard the words before. Furthermore, at sentence or higher levels, poor intelligibility is difficult to trace back to specific sources, and it may be difficult to know whether the test design would help or hinder intelligibility. On the one hand, context or situational redundancy may compensate for poor word intelligibility. But on the other hand, the added layer of morphosyntactic and discoursal differences, on top of lexical differences, may serve to impede intelligibility. That said, previous research has shown that, in general, morphosyntactic differences affect intelligibility to a lesser degree than lexical and phonological differences (Hilton, Gooskens, and Schüppert 2013).

We are also aware of the fact that a word test is ecologically less valid (that is, more constrained and less natural) than a test involving whole sentences or texts. However, a recent investigation (Gooskens and van Heuven 2017) comparing the results of three spoken intelligibility tests used to test mutual intelligibility between 16 different languages in Europe shows that the results of a word translation task correlated highly with the results of a cloze test set up to test the intelligibility of a text of 200 words ($r = .73$). This seems logical, since to understand a text a listener has to be able to understand individual words.

The data presented in this paper were gathered through a word translation task. Only adults (people aged 16 and over) did this task. We were fortunate that, except for some elderly women, all people we tested could speak Bislama, Vanuatu's lingua franca, and could, therefore, translate the test words into this intermediary language. We checked that this was indeed the case by testing subjects' knowledge of Bislama in a picture-pointing task beforehand, as well as in that part of the translation task where participants translated their own variety (and, therefore, would be expected to translate consistently from their own language into Bislama). Participants who recognized fewer than four of the five Bislama words in the initial picture-pointing task were excluded from the analysis of the word translation task.

The words in our test were taken from Gray's compilation of 247 common words (Andrew Gray, pers. comm., May 30, 2012). Since the test would become too long if we were to test all words in this list, we made a selection of 80 words: 40 nouns and 40 verbs.

A few transitive verbs could not occur in isolation; these were presented together with an object. For example, the verb *hit* required an object, so we used the generic noun *someone*. Along similar lines, the intransitive verb *sit* sounded more 'natural' when followed by another word (*sit down*).

The translation test came in eight different versions. The languages in versions 1 to 4 were presented in the mirrored order of the languages in versions 5 to 8. In this way, we made sure that the potential effect of fatigue was the same for all language varieties and all test words in our investigation. The participants listened to the recordings of 15 nouns and 15 verbs in each of three Pentecost varieties: Raga, Suru Kavian, and Suru Mwerani, but they never listened to the same word twice. Each word was followed by a pause of five seconds during which the participant gave a spoken translation of the word into Bislama. The first author, a Bislama speaker, listened to the test words together with the participant via headphones and noted down for each word whether the participant translated it correctly. If the word was translated incorrectly, the wrong translation was noted down. After each pause, a beep signaled the next word. For more information on test design, see Gooskens and Schneider (2016).

Because we are specifically interested in understanding the factors that affect intelligibility across cognates, the data examined in this paper include only the cognate words.[2] The appendix contains a list of both Raga–Suru Mwerani and Suru Kavian–Suru Mwerani cognate word pairs. Since Raga and Apma (of which Suru Mwerani is a dialect) share a lower cognacy rate than do Suru Kavian and Suru Mwerani (two dialects of the same language), the Raga–Suru Mwerani wordlist (42 word pairs) is shorter than the Suru Kavian–Suru Mwerani wordlist (60 word pairs).[3]

2. In Raga, three word tokens are followed by a complement word that is not cognate with its Suru Mwerani counterpart. The noncognate complement word was excluded from the calculations. In Suru Kavian, all complement words were cognate with their Suru Mwerani counterparts, so there was no need to exclude them from calculations.
3. Two Raga–Suru Mwerani and ten Suru Kavian–Suru Mwerani cognate word pairs are excluded from the present analysis because they were found to be "corrupt," in the sense that the test result did not accurately reflect comprehension. For example, when listeners were prompted with Suru Kavian [man] 'he/she laughs', they simply translated it into Bislama/English 'man'.

Although we administered the intelligibility test to speakers of all three varieties (Raga, Suru Kavian, and Suru Mwerani) in three different locations in Pentecost, this paper focuses specifically on how well Suru Mwerani speakers understood Raga and Suru Kavian cognates. Our 2016 study indicated that Suru Mwerani speakers correctly translated 19 percent of Raga noncognates, and 14.9 percent of Suru Kavian noncognates. This suggests that their exposure to these two varieties is low and that, when translating cognates, Suru Mwerani speakers could generally not depend on prior knowledge of Raga and Suru Kavian.[4] Successful translation would depend on a listener's ability to extrapolate information solely from the linguistic cues.

Thirty-two Suru Mwerani-speaking participants (12 females and 20 males) participated in the experiment. Their mean age was 37.5 years (with a range between 18 and 68 years) and their mean educational level was 7.8 years (with a range between 3 and 15 years).

3. LINGUISTIC DISTANCES. We measured linguistic distances between each Raga–Suru Mwerani and Suru Kavian–Suru Mwerani cognate word pair. We based linguistic distance calculations on broad phonetic transcriptions of the words in the intelligibility experiments. We chose a broad phonetic transcription over a phonemic one because the former offers a better characterization of how speakers actually pronounce the words. We also preferred a broad phonetic transcription over a narrow transcription because we wanted to be able to compare linguistic segments at a general level, without being distracted by excessive detail.

Phonetic distance is computed for the aligned cognate word pairs in both pairs of languages. The degree of dissimilarity between cognates is computed by the Levenshtein algorithm. It computes the smallest number of string edit operations needed to convert the phonetic string in language A to the string in B. Possible string operations are deletions, insertions, and substitutions of symbols. We illustrate this algorithm by comparing the Raga phrase [m^wa ^{m}bohai vatu] with the Suru Mwerani cognate [m^wabohni βet] 'he/she throws a stone' in table 1. The phrase [m^wa ^{m}bohai vatu] can be mapped to [m^wabohni βet] in many different ways, but the Levenshtein distance always gives the cost of the cheapest mapping. The minimum cost is based on an alignment in which a vowel matches with a vowel and a consonant matches with a consonant.

In the third slot, [^{m}b] is replaced by [b], in the sixth slot [n] is inserted, in the seventh slot [ai] is replaced by [i], in the eighth slot [v] is replaced by [β], in the ninth slot [a] is replaced by [e], and in the eleventh slot [u] is deleted. The total number of operations is then divided by the length of the alignment (number of alignment slots) to yield a length-normalized

TABLE 1. ILLUSTRATION OF THE LEVENSHTEIN ALGORITHM

Alignments	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
Raga	m^w	a	^{m}b	o	h		ai	v	a	t	u
Suru Mwerani	m^w	a	b	o	h	n	i	β	e	t	
No. of operations			1			1	1	1	1		1

4. Of course, any exposure to Raga and Suru Kavian by Suru Mwerani speakers can lead to results that suggest levels of apparent comprehension of cognate words that are higher than they would be if there was no exposure at all. For more discussion on the role of exposure, see Gooskens and Swarte (2017).

Levenshtein distance. As there are six operations and the alignment has eleven slots, the distance is calculated as (6/11) x 100 = 55%. The measure is symmetrical between word pairs (for more explanation and background, see Nerbonne and Heeringa 2010).

4. FACTORS THAT AFFECT INTELLIGIBILITY. This section introduces variables that have been considered in the literature for their potential to affect intelligibility across cognates in closely related languages. Each of these is measured in our own analysis. We used a binary coding system, either '0' or '1'. This method follows Kürschner, Gooskens, and van Bezooijen (2008:88–90) to an extent, although they use more refined measurements. The methodology for cognate comparison, explained in section 3 above, is also applied here. That is, the "like" parts of each word are mapped against each other for the purpose of comparison.

4.1 WORD STRESS DIFFERENCES ("Stress"). According to Harrington and Cox (2009), listeners are strongly attuned to word stress. Table 2 gives an example of a difference in word stress. After the words in the pair are mapped on to each other, it is evident that stress occurs in different places in the two languages.

TABLE 2. DIFFERENCES IN WORD STRESS

Raga	[bu.tu.ˈbu.tu]	bu	tu	ˈ**bu**	tu	'ant'
Suru Mwerani	['but.but]	**'but**		but		

Weisser (2005) suggests that word stress can be identified, in order of importance, by pitch movement; length of the vowel at the core of the syllable; and intensity (loudness). He observes that it is often the combination of these three features that creates the impression of stress, rather than any single feature in isolation.

Careful listening by the authors was facilitated by Praat software in the analysis and identification of word stress in this study. Only primary stress was marked and compared across cognates. If the cognate pair contained a phrase (a verb followed by a noun, or a possessed noun followed by a possessor noun), then primary stress for each word in the cognate phrases was indicated and compared. Regardless of whether one or both of the two words had word stress differences, the difference was simply coded as '1' for the purpose of simple comparison across variables.

4.2 SYLLABLE DIFFERENCES ("SyllDiff"). Kürschner, Gooskens, and van Bezooijen (2008:93) found that a difference in the number of syllables between Swedish and Danish cognate pairs correlated negatively with intelligibility scores. They scored for the number of syllable differences, but we simply noted whether or not there was a difference in the number of syllables with '0' or '1', respectively. An example of a syllable difference between Suru Kavian and Suru Mwerani is given in table 3.

TABLE 3. SYLLABLE DIFFERENCE

Suru Kavian	['sol.sol]		sol	sol	'he/she/sews'
Suru Mwerani	[mʷo.'sol.sol]	mʷo	sol	sol	

4.3 VOWEL QUALITY DIFFERENCES ACROSS STRESSED SYLLABLES ("VQualStrSyll"). Van Heuven (2008:53) and Gooskens, Heeringa, and Beijering (2008:73) both find that consonants are more important to intelligibility than are vowels. In this study, we have selected a number of individual consonant sounds to investigate their impact upon intelligibility. As for vowel sounds, we decided to use a more general measure, focusing on stressed syllables. This measure tests whether intelligibility is affected by differing vowel qualities in stressed syllables across cognate word pairs. In table 4, the stressed syllable has a different vowel quality (Raga [a] versus Suru Mwerani [e]) across the cognate pair.

TABLE 4. DIFFERENCE IN VOWEL QUALITY IN THE STRESSED SYLLABLE

Raga	[ɣa.'ma.li]	ɣa	'm**a**	li	'meeting house'
Suru Mwerani	[ka.'mel]	ka	'm**e**l		

If there is a difference in word stress across a cognate pair, it is impossible to then compare vowel quality of the stressed syllables because primary stress is placed on different cognate syllables. These pairs could, therefore, not be compared using this measure, and were excluded from the comparison.

Length is not a component of vowel quality and was not considered to be a difference.

4.4 PRESENCE OF "FOREIGN" SOUNDS. "Foreign" sounds refer to sounds in the target variety that do not exist in the test taker's native variety (Kürschner, Gooskens, and van Bezooijen [2008:88]; also see van Heuven [2008:46]).

In this study, the test takers were native speakers of Suru Mwerani. There were three "foreign" sounds that Suru Mwerani participants heard in the recordings: one was the prenasalization that occurs in both Raga and Suru Kavian. The others were the Raga fricatives [ɣ] and [v]. These are discussed below.

4.4.1 Prenasalization ("PreNasal"). Although Suru Mwerani does not itself have prenasalization, the closely related Suru Rabwanga dialect does have prenasalized voiced stops. Many Suru Mwerani speakers would, therefore, be familiar with them. The prenasalized voiced stops in Raga and Suru Kavian are [ᵑg], [ⁿd], and [ᵐb]. Table 5 gives an example of a Suru Kavian-Suru Mwerani cognate pair with a difference in prenasalization.

