Variation in creole languages: insights from a Swadesh list[1]

Marlyse Baptista[2]
University of Michigan

The purpose of this paper is to show that the use of Swadesh lists in language documentation can bring to light the complex and multilayered variation that exists in archipelago settings. The Swadesh list under study reveals that the traditional divide between acrolect and basilect on the one hand and between leeward and windward varieties on the other does not reveal well-defined boundaries across the lects/varieties, except to demonstrate the dramatic variation that occurs within the same oral language. We show that Le Page & Tabouret-Keller's (1985) proposal that speakers consistently mix lects rather than confining themselves to one point of the creole continuum is supported by the empirical evidence found in the Swadesh list. Coseriu's (1981) three-dimensional model of diasystematic variation is also validated: the three dimensions involving diatopic (regional), diastratic and diaphasic (spoken, oral language) variation illustrate that the development of any language can be best described by taking into account the fundamental distinction between written and spoken language which cannot be reduced to diasystematic differences. This ultimately points to the importance of the idiolect as a crucial site of variation (Mufwene 2001).

Keywords: Cabo Verdean Creole, diatopic, basilectal, acrolectal and idiolectal variation, Swadesh list.

1. Introduction

Irrespective of their area of specialization, most linguists have to contend with language variation and the challenges it poses for straightforward linguistic analysis if one wishes to render an accurate description of a linguistic feature or set of features. Variation due to linguistic and extralinguistic factors "affects" every subfield of our discipline. In phonology, the choice of variants is largely (but not always, as discussed in section 5) dictated by the environments in which the phonemes appear. In morphology, variation can emerge from the range of

[1] I am deeply indebted to two anonymous reviewers for their insightful comments on a previous version of this paper. Their input greatly contributed to improving this paper. All errors or shortcomings remain, of course, my own.

[2] With the assistance of the students in the first cohort of the Masters' in creolistics University of Cape Verde.

suffixes a head noun, a head verb or other units can take. For syntax, a noun may change shape when it changes cases, based on its role and position in the sentence (Raumolin-Brunberg 1988). Labov (1972) and before him the sociologist Goffman (1964/1972) were pioneering in identifying a set of extralinguistic factors that can affect a speaker's choice of a variant over another. Such factors include age, sex, ethnic group, social class, level of education, profession and overall social attitudes. In addition, geographic boundaries can give rise to much variation between the dialects of the same language, leading dialectologists to design isoglosses capturing the set of features characterizing the regions they study. LePage & Tabouret-Keller (1985) argued for a dialectal continuum punctuated by a basilect, acrolect and mesolect, such lects cutting across both geographic and social boundaries. Linguists like Halliday (1978) consider dialect variation as conveying the diversity of social structures in society whereas scholars like Coseriu views every variety through the lens of three dimensions, diatopic, diastratic and diaphasic. We return to to these notions and to these scholars' works in section 3.

Just like other linguists, creolists have had to wrestle with variation whether collecting data in the field or analyzing a given creole language based on grammatical descriptions, as descriptions may differ from each other. More recently, a particular emphasis has been placed on variation at the individual level, re-orienting sociolinguistic research towards the study of idiolectal variation (Mufwene 2001).

This paper focuses on dialectal variation across five islands of Cabo Verde, an archipelago situated off the coast of Senegal and long-known for the drastic variation that characterizes its *Kriolu* language, or Cabo Verdean,[3] spoken in its windward and leeward islands.

There are nine inhabited islands and one uninhabited one in the archipelago but for logistical reasons having to do with access to consultants, this paper focuses on five islands only.

The five islands under study include on the one hand Brava, Fogo and Santiago (Sotavento or leeward islands) and on the other Santo Antão and São Vicente (Barlavento or windward islands). The leeward islands have been traditionally viewed as reflecting basilectal variation whereas the windward

[3] As of last year (2018), the Cabo Verdean government of Cabo Verde islands has decided that the English name of the language spoken in Cabo Verde should be Cabo Verdean, a label that is adopted throughout this article.

islands are associated with acrolectal variation. We discuss the reasons for such a taxonomy in sections 2 and 3.

The main objective of this exploratory paper is to show how using a Swadesh list can cast an interesting light on dialectal variation. The use of a Swadesh list allows for systematic comparisons between leeward and windward islands and within the two groups, revealing on which sets of lexical items the members of each group of islands converge and where they diverge. I propose that the Swadesh list reflects both dialectal (diatopic, in Coseriu's terms) and idiolectal variation in the lexical and morpho-phonological domains while revealing that the traditional distinction between basilectal (for the leeward islands) and acrolectal (for the windward islands) varieties is not always upheld. I will show to what extent this Swadesh list (collected in a classroom setting in Cabo Verde islands (Praia, Santiago))[4] reveals variation within and across the five islands under study; I will also discuss the limitations of this methodology and how such limitations could be rectified in future work.

This paper is organized as follows. In section 2, I present a brief history of all five islands showing that distinct periods of settlement involving different populations may be the source of the observed variation. In section 3, I elaborate on the notion of variation discussing its multiple facets and justifying the particular focus on regional dialectal and idiolectal variation that this paper adopts. In section 4, I discuss the objectives of Swadesh lists and introduce the methodology I used to collect the present one. In section 5, I present the Swadesh list and analyze the lexical and morpho-phological variation it displays within each group of islands and across clusters. In section 6, I summarize the key findings as well as the limitations of the current study.