4.4.2 Velar fricative ("VelarFric"). Raga has the velar fricative /ɣ/ (Vari-Bogiri 2011: 25). An example of its sound correspondence with the Suru Mwerani velar stop /k/ is given in table 6.

TABLE 5. DIFFERENCE IN PRENASALIZATION

Suru Kavian	[te. '**ⁿd**ap]	'it is white'
Suru Mwerani	[te. '**d**ap]	

TABLE 6. DIFFERENCE IN VELAR FRICATIVE

Raga	[**ɣ**a.'vi.**ɣ**a]	'Malay apple'
Suru Mwerani	[**k**a.'βi**k**]	

4.4.3 Labiodental fricative ("LabFric"). Raga also has the labiodental fricative /v/. Gray (2012:51) notes that Raga /v/ can be devoiced to [f]. Neither [v] nor [f] exists in Suru Mwerani; the closest Suru Mwerani sound is the voiced bilabial fricative /β/. Since a broad phonetic transcription was employed, for the sake of simplicity all labiodental fricatives in the Raga data are simply transcribed as [v], although some of them arguably sound more like [f]. An example of the correspondence is shown in table 7.

TABLE 7. DIFFERENCE IN LABIODENTAL FRICATIVE

Raga	[va. 'nu.a]	'village'
Suru Mwerani	['**β**i.ni]	

4.5 [s]/[t]/[d] ("s/t/d"). Unlike "foreign" sounds, the Suru Kavian sounds [s]/[t]/[d] also exist in Suru Mwerani. But there is noticeable variation in the way these sounds occur across the two varieties.

There is a complicated historical relationship between [ts] in Suru Mwerani on the one hand, and [s] and [t] in Suru Kavian on the other. There is also a diachronic explanation for the correspondence between Suru Mwerani [d] and Suru Kavian [s]. See Schneider and Gray (2015:207–8) for details.

The resulting variation is potentially more confusing for Suru Mwerani listeners than other types of consonantal variation, due to the fact that familiar sounds occurring in unfamiliar positions in the target word create a sort of "dissonance" for the listener.

Either type of variation (Suru Mwerani [ts] : Suru Kavian [s]/[t]; or Suru Mwerani [d] : Suru Kavian [s]) was coded with '1'. Moreover, if there was more than one token of variation in a word pair, we still coded the difference as '1' for that word pair. An example is given in table 8.

TABLE 8. DIFFERENCES IN s/t/d SOUNDS

Suru Kavian	['si. bos.'wos]		**s**i	bos	wos	'he/she stands up'
Suru Mwerani	['m^{w}i.di. bos.'wos]	m^{w}i	**d**i	bos	wos	

4.6 TOTAL NUMBER OF DIFFERENT FEATURES PER WORD ("#DiffFtrs"). The presence or absence of the above-listed variables for each cognate word pair (Raga–Suru Mwerani or Suru Kavian–Suru Mwerani) is coded with '0' or '1' in the spreadsheet in the appendix. The values for the variables are tallied and their sum is placed in the final column, headed "#DiffFtrs." This measure, which indicates the total number of distinguishing features for each pair of words, can be useful when analyzing the impact of any given variable on intelligibility. If, for example, the number of different features per word pair is low, say '1', then the variable creating that difference may be deemed to have more of an effect on Suru Mwerani comprehension than if, say, there are several differences in features across a word pair.

A limitation of this method is that it only measures negative comprehension (the degree to which a variable affects *non*-comprehension). It has nothing to say about the effect of a variable on a listener's ability to successfully comprehend a word.

4.7 LOCATION OF VARIATION WITHIN THE WORD. Another important factor that can influence listener (mis)comprehension is not a single linguistic feature (like the factors discussed in sections 4.1 to 4.5). Rather it relates to *where* variation occurs within the target word. Van Heuven (2008:51) discusses the "superiority of the word beginning," saying that it is preferable for a listener to reduce the possible number of candidate words as quickly as possible for the sake of processing efficiency. If a feature difference (between the listener's variety and the target variety) occurs at the beginning of the target word, listener comprehension is immediately impaired. On the other hand, if there is no variation until the end of the word, lexical activation is more likely to have already occurred, and there is a greater likelihood of successful comprehension. See Cutler (2012, ch. 3) for more discussion on the greater importance of initial segments in spoken word recognition.

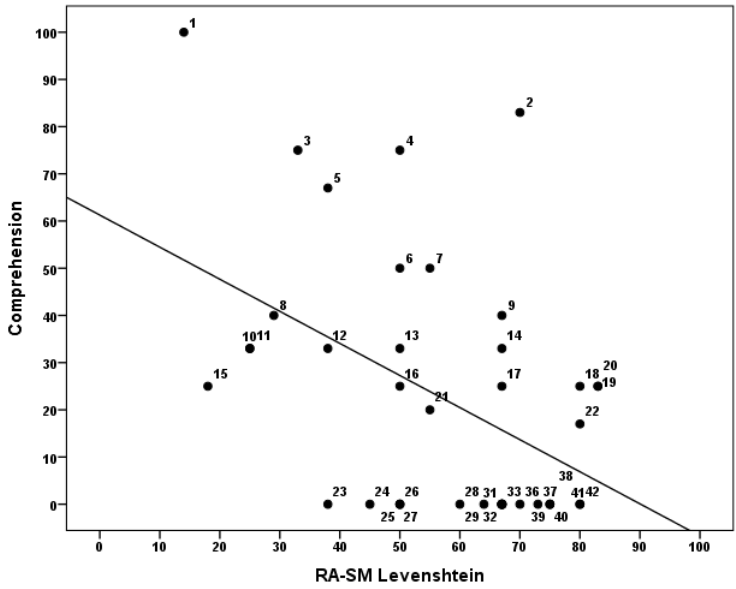
5. RESULTS, ANALYSIS, AND DISCUSSION. If a word was correctly translated, then it was marked as correct and assumed to have been understood by the Suru Mwerani participant. If it was translated incorrectly, or if no response was given, then it was marked as incorrect. The results data in the appendix are split into three categories: High, Medium, and Low (H-M-L). A correct response (comprehension) rate of 67 percent or higher was considered "High." A comprehension rate of 33 percent or less was considered "Low." Everything else fell into the middle group. We wanted the High-Medium-Low groups to be equally represented, so we divided the comprehension categories into equal thirds. Having three major groupings maximized the possibility that a sufficient number of words would fall into each group for meaningful averages to be calculated, and for generalizations to be made. A gradient measurement would provide a more detailed view of the results, but it would be more difficult to make generalizations from the data.

Four main observations can be made. First, there is a rough inverse correlation between comprehension and Levenshtein distance, which is to be expected. However, there are interesting individual exceptions to this pattern: (1) high intelligibility/high distance; and (2) low intelligibility/low distance (5.1). Second, some variables seem to impede intelligibility more than others do (5.2). Third, listeners seem better able to process linguistic differences and understand the target word when variation does not occur in the first syllable (5.3). And finally, Suru Mwerani (SM) speakers have a higher tolerance of linguistic distance with Raga (RA) than with Suru Kavian (SK), as outlined in 5.4.

5.1 DISCREPANCIES BETWEEN LINGUISTIC DISTANCE AND COMPREHENSION. Kürschner, Gooskens, and van Bezooijen (2008:86) state: "at the word level small phonetic distances can be assumed to correlate with high intelligibility scores, while large distances can be expected to correlate with low intelligibility scores." Our data roughly follow that pattern; see figures 1 and 2.

The reference lines in figures 1 and 2 are drawn in such a way that as many points as possible lie as close as possible to the line. The points above the line show cases where intelligibility is higher than could be expected from Levenshtein distances, and the points below the line show cases where intelligibility is lower than could be expected. If there were a perfect relationship between comprehension and distance (if distance could pre-

FIGURE 1. RAGA–SURU MWERANI COMPREHENSION AND LEVENSHTEIN DISTANCE†



† Correlation: r = -0.47. Individual points represent word pairs 1 to 42 listed in the appendix.

dict comprehension 100 percent of the time, or vice versa), all points would be on the line and there would be a correlation of 1.0 (or –1.0). In the present case, there is a weak inverse correlation ($r = –0.47$) between Raga–Suru Mwerani comprehension and Levenshtein distance. The Suru Kavian–Suru Mwerani correlation is stronger ($r = 0.60$). Both correlations are significant at the .01 level. However, there are individual exceptions to this general pattern, which are reviewed below.

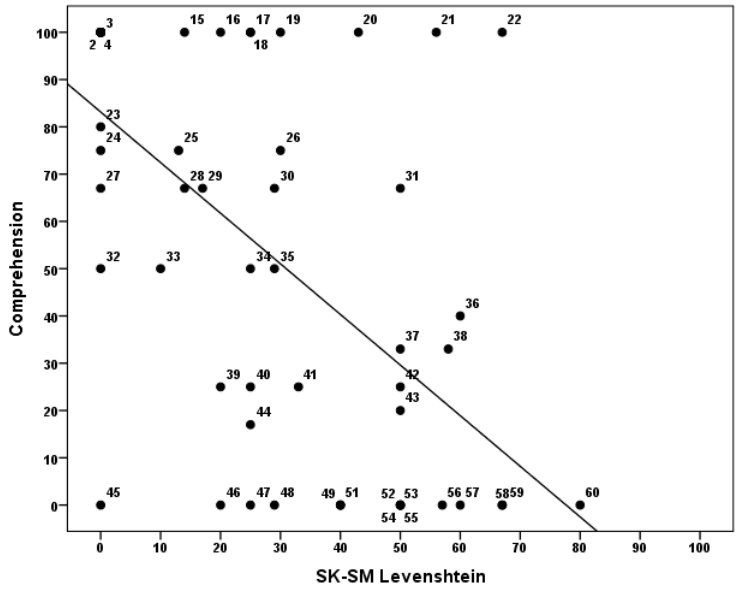
5.1.1 Suru Mwerani comprehension of Raga: High comprehension but high Levenshtein distance. Table 9 shows that only one cognate pair has high comprehension but also high Levenshtein distance. Words falling into this category had a comprehension rate at 67 percent or higher, and a corresponding distance of 67 percent or higher.

The anomaly in table 9 can be explained by the fact that the base form [min] used in Raga is close to the Suru Rabwanga form [mini sileŋ]. Due to high interintelligibility between Suru Rabwanga and Suru Mwerani, many Suru Mwerani speakers would

TABLE 9. SURU MWERANI COMPREHENSION OF RAGA: HIGH COMPREHENSION BUT HIGH LEVENSHTEIN DISTANCE

#	RA	SM	Gloss	Comprehension	Levenshtein Distance	Differences
2	[mʷa. 'min. 'wai]	['mʷam.ni si.'leŋ]	'he/she drinks water'	83%	70%	Stress; SyllDiff

FIGURE 2. SURU KAVIAN–SURU MWERANI COMPREHENSION AND LEVENSHTEIN DISTANCE†



† Correlation: r = -0.60. Individual points represent word pairs 1 to 60 listed in the appendix.

already be familiar with this form. Therefore, Suru Mwerani participants likely recognized Raga [mʷa 'min] from exposure to Suru Rabwanga. They would not, however, have been familiar with Raga [wai] 'water', which is not cognate with Suru Mwerani/Suru Rabwanga [si'leŋ]. Perhaps this explains why the comprehension rate was only 83 percent and not 100 percent.

5.1.2 Suru Mwerani comprehension of Raga: Low comprehension but low Levenshtein distance. Table 10 lists the three Raga words that Suru Mwerani speakers did not understand well, but which also had low Levenshtein distances across the cognate pairs.

Although [10] in table 10 has a stress difference across the cognate pair, what [10], [11], and [15] all share is a difference in the number of syllables. This suggests that syllable differences could be significant in affecting intelligibility. Yet even Raga words that

TABLE 10. SURU MWERANI COMPREHENSION OF RAGA: LOW COMPREHENSION BUT LOW LEVENSHTEIN DISTANCE

#	RA	SM	Gloss	Comprehension	Levenshtein Distance	Differences
10	[bu.tu.'bu.tu]	['but.but]	'ant'	33%	25%	Stress; SyllDiff
11	['bʷe.ta]	[bʷet]	'taro'	33%	25%	SyllDiff
15	[bʷan.se.re.'se.re]	[bʷan.se:.'se:]	'centipede'	25%	18%	SyllDiff

were well comprehended had syllable differences with their Suru Mwerani counterparts (see Raga–Suru Mwerani word pairs [1]–[5] in the appendix). With its CVCV structure, Raga words simply tend to have more syllables, while their Suru Mwerani counterparts have a collapsed syllable structure due to vowel loss. There is, therefore, no clear reason why these particular words were poorly understood.

5.1.3 Suru Mwerani comprehension of Suru Kavian: High comprehension but high Levenshtein distance. The high comprehension of Suru Kavian shown in table 11 has the same explanation as given in section 5.1.1 for Raga.

TABLE 11. SURU MWERANI COMPREHENSION OF SURU KAVIAN: HIGH COMPREHENSION BUT HIGH LEVENSHTEIN DISTANCE

#	SK	SM	Gloss	Comprehension	Levenshtein Distance	Differences
22	['min. si.'leŋ]	['mʷam.ni. si.'leŋ]	'he/she drinks water'	100%	67%	Stress; SyllDiff

5.1.4 Suru Mwerani comprehension of Suru Kavian: Low comprehension but low Levenshtein distance. The data in table 12 present a mixed bag of explanations. The variables at play here are: vowel quality differences ([41], [44]); different number of syllables ([40]); prenasalization ([47]); and s/t/d ([48]). There are also differences that the variables do not take account of, for example, a difference in vowel length plus the addition of a final consonant ([46]). There is no obvious pattern behind listener confusion if one considers the linguistic features alone.

TABLE 12. SURU MWERANI COMPREHENSION OF SURU KAVIAN: LOW COMPREHENSION BUT LOW LEVENSHTEIN DISTANCE

#	SK	SM	Gloss	Comprehension	Levenshtein Distance	Differences
39	[me.'sin]	[mi.'sin]	'his/her urine'	25%	20%	None
40	['gil.kil]	[mʷi.'gil.kil]	'he/she digs'	25%	25%	SyllDiff
41	[laŋ]	[leŋ]	'fly'	25%	33%	VQualStrSyll
44	['mʷi.ah]	['mʷai.ah]	'he/she falls down'	17%	25%	VQualStrSyll
45	[nap]	[nap]	'wave'	0%	0%	None
46	[te. 'met]	[te. 'me:]	'it is black'	0%	20%	None
47	['me.ᵐbe]	['me.be]	'he/she rests'	0%	25%	PreNasal
48	['da.tu.'wan]	['ta.tsu.'wan]	'his/her sweat'	0%	29%	s/t/d

Of particular interest is [45], which takes the identical form in Suru Kavian and Suru Mwerani, yet it has zero percent comprehension. Looking at the data in the appendix more generally, there are 19 Suru Kavian–Suru Mwerani cognate pairs that have a Levenshtein distance of zero (and are, therefore, identical across both varieties). Of these, 14 have a comprehension of 100 percent, which is what we would expect. Of the five remaining pairs having less than 100 percent comprehension ([23], [24], [27], [32], [45]), four (80 percent) are monosyllables. On the other hand, of the 14 pairs that have 100 percent comprehension, only two ([4], [7]), or 14 percent, are monosyllables. This suggests

that Suru Mwerani speakers generally have a more difficult time understanding monosyllables. This may be due to the simple fact that single syllables are shorter in duration than bisyllabic and multisyllabic words, and so listeners have less time to process them under test conditions. Of course, this is not an issue when the monosyllable occurs within the wider context of natural speech. Single-syllable words also have more phonetic neighbors than longer words, and, therefore, it is easier for them to be confused with other words (van Heuven [2008:52]).

5.2 SIGNIFICANCE OF INDIVIDUAL VARIABLES. We now examine in more detail the individual variables that can affect intelligibility. Measures reflect the average presence of a given variable in each of the High-Medium-Low groups listed in the appendix. Table 13 shows the average presence of individual features for each of the High-Medium-Low groups for Suru Mwerani comprehension of Raga and Suru Kavian. One indicator that any given linguistic feature impacts upon intelligibility is that its value in the High group is markedly lower relative to the Medium-Low groups. In other words, comprehension is higher when the variable occurs less frequently, and lower when it occurs more frequently.

TABLE 13. SURU MWERANI COMPREHENSION AND FEATURE DIFFERENCES WITH RAGA AND SURU KAVIAN

	RA			SK		
	H	**M**	**L**	**H**	**M**	**L**
Average Comprehension (%)	80	45	11	92	48	8
Average Number of Different Features per word	1.4	2.5	2.2	0.42	1.4	1.58
Stress: Average Difference	0.4	0.5	0.49	0.06	0.2	0.25
SyllDiff: Average Difference	1	1	0.85	0.16	0.4	0.42
VQualStrSyll: Average Difference	0	0	0.33	0.07	0.5	0.5
PreNasal: Average Difference	0	0.25	0.18	0.06	0	0.13
VelarFric: Average Difference	0	0.25	0.33			
LabFric: Average Differences	0	0.5	0.21			
s/t/ts/d: Average Difference				0.07	0.4	0.42

There is a general inverse correlation between Suru Mwerani comprehension of Suru Kavian in the High-Medium-Low groups and the corresponding average number of different features per word. For Raga, the pattern is less clear; the group of words that Suru Mwerani speakers comprehended the least had a lower average number of different features per word (2.2) than the Medium group (2.5).

At this point we consider the impact of each of the individual linguistic variables in more detail.

5.2.1 Differences in word stress ("Stress")

5.2.1.1 Raga. In total, 20 Raga words had different stress patterns from Suru Mwerani words. There does not appear to be any pattern, with mean presence of this variable per word relatively stable across each of the High, Medium, and Low groupings at 0.4, 0.5, and 0.49, respectively. Table 14 shows the two pairs in the high group with word stress

differences, while table 15 lists all pairs with word stress differences in the Medium and Low groups.

Only three of eighteen word pairs (17 percent) with medium or low comprehension had stress as the sole distinguishing feature. This fact, combined with the stability of this feature across all three groups, provides evidence that word stress difference, on its own, does not impede Suru Mwerani comprehension of Raga words.

TABLE 14. SURU MWERANI–RAGA HIGH GROUP: DIFFERENCES IN WORD STRESS

#	RA	SM	Gloss	Comprehension	Other Factors
1	[tel.'tel.e]	['tel.tel]	'snake'	100%	SyllDiff
2	[mʷa. 'min. 'wai]	['mʷam.ni. si.'leŋ]	'he/she drinks water'	83%	SyllDiff

TABLE 15. SURU MWERANI–RAGA MEDIUM /LOW GROUPS: DIFFERENCES IN WORD STRESS

#	RA	SM	Gloss	Comprehension	Other Factors
7	[mʷa. 'ᵑgi.ta. a.'ta.tu]	['mʷi.gi.ta. 'a.tsi]	'he/she sees someone'	50%	SyllDiff; PreNasal
8	[mʷa. ɾo.'ŋo.e]	[mʷo.'ɾo.ŋo]	'he/she hears'	40%	SyllDiff
10	[bu.tu.ˈbu.tu]	['but.but]	'ant'	33%	SyllDiff
17	[mʷa. 'lu.a]	[mʷi.li.'aut]	'he/she vomits'	25%	None
18	['me.ɾe]	[mi.'sin]	'his/her urine'	25%	None
20	['u.hi]	['bʷa.ɾus]	'pawpaw'	25%	None
21	[mʷa. 'we.hi. a.'ta.tu]	['mʷa.hi. 'a.tsi]	'he/she hits someone'	20%	SyllDiff
22	[va.'nu.a]	['βi.ni]	'village'	17%	SyllDiff; LabFric
23	[mʷa. 'lai. 'wa.ᵑga]	['mʷa.li. 'a.ga]	'he/she takes a ship'	0%	PreNasal
24	[mʷa. ᵐbo.'hai. 'va.tu]	['mʷa.boh.ni. 'βet]	'he/she throws a stone'	0%	SyllDiff; PreNasal; LabFric
25	['i.do.lin. 'ma.nu]	['du.lun. bʷi.'hil]	'(chicken) egg'	0%	SyllDiff
28	[ba.'ɣe.o]	['be.ke]	'shark'	0%	SyllDiff; VelarFric
29	[mʷa. 'ɾav. 'ɣao]	['mʷe.ɾa.βa. 'ka:.wa]	'he/she pulls a rope'	0%	SyllDiff; Velar-Fric; LabFric
30	[mʷa. ᵑgel.'ge.li]	[mʷi.'gil.kil]	'he/she digs'	0%	SyllDiff; PreNasal
32	[ba.'ha.ɾa]	[bas]	'cliff'	0%	SyllDiff
35	[mʷa. so.'ɣai. a.'ta.tu]	['mʷe.da.ba.so. ni. 'a.tsi]	'he/she pushes someone'	0%	SyllDiff; VelarFric
37	[nu. ma.nev.'nev.i]	[tem.'nip.nip]	'it is thin'	0%	SyllDiff; LabFric
39	[mʷa. ⁿdo.ɣo. 'ɾa.du]	['mʷe.sa.dok. 'ŋa.mʷa]	'he/she is still sitting down'	0%	PreNasal; VelarFric

5.2.1.2 Suru Kavian. The high comprehension group has only two words with word stress differences, shown in table 16. The mean average presence of words in the High group with differences in word stress is 0.06.

The stress pattern of Suru Kavian ['si bos'wos] 'he/she stands up' differs from Suru Mwerani ['mʷidi bos'wos] in the first word but not the second. The forms ['si] and ['mʷidi] 'stand' are cognate. The existence of [bos'wos] 'straight' as an identical form across both

TABLE 16. SURU MWERANI–SURU KAVIAN HIGH GROUP: DIFFERENCES IN WORD STRESS

#	SK	SM	Gloss	Comprehension	Other Factors
19	['si. bos.'wos]	['mʷi.di. bos.'wos]	'he/she stands up'	100%	SyllDiff; s/t/d
22	['min. si.'leŋ]	['mʷam.ni. si.'leŋ]	'he/she drinks water'	100%	SyllDiff

varieties would have assisted Suru Mwerani speakers in their comprehension, and they correctly translated this phrase 100 percent of the time. Pair [22], also translated correctly 100 percent of the time, was discussed in 5.1.3.

The occurrence of pairs having differences in word stress more than trebles from 0.06 in the High group to 0.2 and 0.25 in the Medium and Low groups, respectively. However, as shown in table 17, word stress is not the sole contributor to unintelligibility; it is one of two or three linguistic features that distinguish Suru Mwerani and Suru Kavian. Stress does appear to have some impact on intelligibility, but mitigating this impact is the fact that it occurs in combination with other distinguishing linguistic features in the Medium-Low groups.

TABLE 17. SURU MWERANI–SURU KAVIAN MEDIUM/LOW GROUPS: DIFFERENCES IN WORD STRESS

#	SK	SM	Gloss	Comprehension	Other Factors
35	['boh.ni. 'βas]	['mʷa.boh.ni. 'βet]	'he/she throws a stone'	50%	SyllDiff; s/t/d
37	['gi.da. 'a.si]	['mʷi.gi.ta. 'a.tsi]	'he/she sees someone'	33%	SyllDiff; s/t/d
38	['u.ⁿde.lin. bʷe:.'il]	['du.lun. bʷi.'hil]	'(chicken) egg'	33%	SyllDiff; PreNasal
54	['buh. 'ka:.wa]	['mʷa.bu.hu. 'ka:.wa]	'he/she holds a rope'	0%	SyllDiff
56	['da.be.sok. 'a.si]	['mʷe.da.ba.so.ni. 'a.tsi]	'he/she pushes someone'	0%	SyllDiff; s/t/d
57	['dok. 'ŋa.mʷa]	['mʷe.sa.dok. 'ŋa.mʷa]	'he/she is still sitting down'	0%	SyllDiff
60	['lap. 'a.ᵑga]	['mʷa.li. 'a.ga]	'he/she takes a ship'	0%	SyllDiff; PreNasal

In cases where Suru Mwerani comprehension of Suru Kavian is high, cognate words usually share the same stress pattern. Where comprehension is lower, then other features besides just stress distinguish the Suru Kavian and Suru Mwerani cognates.

5.2.2 Different number of syllables ("SyllDiff")

5.2.2.1 Raga. Across all three categories (High-Medium-Low), there is almost always a difference in the number of syllables between Suru Mwerani and Raga. Thus, a comparison across High-Medium-Low groups is meaningless. Every word pair in the high group has a different number of syllables, as shown in table 18.

In the Medium and Low groups, syllable differences occurred at the level of 1 and 0.85, respectively. These word pairs are listed in table 19.

TABLE 18. SURU MWERANI–RAGA HIGH GROUP: DIFFERENCES IN WORD STRESS

#	RA	SM	Gloss	Comprehension	Other Factors
1	[tel.'tel.e]	['tel.tel]	'snake'	100%	Stress
2	[mʷa. 'min. 'wai]	['mʷam.ni. si.'leŋ]	'he/she drinks water'	83%	Stress
3	['bo.e]	[bo]	'pig'	75%	None
4	[mʷa. 'ma.na]	[man]	'he/she laughs'	75%	None
5	[mʷa. ma.'tu.ɾu]	[mʷam.'tsuː]	'he/she sleeps'	67%	None

TABLE 19. SURU MWERANI–RAGA MEDIUM /LOW GROUPS: DIFFERENCES IN WORD STRESS

#	RA	SM	Gloss	Comprehension	Other Factors
6	[mʷa. 'ɾo.vo]	[mʷo.'rop]	'he/she runs'	50%	LabFric
7	[mʷa. ⁿgi.ta. a.'ta.tu]	['mʷi.gi.ta. 'a.tsi]	'he/she sees s.o.'	50%	Stress; PreNasal
8	[mʷa. ɾo.'ŋo.e]	[mʷo.'ɾo.ŋo]	'he/she hears'	40%	Stress
9	[ɣa.'vi.ɣa]	[ka.'βik]	'Malay apple'	40%	VelarFric; LabFric
10	[bu.tu.ˈbu.tu]	['but.but]	'ant'	33%	Stress
11	['bʷe.ta]	[bʷet]	'taro'	33%	None
12	[mʷa. ma.'ta.ɣu]	[mʷam.'ta.tsi]	'he/she is afraid'	33%	VelarFric
13	[ɣa.'ma.li]	[ka.'mel]	'meeting house'	33%	VQualStrSyll; VelarFric
14	[nu. 'no.ɣo]	[te. 'nok]	'it is finished'	33%	VelarFric
15	[bʷan.se.ɾe.'se.ɾe]	[bʷan.seː.'seː]	'centipede'	25%	None
16	[bʷa.'ɾa.tu]	[bʷe.'ɾet]	'flying fox'	25%	VQualStrSyll
19	[mʷa. 'ma.ᵐbu]	['me.be]	'he/she rests'	25%	VQualStrSyll; PreNasal
21	[mʷa. 'we.hi. a.'ta.tu]	['mʷa.hi. 'a.tsi]	'he/she hits s.o.'	20%	Stress
22	[va.'nu.a]	['βi.ni]	'village'	17%	Stress; LabFric
24	[mʷa. ᵐbo.'hai. 'va.tu]	['mʷa.boh.ni. 'βet]	'he/she throws a stone'	0%	Stress; PreNasal; LabFric
25	['i.do.lin. 'ma.nu]	['du.lun. bʷi.'hil]	'(chicken) egg'	0%	Stress
26	['la.ŋo]	[leŋ]	'fly'	0%	VQualStrSyll
27	['na.vo]	[nap]	'wave'	0%	LabFric
28	[ba.'ɣe.o]	['be.ke]	'shark'	0%	Stress; VelarFric
29	[mʷa. 'ɾav. 'ɣao]	['mʷe.ɾa.βa. 'kaː.wa]	'he/she pulls a rope'	0%	Stress; VelarFric; LabFric
30	[mʷa. ⁿgel.'ge.li]	[mʷi.'gil.kil]	'he/she digs'	0%	Stress; PreNasal
31	[mʷa. ⁿga.ɣa]	[mʷe.'gak]	'he/she flies'	0%	PreNasal; VelarFric
32	[ba.'ha.ɾa]	[bas]	'cliff'	0%	Stress
33	[nu. 'me.to]	[te. 'meː]	'it is black'	0%	None
34	[nu 'meho]	[te 'mes]	'it is wet'	0%	None
35	[mʷa. so.'ɣai. a.'ta.tu]	['mʷe.da.ba.so.ni. 'a.tsi]	'he/she pushes someone'	0%	Stress; VelarFric
36	[nu. ma.ma.'ɣa.ni]	[tem.'kan]	'it is sharp'	0%	VelarFric
37	[nu. ma.nev.'ne.vi]	[tem.'nip.nip]	'it is thin'	0%	Stress; LabFric
38	[ɣa.ti.'ɣu.na]	[tsu.'kun]	'his/her back'	0%	VelarFric
40	['va.tu]	[βet]	'stone'	0%	VQualStrSyll; LabFric
41	[ɣao]	['kaː.wa]	'rope'	0%	VQualStrSyll; VelarFric
42	[ta.'i.va]	['tsi.βi]	'conch shell'	0%	LabFric

In addition to the high rate of occurrence of syllable differences, only 10.8 percent (4/37 pairs) with medium or low comprehension had "SyllDiff" as the sole distinguishing feature. This makes it difficult to attribute lack of comprehension just to syllable difference.

5.2.2.2 Suru Kavian. Table 20 shows word pairs in the High group with differences in the number of syllables. All five Suru Mwerani words in the High group have third person singular imperfective ([mʷa]) marking. In Suru Kavian, the 3SG imperfective is (with some exceptions) zero-marked, but in Suru Mwerani (also with some exceptions) it is cliticized to the verb stem. One reason that Suru Mwerani speakers are able to recognize these Suru Kavian forms is because the word stress for [18], [21], and [30] remains consistent across the pairs. The high comprehension of [19] and [22] has been explained previously. While the average presence of syllable differences is 0.16 for the High group, this more than doubles to 0.40 and 0.42 for the Medium and Low groups, respectively.

Table 21 lists words in the Medium-Low groups that had syllable differences between the word pairs. More examples of differences created by imperfective marking are shown by [35], [37], [40], [54], [55], [56], [57], and [60]. These pairs also had other linguistic differences that further impeded intelligibility. The sole exception to this is [40].

TABLE 20. SURU KAVIAN–SURU MWERANI HIGH GROUP: DIFFERENCES IN NUMBER OF SYLLABLES

#	SK	SM	Gloss	Comprehension	Other Factors
18	['sol.sol]	[mʷo.'sol.sol]	'he/she sews'	100%	None
19	['si. bos.'wos]	['mʷi.di. bos.'wos]	'he/she stands up'	100%	Stress; s/t/d
21	['sa.sah]	[mʷe.'sa:.sa:]	'he/she sings'	100%	None
22	['min. si.'leŋ]	['mʷam.ni. si.'leŋ]	'he/she drinks water'	100%	Stress
30	[ga:.'su:]	[mʷe.ga.'su:]	'he/she spits'	67%	None

TABLE 21. SURU KAVIAN–SURU MWERANI MEDUIM/LOW GROUPS: DIFFERENCES IN NUMBER OF SYLLABLES

#	SK	SM	Gloss	Comprehension	Other Factors
33	[te. me.'nip.nip]	[tem.'nip.nip]	'it is thin'	50%	None
35	['boh.ni. 'βas]	['mʷa.boh.ni. 'βet]	'he/she throws a stone'	50%	Stress; s/t/d
37	['gi.da. 'a.si]	['mʷi.gi.ta. 'a.tsi]	'he/she sees s,o.'	33%	Stress; s/t/d
38	['uⁿde.lin. bʷe:.'il]	['du.lun. bʷi.'hil]	'(chicken) egg'	33%	Stress; PreNasal
40	['gil.kil]	[mʷi.'gil.kil]	'he/she digs'	25%	None
43	['mʷe.ra.βa. 'ka:.wa]	['mʷo.rap. 'ka:.wa]	'he/she pulls a rope'	20%	VQualStrSyll
53	['mʷih. 'a.si]	['mʷa.hi. 'a.tsi]	'he/she hits s.o.'	0%	VQualStrSyll; s/t/d
54	['buh. 'ka:.wa]	['mʷa.bu.hu. 'ka:.wa]	'he/she holds a rope'	0%	Stress
55	[li.'ait]	[mʷi.li.'aut]	'he/she vomits'	0%	VQualStrSyll
56	['da.be.sok. 'a.si]	['mʷe.da.ba.so. ni. 'a.tsi]	'he/she pushes s.o.'	0%	Stress; s/t/d
57	['dok. 'ŋa.mʷa]	['mʷe.sa.dok. 'ŋa.mʷa]	'he/she is still sitting down'	0%	Stress
60	['lap. 'a.ᵑga]	['mʷa.li. 'a.ga]	'he/she takes a ship'	0%	Stress; PreNasal

There is no obvious reason why Suru Mwerani listeners would have difficulty in understanding this Suru Kavian word, in comparison to the relative ease with which they understood a similar set of words in the High group.

The higher average quantity of syllable differences in Medium and Low groups, compared to the High group, suggests that syllable difference does play a role in Suru Mwerani comprehension of Suru Kavian. However, only two of the twelve pairs (17 percent) with medium or low comprehension had "SyllDiff" as the sole distinguishing feature. This suggests that, while syllable difference does appear to have some impact on intelligibility, it usually works in concert with other variables.

5.2.3 Vowel quality differences in stressed syllables ('VQualStrSyll')

5.2.3.1 Raga. None of the words in the High or Medium groups had differences in the vowel quality of the stressed syllable. However, in the Low group, one-third of the dataset had differences in vowel quality. Vowel quality was never the sole linguistic feature to cause miscomprehension; it was one of between two and three differences, as table 22 shows.

TABLE 22. RAGA–SURU MWERANI LOW GROUP: DIFFERENCES IN VOWEL QUALITY OF STRESSED SYLLABLE

#	RA	SM	Gloss	Comprehension	Other Factors
13	[ɣa.'ma.li]	[ka.'mel]	'meeting house'	33%	SyllDiff; VelarFric
16	[bʷa.'ra.tu]	[bʷe.'ret]	'flying fox'	25%	SyllDiff
19	[mʷa. 'ma.ᵐbu]	['me.be]	'he/she rests'	25%	SyllDiff; PreNasal
26	['la.ŋo]	[leŋ]	'fly'	0%	SyllDiff
40	['va.tu]	[βet]	'stone'	0%	SyllDiff; LabFric
41	[ɣao]	['ka:.wa]	'rope'	0%	SyllDiff; VelarFric

Vowel quality on its own is probably not an important trigger of Suru Mwerani speakers' miscomprehension. However, in conjunction with other factors, it may play some role in listener miscomprehension, given that vowel quality differences appear exclusively in the Low group.

5.2.3.2 Suru Kavian. In the High group, a mean of 0.07 word pairs had differences in VQualStrSyll. This increased sevenfold to a mean of 0.5 for both Medium and Low groups.

Of the word pairs in the High group, both instances of differences in vowel quality of the stressed syllable involved an alternation between Suru Mwerani [i] and Suru Kavian [e]. In neither case did the slightly lower vowel height in the Suru Kavian word create comprehension problems for Suru Mwerani listeners. These cases are shown in table 23.

TABLE 23. SURU KAVIAN–SURU MWERANI HIGH GROUP: DIFFERENCES IN VOWEL QUALITY OF STRESSED SYLLABLE

#	SK	SM	Gloss	Comprehension	Other Factors
15	[wa. 'βe.lih]	[wa. 'βi.lih]	'grass'	100%	None
17	['se.ni]	['si.ni]	'kava'	100%	None

However, vowel quality of the stressed syllable became more divergent in the Medium-Low groups, and comprehension reduced to 50% or less. This is shown in table 24.

TABLE 24. SURU KAVIAN–SURU MWERANI MEDIUM/LOW GROUPS: DIFFERENCES IN VOWEL QUALITY OF STRESSED SYLLABLE

#	SK	SM	Gloss	Comprehension	Other Factors
34	['βe.ni]	['βi.ni]	'village'	50%	None
36	[bʷa.'ɾas]	[bʷe.'ɾet]	'flying fox'	40%	s/t/d
41	[laŋ]	[leŋ]	'fly'	25%	None
42	['a.si. ta.'bʷas]	['a.tsi te.'bʷet]	'elderly person'	25%	s/t/d
43	['mʷo.ɾap. 'ka:.wa]	['mʷe.ɾa.βa. 'ka:.wa]	'he/she pulls a rope'	20%	SyllDiff
44	['mʷi.ah]	['mʷai.ah]	'he/she falls down'	17%	None
49	[ŋo.'sin]	[ŋu.'sun]	'his/her nose'	0%	None
50	[ta.'ɾis]	[ta.'ɾut]	'people'	0%	s/t/d
53	['mʷih. 'a.si]	['mʷa.hi. 'a.tsi]	'he/she hits someone'	0%	SyllDiff; s/t/d
55	[li.'ait]	[mʷi.li.'aut]	'he/she vomits'	0%	SyllDiff
59	[βas]	[βet]	'stone'	0%	s/t/d

Examples [15] and [17] in table 23 and [34] in table 24 all involve alternations between [e] in Suru Kavian and [i] in Suru Mwerani. Perhaps an alternation between two front vowels [e] and [i], which are in relatively close proximity to each other on the vowel chart, causes less confusion for Suru Mwerani speakers than does an alternation between the other vowel pairs that have greater distance from each other. This would explain why [15] and [17] are well understood, but it does not explain why [34] is understood by only 50 percent of Suru Mwerani speakers.

Other word pairs that are distinguished solely by the vowel quality of the stressed syllable are [41], [44], and [49]. These involve alternations between a front and central vowel ([41]), a diphthong and front vowel ([44]), and a back and front vowel ([49]).

As already noted, there is a much higher incidence of words containing differences in the vowel quality of the stressed syllable in the Medium-Low groups than in the High group. Furthermore, four of eleven pairs (36 percent) with medium or low comprehension had "VQualStrSyll" as the sole distinguishing feature. This suggests that vowel quality of the stressed syllable has some impact on comprehension. There is also some evidence that variation across two front vowels impedes intelligibility less than do other types of variation.

5.2.4 Prenasalisation ("PreNasal")

5.2.4.1 Raga. None of the Raga words in the High group have prenasalization. In the Medium-Low groups, all words that had differences between Raga and Suru Mwerani in prenasalization also had other differences. See table 25.

The average presence of prenasalization—0.25 in the Medium group, lowering to 0.18 in the Low group—does not give evidence of any pattern. All examples where prenasalization existed in Raga had at least one other difference as well. Therefore, it is not possible to single out prenasalization as a cause of miscomprehension.

5.2.4.2 Suru Kavian. Only two Suru Kavian words with prenasalization fell into the High group, as shown in table 26.

TABLE 25. RAGA–SURU MWERANI MEDIUM/LOW GROUPS: PRENASALIZATION IN RAGA

#	RA	SM	Gloss	Comprehension	Other Factors
7	[mʷa. 'ᵑgi.ta. a.'ta.tu]	['mʷi.gi.ta. 'a.tsi]	'he/she sees someone'	50%	Stress; SyllDiff
19	[mʷa. 'ma.ᵐbu]	['me.be]	'he/she rests'	25%	SyllDiff; VQualStrSyll
23	[mʷa. 'lai. 'wa.ᵑga]	['mʷa.li. 'a.ga]	'he/she takes a ship'	0%	Stress
24	[mʷa. ᵐbo.'hai. 'va.tu]	['mʷa.boh.ni. 'βet]	'he/she throws a stone'	0%	Stress; SyllDiff; LabFric
30	[mʷa. ᵑgel.'ge.li]	[mʷi.'gil.kil]	'he/she digs'	0%	Stress; SyllDiff
31	[mʷa. 'ᵑga.ɣa]	[mʷe.'gak]	'he she flies'	0%	SyllDiff; VelarFric
39	[mʷa. 'ⁿdo.ɣo 'ɾa.du]	['mʷe.sa.dok. 'ŋa.mʷa]	'he/she is still sitting down'	0%	Stress; VelarFric

TABLE 26. SURU KAVIAN–SURU MWERANI HIGH GROUP: PRENASALIZATION IN SURU KAVIAN

#	SK	SM	Gloss	Comprehension	Other Factors
16	[te. 'ⁿdap]	[te. 'dap]	'it is white'	100%	None
25	[te. se.'se.ⁿde]	[te. Se.'se.de]	'it is yellow'	75%	None

The average presence of Suru Kavian prenasalization in the Medium and Low groups is 0 and 0.13, respectively; these word pairs are listed in table 27.

One of three pairs (33 percent) with medium or low comprehension had "PreNasal" as the sole distinguishing feature. The numbers across all three High-Medium-Low groups are too low to be informative. Therefore, it is not possible to single out prenasalisation as a cause of miscomprehension.

TABLE 27. SURU KAVIAN–SURU MWERANI LOW GROUP: PRENASALIZATION IN SURU KAVIAN

#	SK	SM	Gloss	Comprehension	Other Factors
38	['u.ⁿde.lin. bʷe:.'il]	['du.lun. bʷi.'hil]	'(chicken) egg'	33%	Stress; SyllDiff
47	['me.ᵐbe]	['me.be]	'he/she rests'	0%	None
60	['lap. 'a.ᵑga]	['mʷa.li. 'a.ga]	'he/she takes a ship'	0%	Stress; SyllDiff

5.2.5 Velar fricative ("VelarFric") in Raga. The average occurrence of the velar fricative in High, Medium, and Low groups was 0, 0.25, and 0.33, respectively. There are only four occurrences in total. As the Medium-Low groups in table 28 show, velar fricatives are one of between two and three variables that contribute to Suru Mwerani listener difficulties. It is safe to conclude that velar fricatives are one of a mix of variables that together contribute to listener miscomprehension.

TABLE 28. RAGA–SURU MWERANI MEDIUM/LOW GROUPS : VELAR FRICATIVE IN RAGA

#	RA	SM	Gloss	Comprehension	Other Factors
9	[ɣa.'vi.ɣa]	[ka.'βik]	'Malay apple'	40%	SyllDiff; LabFric
12	[mʷa. ma.'ta.ɣu]	[mʷam.'ta.tsi]	'he/she is afraid'	33%	SyllDiff
13	[ɣa.'ma.li]	[ka.'mel]	'meeting house'	33%	SyllDiff;VQualStrSyll
14	[nu. 'no.ɣo]	[te. 'nok]	'it is finished'	33%	SyllDiff

5.2.6 Labiodental fricative ("LabFric") in Raga. The average occurrence of the velar fricative in High, Medium, and Low groups was 0, 0.5, and 0.21, respectively. The higher average incidence of the labiodental fricative in the Medium over the Low group runs counter to the general pattern in the data of the occurrence of the variable increasing as comprehension reduces. There is no clear explanation for this. Table 29 shows the full list of word pairs distinguished by the labiodental fricative.

TABLE 29. RAGA–SURU MWERANI MEDIUM/LOW GROUPS: LABIODENTAL FRICATIVE IN RAGA

#	RA	SM	Gloss	Comprehension	Other Factors
6	[mʷa. 'ro.vo]	[mʷo.'rop]	'he/she runs'	50%	SyllDiff
9	[ɣa.'vi.ɣa]	[ka.'βik]	'Malay apple'	40%	SyllDiff; VelarFric
22	[va.'nu.a]	['βi.ni]	'village'	17%	Stress; SyllDiff
24	[mʷa. ᵐbo.'hai. 'va.tu]	['mʷa.boh.ni. 'βet]	'he/she throws a stone'	0%	Stress; SyllDiff; PreNasal
27	['na.vo]	[nap]	'wave'	0%	SyllDiff
29	[mʷa. 'rav. 'ɣao]	['mʷe.ra.βa. 'ka:.wa]	'he/she pulls a rope'	0%	Stress; SyllDiff; VelarFric
37	[nu. ma.nev.'ne.vi]	[tem.'nip.nip]	'it is thin'	0%	Stress; SyllDiff
40	['va.tu]	[βet]	'stone'	0%	SyllDiff; VQualStrSyll
42	[ta.'i.va]	['tsi.βi]	'conch shell'	0%	SyllDiff

None of the nine pairs in the Medium or Low categories distinguished between Raga and Suru Mwerani solely by the presence of the labiodental fricative. Given the unclear pattern of occurrence of the labiodental fricative across the High-Medium-Low groups, it is not possible to single it out as a cause of miscomprehension.

5.2.7 s/t/d in Suru Kavian. Although s/t/d sounds exist in Suru Mwerani as well as in Suru Kavian, the occurrence of s/t/d in Suru Kavian occurs in places that Suru Mwerani speakers might not expect. Only two pairs in the high category (an average presence of 0.07) distinguish s/t/d between Suru Kavian and Suru Mwerani; these are given in table 30. The pair in [19] has 100 percent comprehension, but, as explained in 5.2.1.2, the form [boswos] likely assisted Suru Mwerani listeners because it is constant across both varieties. In [28], s/t/d is the only linguistic factor distinguishing the pair, and the word has three syllables (and, thus, relatively few competitors with which Suru Mwerani speakers might confuse it). Yet Suru Mwerani comprehension of this word falls into the bottom margin of the high group, with only 67 percent comprehension.

TABLE 30. SURU KAVIAN–SURU MWERANI HIGH GROUP: DIFFERENCES IN s/t/d

#	SK	SM	Gloss	Comprehension	Other Factors
19	['si. bos.'wos]	['mʷi.di. bos.'wos]	'he/she stands up'	100%	Stress; SyllDiff
28	[βas.'ku:.bu]	[βat.'ku.bu]	'bamboo'	67%	None

The occurrence of s/t/d variation between Suru Kavian and Suru Mwerani increases for the Medium-Low groups, increasing fivefold from 0.07 in the High group to 0.4 in the Medium group and 0.42 in the Low group. Table 31 lists the pairs from the Medium-Low groups.

TABLE 31. SURU KAVIAN–SURU MWERANI MEDIUM/LOW GROUPS: DIFFERENCES IN s/t/d

#	SK	SM	Gloss	Comprehension	Other Factors
35	['boh.ni. 'βas]	['mʷa.boh.ni. 'βet]	'he/she throws a stone'	50%	Stress; SyllDiff
36	[bʷa.'ras]	[bʷe.'ret]	'flying fox'	40%	VQualStrSyll
37	['gi.da. 'a.si]	['mʷi.gi.ta. 'a.tsi]	'he/she sees s.o.'	33%	Stress; SyllDiff
42	['a.si. ta.'bʷas]	['a.tsi. te.'bʷet]	'elderly person'	25%	VQualStrSyll
48	[da.tu.'wan]	[ta.tsu.'wan]	'his/her sweat'	0%	None
50	[ta.'ris]	[ta.'rut]	'people'	0%	VQualStrSyll
51	[te.'kun]	[tsu.'kun]	'his/her back'	0%	None
52	[ma.'ta:.si]	[mʷam.'ta.tsi]	'he/she is afraid'	0%	None
53	['mʷih. 'a.si]	['mʷa.hi. 'a.tsi]	'he/she hits s.o.'	0%	SyllDiff; VQualStrSyll
56	['da.be.sok. 'a.si]	['mʷe.da.ba.so.ni. 'a.tsi]	'he/she pushes s.o.'	0%	Stress; SyllDiff
58	[me.'su:]	[mʷam.'tsu:]	'he/she sleeps'	0%	None
59	[βas]	[βet]	'stone'	0%	VQualStrSyll

Outside of the High group, s/t/d variation occurred as the sole linguistic difference in four of the twelve pairs (33.3 percent), which is relatively high compared to other variables. Since the overall prevalence of s/t/d variation is also much higher in the Medium-Low groups than it is in the High group, it probably has some impact on the ability of Suru Mwerani speakers to understand Suru Kavian words.

5.2.8 Summary. Of the six linguistic variables investigated for their effect on intelligibility of Raga for Suru Mwerani speakers, there is no "smoking gun." Rather, it is probably the accumulation of differences that affects intelligibility. Differences in vowel quality of the stressed syllable may make a greater contribution to miscomprehension.

The primary factors impeding Suru Mwerani understanding of Suru Kavian are the existence of differences in s/t/d and differing vowel quality of the stressed syllable. Word stress and a difference in the number of syllables also appear to have a lesser impact on intelligibility. The final linguistic feature, prenasalization, has no clear impact on intelligibility. And as with Raga, multiple linguistic differences across a word pair also tend to lead to lower Suru Mwerani comprehension of Suru Kavian words.

5.3 LOCATION OF VARIATION WITHIN THE WORD. Suru Mwerani listeners generally understood less if a linguistic difference between the Suru Mwerani word and its cognate counterpart occurred in the first syllable, and more if the first syllable was the same across cognates. A "linguistic difference" is any linguistic differentiation between the first syllable of the two words. This includes a difference in syllable stress.

For the Raga–Suru Mwerani High Group, three ([1], [2], [4]) of the five word pairs (60 percent) have a linguistic difference in the first syllable. This contrasts with the Low Group, where 31 of the 33 total word pairs (94 percent) have a feature contrast in the first syllable. If we exclude syllable stress from this measure, then the number of initial-syllable differences reduces to 1/5 (20 percent) in the High Group, and to 27/33 (82 percent) in the Low Group.

For the Suru Kavian–Suru Mwerani High Group, 10/31 (32 percent) of the word pairs have a linguistic difference in the first syllable. This contrasts with the Low Group, where 19 of the 24 total word pairs (79 percent) have a feature contrast in the first syllable.

These results are summarized in table 32. In all cases, initial syllable difference appears to play a role in Suru Mwerani listener comprehension.

TABLE 32. OCCURRENCE OF LINGUISTIC DIFFERENCES IN INITIAL SYLLABLE IN PROPORTION TO TOTAL NO. OF COGNATE WORD PAIRS

	RA–SM		SK–SM	
	Any difference in feature	Any difference in feature, excluding syllable stress	Any difference in feature	Any difference in feature, excluding syllable stress
High Group	3/5 (60%)	1/5 (20%)	10/31 (32%)	N/A
Low Group	32/33 (97%)	28/33 (85%)	19/24 (79%)	N/A

5.4 RELATIVE INCOMPREHENSION OF SURU KAVIAN OVER RAGA. Overall, Suru Mwerani speakers have a better understanding of Suru Kavian than of Raga. This is unsurprising, because Suru Kavian has a closer genetic relationship to Suru Mwerani than Raga does. Yet, relative to Levenshtein distances, Suru Mwerani speakers have a poorer understanding of Suru Kavian than they do of Raga. This is shown in table 33.

TABLE 33. SURU MWERANI COMPREHENSION AND LEVENSHTEIN DISTANCE WITH RAGA AND SURU KAVIAN

	RA			SK		
Grouping	H	M	L	H	M	L
Average Suru Mwerani Comprehension (%)	80	45	11	92	48	8
Average Levenshtein Distance (%)	41	50	61	14	25	43

When Raga–Suru Mwerani word pairs have an average Levenshtein distance of 41, Suru Mwerani speakers are able to correctly translate 80 percent of the words they hear. In contrast, with the slightly higher average Levenshtein distance of 43 between Suru Mwerani and Suru Kavian, Suru Mwerani speakers successfully translate a mere 8 percent of the Suru Kavian words they hear. In other words, Suru Mwerani speakers' com-

prehension of Raga is about ten times as high as that of Suru Kavian, relative to Levenshtein distance.

Why is this so? The anomaly is probably not due to greater exposure to Raga because, as noted in section 2, testing has shown that Suru Mwerani comprehension of Raga and Suru Kavian noncognates is 19 percent and 14.9 percent, respectively, which is quite low. (However, it is also surprising that Suru Mwerani speakers understand Raga noncognates better than Suru Kavian noncognates.)

One possible explanation for this paradox hinges on the fact that there is indeed a close linguistic relationship between Suru Kavian and Suru Mwerani. As stated earlier, the two varieties share 90 percent cognacy. Since the two varieties have much in common (19 of the 60 word pairs are, in fact, identical), then when unpredictable linguistic variation does occur, Suru Mwerani speakers seem to become even more confused than they do when they hear more distantly related Raga words. In other words, the "unfamiliar," interwoven with the "familiar" in unanticipated ways, creates linguistic dissonance for Suru Mwerani speakers. The practical effect of this is that low comprehension levels for Suru Kavian are correlated with lower Levenshtein distances than is the case for Raga. The Medium grouping in table 33 is a case in point: Suru Mwerani comprehension of Raga and Suru Kavian sits at 45 percent and 48 percent, respectively, yet the Suru Kavian–Suru Mwerani Levenshtein distance is half that of Raga–Suru Mwerani.

A social explanation is that Raga speakers occupy a place of relative privilege not only in Pentecost but in Vanuatu more widely, and the prestige enjoyed by this community may have flow-on effects to the language itself. Outsiders may be more inclined to take notice of Raga when they hear it, even if their overall exposure to (and knowledge of) Raga is low. This, of course, assumes that Suru Mwerani speakers were able to recognize Raga as a test language: before the test, participants were told which languages they would hear, but at any given point they were not told which specific language they were listening to.

An additional consideration is that Suru Mwerani speakers generally view Suru Kavian with some ambivalence, considering it to be a "different" and "difficult" variety that is not easily understood by outsiders (see Schneider (2017:6–7)). It is possible that when some Suru Mwerani participants heard Suru Kavian in the recording, their preconceived notions about the difficulty of this variety led them to dismiss it, leading to lower comprehension scores. For example, the first author recalls a male participant listening to the recordings with headphones on and muttering "Suru Kavian…" while shaking his head in dismay!

It is useful to examine other studies with regard to the impact of social factors on intelligibility. Wolff (1959) studied mutual intelligibility between two Nigerian languages and proposed that a negative attitude toward the other language might mean that speakers would be less willing to decode it or, conversely, a positive attitude would motivate listeners to exert greater effort to understand. As Schüppert and Gooskens (2012) point out, however, the argument could be turned around: rather than a positive/negative attitude bringing about higher/lower intelligibility, it could instead be that higher/lower intelligibility engenders a positive/negative attitude. As a case in point, Delsing and Lundin

Åkesson (2005) found a relationship between attitude and intelligibility for Danish and Swedish speakers, but the direction of causality was unclear. Gooskens (2006) studied inter-Scandinavian intelligibility and found that the group with negative attitudes toward the other language had more difficulties in decoding that language. However, the correlation is weak, and another study by Schüppert and Gooskens (2011:135) concluded that attitude played no role in comprehension across Danish and Swedish. The results in this area are, therefore, mixed, and more research is needed.

6. CONCLUSION. We have explained instances of high intelligibility and high Levenshtein distance for both Raga–Suru Mwerani and Suru Kavian–Suru Mwerani. In both cases, Suru Mwerani speakers were likely familiar with the target word through their exposure to a similar form in Suru Rabwanga. Therefore, the lesson for the design of future tests is to actively anticipate the unexpected. One way to do this would be to conduct pilot tests before conducting the actual survey.[5]

For low intelligibility/low Levenshtein distance in both the Raga–Suru Mwerani and Suru Kavian–Suru Mwerani word pairs, we have no clear explanation as to why some Raga and Suru Kavian words that had low intelligibility for Suru Mwerani speakers also had a low Levenshtein distance across the cognate pair. When test takers did not understand a word, they generally provided no response at all to the prompt. In future tests, if time allows, it could be very useful to revisit nonresponses with the listener once the test was completed. By probing listeners, we could gain some insight into why they did not understand words that they would have been expected to easily comprehend.

There is also evidence from the Suru Kavian data to suggest that monosyllables caused more difficulty for Suru Mwerani listeners in comparison to longer words, at least under test conditions. When designing future intelligibility tests, we may, therefore, want to consider avoiding the use of monosyllables as test words. However, a major disadvantage of such an approach would inevitably be that common words in the language would not be represented in the sampling.

With regard to the effect of individual linguistic variables, differences in the vowel quality of the stressed syllable appear to impede Suru Mwerani comprehension of the Raga data to a limited extent. It is probably safe to conclude, however, that Suru Mwerani comprehension of Raga is compromised when more than one feature is present in the Raga word form to set it apart from the Suru Mwerani forms. It is this combination of differences that appears to maximize Suru Mwerani listener confusion.

For the Suru Kavian–Suru Mwerani word pairs, we have found some evidence that if the vowel quality of the stressed Suru Kavian syllable differs from its Suru Mwerani counterpart, then Suru Mwerani listener comprehension is affected. There is also limited evidence that if the Suru Kavian–Suru Mwerani variation occurs across two front vowels, intelligibility is impeded less than for other types of vocalic variation. This, therefore, suggests that vowel quality is an important factor for listener comprehension.

Another linguistic variable that affects Suru Mwerani comprehension is the presence of the native sounds [s]/[t]/[d] in places that are unpredictable for Suru Mwerani speakers. On the other hand, prenasalization, which is technically a "foreign" sound for Suru

5. The present study is itself a pilot test for a planned larger study.

Mwerani speakers, has no discernible impact on comprehension. This would appear to be at odds with the findings of Kürschner, Gooskens, and van Bezooijen (2008:88), where foreign sounds were indeed correlated significantly with intelligibility. However, as van Heuven (2008:46) points out, "only when the discrepancy between an incoming sound and any existing prototype is very large, will the listener refuse to categorize the incoming sound." As already noted, most Suru Mwerani speakers would be familiar with prenasalized sounds through their exposure to Suru Rabwanga. And although the Raga velar fricative /ɣ/ and the labiodental fricative /v/ are not used in Suru Mwerani or any dialect of Apma, the sounds have close neighbors that do exist in Apma, namely, /k/ and /β/ respectively.

Word stress and a difference in number of syllables also appear to affect Suru Mwerani listener comprehension, though to a lesser extent. These results, therefore, support the assertions of Harrington and Cox (2009) and Kürschner, Gooskens, and van Bezooijen (2008: 93) (as discussed in 4.1 and 4.2, respectively). And as with Suru Mwerani comprehension of Raga, multiple differences across word pairs usually lead to lower intelligibility.

Finally, in accordance with van Heuven (2008:1) and Cutler (2012), there is evidence that if linguistic variation occurs in the initial syllable of a cognate word pair, then lexical activation is more likely to be disrupted, and listener comprehension consequently impaired, than when there are no differences in the initial syllable.

Suru Mwerani speakers understand Raga much better than Suru Kavian, relative to Levenshtein distance. This is unexpected because Suru Kavian is genetically closer to Suru Mwerani than Raga is. One explanation is that, since Suru Kavian and Suru Mwerani are linguistically very similar, when differences such as s/t/d usage or vowel quality of the stressed syllable occur, they cause a disproportionate amount of miscomprehension for Suru Mwerani speakers. Another possibility is that the high prestige of the Raga community gives the Raga language a high profile and this increases outsiders' relative receptivity to and comprehension of the language. Conversely, Suru Mwerani speakers' ambivalent attitude toward the Suru Kavian variety may contribute to their relatively low comprehension of Suru Kavian. However, more research is needed to determine whether it is indeed attitude that affects intelligibility, or the other way around.

It is hoped that the knowledge gained from this study will serve as a useful starting point for similar research on other Pacific languages. There are numerous examples in the Pacific where small, vulnerable languages are in danger of being overtaken by larger and stronger ones. The perilous status of Suru Kavian is a case in point. An understanding of the specific linguistic factors that contribute to miscommunication will give linguists and school teachers the tools they need to highlight these differences to others by way of contrastive analysis. This will improve general metalinguistic awareness of similarities and differences across related varieties. In this way, then, this study and others like it can facilitate language awareness and language maintenance.

APPENDIX

1. RAGA-SURU MWERANI COMPREHENSION AND LEVENSHTEIN DISTANCE MEASUREMENTS, AND VARIABLES THAT MAY AFFECT INTELLIGIBILITY

#	RA	SM	Gloss	Comp	Lev	Stress	Syll Diff	VQual StrSyll	Pre Nasal	Velar Fric	Lab Fric	#Diff Ftrs
1	[tel.ˈtele]	[ˈtel.tel]	'snake'	100	14	1	1	N/A	0	0	0	2
2	[mʷa. ˈmin. ˈwai]	[ˈmʷam.ni si.ˈleŋ]	'he/she drinks water'	83	70	1	1	N/A	0	0	0	2
3	[ˈbo.e]	[bo]	'pig'	75	33	0	1	0	0	0	0	1
4	[mʷa. ˈma.na]	[man]	'he/she laughs'	75	50	0	1	0	0	0	0	1
5	[mʷa. ma.ˈtu.ru]	[mʷam.ˈtsɯː]	'he/she sleeps'	67	38	0	1	0	0	0	0	1
	RA-SM Mean Intelligibility – High Group			**80**	**41**	**0.4**	**1**	**0**	**0**	**0**	**0**	**1.4**
6	[mʷa. ˈro.vo]	[mʷo.ˈrop]	'he/she runs'	50	50	0	1	0	0	0	1	2
7	[mʷa. ᵑgi.ɹa. a.ˈta.ɯ]	[ˈmʷi.gi.ɹa ˈa.tsi]	'he/she sees s.o.'	50	55	1	1	N/A	1	0	0	3
8	[mʷa. ro.ˈŋo.e]	[mʷo.ˈro.ŋo]	'he/she hears'	40	29	1	1	N/A	0	0	0	2
9	[ɣa.ˈvi.ɣa]	[ka.ˈβik]	'Malay apple'	40	67	0	1	0	0	1	1	3
	RA-SM Mean Intelligibility – Middle Group			**45**	**50**	**0.5**	**1**	**0**	**0.25**	**0.25**	**0.5**	**2.5**
10	[bu.tu.ˈbu.tu]	[ˈbut.but]	'ant'	33	25	1	1	0	0	0	0	2
11	[ˈbʷe.ta]	[bʷet]	'taro'	33	25	0	1	0	0	0	0	1
12	[mʷa. ma.ˈta.ɣu]	[mʷam.ˈtatsi]	'he/she is afraid'	33	38	0	1	0	0	1	0	2
13	[ɣa.ˈma.li]	[ka.ˈmel]	'meeting house'	33	50	0	1	1	0	1	0	3
14	[nu ˈneɣo]	[te ˈnɔk]	'it is finished'	33	67	0	1	0	0	1	0	2
15	[bʷan.se.re.ˈse.re]	[bʷan.seː.ˈseː]	'centipede'	25	18	0	1	0	0	0	0	1
16	[bʷa.ˈra.tu]	[bʷe.ˈret]	'flying fox'	25	50	0	1	1	0	0	0	2
17	[mʷa. ˈlu.a]	[mʷi.li.ˈaut]	'he/she vomits'	25	67	1	0	N/A	0	0	0	1
18	[ˈme.re]	[mi.ˈsin]	'his/her urine'	25	80	1	0	N/A	0	0	0	1
19	[mʷa. ˈma.^{n}bu]	[ˈme.be]	'he/she rests'	25	83	0	1	1	1	0	0	3
20	[ˈu.hi]	[ˈbʷa.rus]	'pawpaw'	25	83	1	0	N/A	0	0	0	1
21	[mʷa. ˈwe.hi. a.ˈta.tu]	[ˈmʷa.hi. ˈa.tsi]	'he/she hits s.o.'	20	55	1	1	N/A	0	0	0	2
22	[va.ˈnu.a]	[ˈβi.ni]	'village'	17	80	1	1	N/A	0	0	1	3

#	RA	SM	Gloss	Comp	Lev	Stress	Syll Diff	VQual Str Syll	Pre Nasal	Velar Fric	Lab Fric	#Diff Ftrs
23	[mʷa. 'lai 'wa.ᵑga]	['mʷa.li. 'a.ga]	'he/she takes a ship'	0	38	1	9	N/A	1	0	0	2
24	[mʷa. ^{n}bo.'hai. 'va.tu]	['mʷa.boh.ni. 'βet]	'he/she throws a stone'	0	55	1	1	N/A	1	0	1	4
25	['i.do.lin. 'ma.nu]	['du.lun. bʷi.'hil]	'(chicken) egg'	0	50	1	1	N/A	0	0	0	2
26	['la.ŋo]	[leŋ]	'fly'	0	50	0	1	1	0	0	0	2
27	['na.vo]	[nap]	'wave'	0	50	0	1	0	0	0	1	2
28	[ba.'ɣe.o]	['be.ke]	'shark'	0	60	1	1	N/A	0	1	0	3
29	[mʷa. 'rav. 'ɣao]	['mʷe.ra.βa. 'ka:.wa]	'he/she pulls a rope'	0	64	1	1	N/A	0	1	1	4
30	[mʷa. ᵑgel.'ge.li]	[mʷi.'gil.kil]	'he/she digs'	0	67	1	1	N/A	1	0	0	3
31	[mʷa. ᵑga.ɣa]	[mʷe.'gak]	'he/she flies'	0	67	0	1	0	1	1	0	3
32	[ba.'ha.ra]	[bas]	'cliff'	0	67	1	1	N/A	0	0	0	2
33	[nu. 'me.to]	[te. 'me:]	'it is black'	0	67	0	1	0	0	0	0	1
34	[nu. 'me.ho]	[te. 'mes]	'it is wet'	0	67	0	1	0	0	0	0	1
35	[mʷa. so.'ɣai. a.'ta.tu]	['mʷe.da.ba.so.ni. 'a.tsi]	'he/she pushes s.o.'	0	67	1	1	N/A	0	1	0	3
36	[nu. ma.ma.'ɣa.ni]	[tem.'kan]	'it is sharp'	0	70	0	1	0	0	1	0	2
37	[nu. ma.nev.'ne.vi]	[tem.'nip.nip]	'it is thin'	0	73	1	1	N/A	0	0	1	3
38	[ɣa.ti.'ɣu.na]	[tsu.'kun]	'his/her back'	0	75	0	1	0	0	1	0	2
39	[mʷa. ^{n}do.ɣo. 'ra.du]	['mʷe.sa.dok. 'ŋa.mʷa]	'he/she is still sitting down'	0	75	1	0	N/A	1	1	0	3
40	['va.tu]	[βet]	'stone'	0	75	0	1	1	0	0	1	3
41	[ɣao]	['ka:.wa]	'rope'	0	80	0	1	1	0	1	0	3
42	[ta.'i.va]	['tsi.βi]	'conch shell'	0	80	0	1	0	0	0	1	2
	RA-SM Mean Intelligibility – Low Group			**11**	**61**	**0.49**	**0.85**	**0.33**	**0.18**	**0.33**	**0.21**	**2.2**

2. SURU KAVIAN–SURU MWERANI COMPREHENSION AND LEVENSHEIN DISTANCE MEASUREMENTS, AND VARIABLES THAT MAY AFFECT INTELLIGIBILITY

#	SK	SM	Gloss	Comp	Lev	Stress	Syll Diff	VQual StrSyll	Pre Nasal	s/t/d	#Diff Ftrs
1	['ku.li]	['ku.li]	'dog'	100	0	0	0	0	0	0	0
2	['be.kɛ]	['be.kɛ]	'shark'	100	0	0	0	0	0	0	0
3	[te. 'nok]	[te. 'nok]	'it is finished'	100	0	0	0	0	0	0	0
4	[bʷet]	[bʷet]	'taro'	100	0	0	0	0	0	0	0
5	['te. ka.'bis]	['te. ka.'bis]	'it is good'	100	0	0	0	0	0	0	0
6	[si.'leŋ]	[si.'leŋ]	'water'	100	0	0	0	0	0	0	0
7	[rep]	[rep]	'hill'	100	0	0	0	0	0	0	0
8	['ka:.wa]	['ka:.wa]	'rope'	100	0	0	0	0	0	0	0
9	[ka.'βik]	[ka.'βik]	'Malay apple'	100	0	0	0	0	0	0	0
10	['mʷap.mʷap]	['mʷap.mʷap]	'it is hot'	100	0	0	0	0	0	0	0
11	['tel.tel]	['tel.tel]	'snake'	100	0	0	0	0	0	0	0
12	[ta.bʷa.'ken]	[ta.bʷa.'ken]	'mosquito'	100	0	0	0	0	0	0	0
13	[mʷo.'rop]	[mʷo.'rop]	'he/she runs'	100	0	0	0	0	0	0	0
14	[te. 'mes]	[te. 'mes]	'it is wet'	100	0	0	0	0	0	0	0
15	[wa. 'βe.lih]	[wa. 'βi.lih]	'grass'	100	14	0	0	1	0	0	1
16	[te. ⁿdap]	[te. 'dap]	'it is white'	100	20	0	0	0	1	0	1
17	['se.ni]	['si.ni]	'kava'	100	25	0	0	1	0	0	1
18	['sol.sol]	[mʷo.'sol.sol]	'he/she sews'	100	25	0	1	0	0	0	1
19	['si. bos.'wos]	['mʷi.di. bos.'wos]	'he/she stands up'	100	30	1	1	N/A	0	1	3
20	[ma.'de:.de]	[mʷa.'te:.te]	'chicken'	100	43	0	0	0	0	0	0
21	['sa.sah]	[mʷe.'sa:.sa:]	'he/she sings'	100	56	0	1	0	0	0	1
22	['min. si.'leŋ]	['mʷam.ni. si.'leŋ]	'he/she drinks water'	100	67	1	1	N/A	0	0	2
23	[ma.map.'lel]	[ma.map.'lel]	'he/she plays'	80	0	0	0	0	0	0	0
24	[mu:m]	[mu:m]	'he/she works'	75	0	0	0	0	0	0	0
25	[te. se.'se.ⁿde]	[te. se.'se.de]	'it is yellow'	75	13	0	0	0	1	0	1
26	[sa.sa.ra.'kan]	[se.ses.ra.'kan]	'teacher'	75	30	0	0	0	0	0	0
27	[bu]	[bu]	'knife'	67	0	0	0	0	0	0	0
28	[βas.'ku:.bu]	[βat.'ku.bu]	'bamboo'	67	14	0	0	0	0	1	1
29	[mʷo.'ro.ŋa]	[mʷo.'ro.ŋo]	'he/she hears'	67	17	0	0	0	0	0	0
30	[ga:.'su:]	[mʷe.ga.'su:]	'he/she spits'	67	29	0	1	0	0	0	1
31	[bʷe:.'il]	[bʷi.'hil]	'bird'	67	50	0	0	0	0	0	0
	SK-SM Mean Intelligibility – High Group			**92**	**14**	**0.06**	**0.16**	**0.07**	**0.06**	**0.07**	**0.42**

#	SK	SM	Gloss	Comp	Lev	Stress	Syll Diff	VQual StrSyll	Pre Nasal	s/t/d/	#Diff Ftrs
32	[bas]	[bas]	'cliff'	50	0	0	0	0	0	0	0
33	[te. me.'nip.nip]	[tem.'nip.nip]	'it is thin'	50	10	0	1	0	0	0	1
34	['βe.ni]	['βi.ni]	'village'	50	25	0	0	1	0	0	1
35	['boh.ni. 'βas]	['mʷa.boh.ni. 'βet]	'he/she throws a stone'	50	29	1	1	N/A	0	1	3
36	[bʷa.'ras]	[bʷe.'ret]	'flying fox'	40	60	0	0	1	0	1	2
	SK-SM Mean Intelligibility – Middle Group			**48**	**25**	**0.2**	**0.4**	**0.5**	**0**	**0.4**	**1.4**
37	['gi.da. 'a.si]	['mʷi.gi.ta. 'a.tsi]	'he/she sees someone'	33	50	1	1	N/A	0	1	3
38	['u.ⁿde.lin. bʷe:.'il]	['du.lun. bʷi.'hil]	'(chicken) egg'	33	58	1	1	N/A	1	0	3
39	[me.'sin]	[mi.'sin]	'his/her urine'	25	20	0	0	0	0	0	0
40	['gil.kil]	[mʷi.'gil.kil]	'he/she digs'	25	25	0	1	0	0	0	1
41	[laŋ]	[leŋ]	'fly'	25	33	0	0	1	0	0	1
42	['a.si. ta.'bʷas]	['a.tsi. te.'bʷet]	'elderly person'	25	50	0	0	1	0	1	2
43	['mʷo.rap. 'ka:.wa]	['mʷe.ra.βa. 'ka:.wa]	'he/she pulls a rope'	20	50	0	1	1	0	0	2
44	['mʷi.ah]	['mʷai.ah]	'he/she falls down'	17	25	0	0	1	0	0	1
45	[nap]	[nap]	'wave'	0	0	0	0	0	0	0	0
46	[te. 'met]	[te. 'me:]	'it is black'	0	20	0	0	0	0	0	0
47	['me.ᵐbe]	['me.be]	'he/she rests'	0	25	0	0	0	1	0	1
48	[da.tu.'wan]	[ta.tsu.'wan]	'his/her sweat'	0	29	0	0	0	0	1	1
49	[ŋo.'sin]	[ŋu.'sun]	'his/her nose'	0	40	0	0	1	0	0	1
50	[ta.'ris]	[ta.'rut]	'people'	0	40	0	0	1	0	1	2
51	[te.'kun]	[tsu.'kun]	'his/her back'	0	40	0	0	0	0	1	1
52	[ma.'ta:.si]	[mʷam.'ta.tsi]	'he/she is afraid'	0	50	0	0	0	0	1	1
53	['mʷih. 'a.si]	['mʷa.hi. 'a.tsi]	'he/she hits someone'	0	50	0	1	1	0	1	3
54	['buh. 'ka:.wa]	['mʷa.bu.hu. 'ka:.wa]	'he/she holds a rope'	0	50	1	1	N/A	0	0	2
55	[li.'ait]	[mʷi.li.'aut]	'he/she vomits'	0	50	0	1	1	0	0	2
56	['da.be.sok. 'a.si]	['mʷe.da.ba.so.ni. 'a.tsi]	'he/she pushes someone'	0	50	1	1	N/A	0	1	3
57	['dok. 'ŋa.mʷa]	['mʷe.sa.dok. 'ŋa.mʷa]	'he/she is still sitting down'	0	57	1	1	N/A	0	0	2
58	[me.'su:]	[mʷam.'tsu:]	'he/she is sleeping'	0	67	0	0	0	0	1	1
59	[βas]	[βet]	'stone'	0	67	0	0	1	0	1	2
60	['lap. 'a.ᵑga]	['mʷa.li. 'a.ga]	'he/she takes a ship'	0	80	1	1	N/A	1	0	3
	SK-SM Mean Intelligibility – Low Group			**8**	**43**	**0.25**	**0.42**	**0.5**	**0.13**	**0.42**	**1.58**

REFERENCES

Casad, Eugene H. 1974. *Dialect intelligibility testing.* Summer Institute of Linguistics Publications in Linguistics and Related Fields, 38. Norman: Summer Institute of Linguistics of the University of Oklahoma.

Cutler, Anne. 2012. *Native listening: Language experience and the recognition of spoken words*. Cambridge, MA: Massachusetts Institute of Technology.

Delsing, Lars Olof, and Katarina Lundin Akesson. 2005. *Håller språket ihop Norden? En forskningsrapport om ungdomars förståelse av danska, svenska och norska* [*Does the language keep together the Nordic countries? A research report of mutual comprehension between young Danes, Swedes and Norwegians*]. TemaNord 573. Copenhagen: Nordic Council of Ministers.

François, Alexandre, Michael Franjieh, Sébastien Lacrampe, and Stefan Schnell. 2015. The exceptional linguistic density of Vanuatu. In *The languages of Vanuatu: Unity and diversity*, ed. by Alexandre François, Sébastien Lacrampe, Michael Franjieh, and Schnell, 1–21. Studies in the Languages of Island Melanesia, 5. Canberra: Asia-Pacific Linguistics Open Access.

Gooskens, Charlotte. 2006. Linguistic and extra-linguistic predictors of inter-Scandinavian intelligibility. *Linguistics in the Netherlands 2006*, ed. by Jeroen van de Weijer and Bettelou Los, 101–13. Amsterdam: John Benjamins.

———. 2013. Experimental methods for measuring intelligibility of closely related language varieties. In *Handbook of sociolinguistics*, ed. by Robert Bayley, Richard Cameron, and Ceil Lucas, 195–213. Oxford University Press.

Gooskens, Charlotte, Wilbert Heeringa, and Karin Beijering. 2008. Phonetic and lexical predictors of intelligibility. *International Journal of Humanities and Arts Computing* 2(1–2):63–81.

Gooskens, Charlotte, and Cindy Schneider. 2016. Testing mutual intelligibility between closely related languages in an oral society. *Language Documentation and Conservation* 10:278–305.

Gooskens, Charlotte, Renée van Bezooijen, and Vincent J. van Heuven. 2015. Mutual intelligibility of Dutch-German cognates by children: The devil is in the detail. *Linguistics* 53(2):255–83.

Gooskens, Charlotte, and Vincent van Heuven. 2017. Measuring cross-linguistic intelligibility in the Germanic, Romance and Slavic language groups. *Speech Communication* 89:25–36.

Gray, Andrew. 2012. *The languages of Pentecost Island*. Ashford, UK: Manples (British Friends of Vanuatu) Publishing.

Harrington, Jonathan, and Felicity Cox. 2009. Phonetics and phonology: The foot and word stress. Department of Linguistics, Macquarie University. Online at http://clas.mq.edu.au/speech/phonetics/phonology/syllable/syll_foot.html. [Accessed August 17, 2017.]

Hickerson, Harold, Glen D. Turner, and Nancy. P. Hickerson. 1952. Testing procedures for estimating transfer of information among Iroquois dialects and languages. *International Journal of American Linguistics* 18:1–8.

Hilton, Nanna H., Charlotte Gooskens, and Anja Schüppert. 2013. The influence of non-native morphosyntax on the intelligibility of a closely related language. *Lingua* 137:1–18.

Kürschner, Sebastian, Charlotte Gooskens, and Renée van Bezooijen. 2008. Linguistic determinants of the intelligibility of Swedish words among Danes. *International Journal of Humanities and Arts Computing* 2(1–2):83–100.

Lynch, John, and Terry Crowley. 2001. *Languages of Vanuatu: A new survey and bibliography.* Canberra: Pacific Linguistics.

Nahhas, Ramzi W. 2006. *The steps of recorded text testing: A practical guide.* Chiang Mai: Payap University.

Nerbonne, John, and Wilbert Heeringa. 2010. Measuring dialect differences. In *Language and space: Theories and methods*, ed. by J. E. Schmidt and P. Auer, 550–67. Handbooks of Linguistics and Communication Science. Berlin: Mouton De Gruyter,

Pierce, Joe E. 1952. Dialect distance testing in Algonquian. *International Journal of American Linguistics* 18:208–18.

Schneider, Cindy. 2017. Ideologies, practices, and their effect on dialect vitality: A case study from Vanuatu. *Journal of Multilingual and Multicultural Development*, DOI: 10.1080/01434632.2017.1311904.

Schneider, Cindy, and Andrew Gray. 2015. Is it worth documenting "just a dialect"? Making the case for Suru Kavian (Pentecost island). In *The languages of Vanuatu: Unity and diversity*, edited by Alexandre François, Sébastian Lacrampe, Michael Franjieh, and Stefan Schnell, 197–216. Studies in the Language of Island Melanesia, 5. Canberra: Asia-Pacific Linguistics Open Access.

Schüppert, Anja, and Charlotte Gooskens. 2011. Investigating the role of language attitudes for perception abilities using reaction time. *Dialectologia Special Issue II*, 119–40.

———. 2012. The role of extra-linguistic factors in receptive bilingualism: Evidence from Danish and Swedish pre-schoolers. *International Journal of Bilingualism* 16(3):332–47.

Stahl, Jim. 1994. Intelligibility survey of the Pango, Eton and Lelepa speech communities. Unpublished MS, SIL Vanuatu.

Stahl, Jim. n.d.. A language survey of north central Santo. Unpublished MS, SIL Vanuatu.

Tryon, D. T. 1976. *New Hebrides languages: An internal classification*. Canberra: Pacific Linguistics.

Van Heuven, Vincent J. 2008. Making sense of strange sounds: (Mutual) intelligibility of related language varieties. A review. *International Journal of Humanities and Arts Computing* 2(1–2):39–62.

Vanhove, Jan. 2014. Receptive multilingualism across the lifespan: Cognitive and linguistic factors in cognate guessing. PhD thesis, University of Fribourg.

Vari-Bogiri, Hannah. 2011. Phonology and morpho-syntax of Raga, Vanuatu. PhD thesis, University of the South Pacific.

Voegelin, C. F., and Zellig S. Harris. 1951. Methods for determining intelligibility among dialects of natural languages. *Proceedings of the American Philosophical Society* 45:322–29.

Weisser, Martin. 2005. 24.1: Word stress. (Last updated April 24, 2013.) Online at http://martinweisser.org/courses/phonetics/supra/stress.html. [Accessed May 31, 2017.]

Wolff, Hans. 1959. Intelligibility and inter-ethnic attitudes. *Anthropological Linguistics* 1:4-41.

cschnei3@une.edu.au c.s.gooskens@rug.nl