2. A brief historical sketch of the five islands under study[5]

2.1. The Sotavento, leeward islands: Santiago, Fogo, and Brava

Santiago was officially discovered by António de Noli, a Genoese nobleman and navigator, and was the first island to be settled in 1460 under the reign of Afonso V of Portugal. Ideally located, Santiago is at the cross-roads of Africa

[4] I am deeply indebted to the Masters' in Creolistics first cohort of students who attended the Field Methods course I taught at the University of Cape Verde in December 2011. I would not have been able to compile this Swadesh list without their valuable contribution and insightful discussion. Their names are listed in the appendix at the end of this paper.

[5] Most of the information in this section is taken from Requedaz (1999) and Andrade (1996).

(Guinea) and Europe (Portugal). It is viewed as the most "African" of all islands due to the slave trade and its commerce with the African continent (Requedaz, 1999). Andrade (1996) (informed by Brásio 1962) states that the majority of slaves who arrived in Santiago came from Guinea and were composed of Mandingas, Balantas, Bijagos, Feloupes, Beafadas, Pepels, Quissis, Brames, Banhuns, Peuls, Jalofos, Bambaras, Bololas, and Manjakus; the regions of Cacheu and Bissau in Guinea supplied most of the slaves (Andrade 1996: 40). The slaves were first brought to cultivate cotton (a labor-intensive crop) before cotton gave way to cattle raising which became the pivot of commercial transactions. The island of Fogo was discovered at the same time as the island of Santiago in May 1460 and was rapidly settled with slaves (presumably the same slave populations as in Santiago) to increase agricultural production (Requedaz 1999). As for Brava, the Portuguese set foot for the first time on the island on June 24 1462 and originally called it São João Baptista. However, the first inhabitants started to settle only in 1573 and their numbers increased dramatically in the 17th century after the eruption of the volcano in Chã das Caldeiras in Fogo forced many Fogo inhabitants to take refuge in Brava in 1680 (Requedaz 1999).

2.2. The Barlavento, windward islands: Santo Antão and São Vicente

The Portuguese Diogo Afonso first discovered the island of Santo Antão in 1462 although some archeological findings point to earlier inhabitants, as claimed by Portuguese historian António Carreira. The first settlers of Santo Antão were sent by the Portuguese Infante Henrique and originated from Algarve and Alentejano in Portugal. The island's excellent climate promoted a flourishing agriculture which necessitated slave labor. The climate and fertile soil of the island soon after attracted other settlers, including Spaniards, Italians and French individuals who mixed with the local population and gave rise to a genuine creole society. In 1548, settlers from Madeira and Portuguese Jews also arrived, as attested by toponyms like Sinagoga and several Jewish cemetaries in towns like Ponta do Sol (Requedaz 1999). While the first wave of settlement is marked by the arrival of these exogenous populations, the second wave starting at the end of the 18th century through the 19th century is characterized by arrivals of endogenous populations coming from the other islands of the arquipelago (Requedaz 1999).

São Vicente was discovered like Santo Antão, its next-door neighbor, in 1462 but in contrast to Santo Antão, the island was only settled in the middle of the 19th century due to a poor aquifer, hence limited water supply. Prior attempts

at settlements occurred first in 1795 with a land owner, a few slaves and white families and then in 1820 with a population of 300 individuals composed of a land owner, some slaves and free men but both attempts failed due to harsh life conditions on the island. It is only in June 1838 that the settlement took off with the Minister of Colonies signing a decree leading to the creation of Mindelo which was supposed to be the capital of Cabo Verde and which officially became a city in 1879 (Requedaz 1999; Swolkien 2015)). Mindelo attracted a number of foreign powers such as the Netherlands that regularly stopped there for supplies on their way to Brazil but no country had as much influence on the economic development of São Vicente as England which developed close economic ties with Portugal and set up manufacturing infrastructures in 1790. Such initiatives greatly benefited the prosperity of the island. England also built a coal plant in 1838 to supply ships that were crossing the Atlantic. Such economic developments attracted populations from neighboring islands such as São Nicolau and Santo Antão, and from abroad, mainly Portugal, England and Italy. These three European countries have left their imprint on the population of São Vicente, considered to be one of the whitest islands of the archipelago. The English language in particular has passed on numerous lexical items to the variety of *Kriolu* spoken in São Vicente. On this topic, Delgado (2008) has observed that the back and forth travels of Cabo Verdean immigrants working in the United States have brought back to the windward islands a plethora of English words such as *bróda* for "brother", *kofebreke* for "coffee break", *orkxope* for "workshop" (see Delgado 2008: 68 for an extensive list).

The original settlement dates with the origins of the populations for the Sotavento, leeward and Barlavento, windward islands are summarized in Table 1. This brief historical sketch highlights that although all five islands were discovered around the same time (1460-1462), their periods of settlement and the populations involved in such settlements vary dramatically, giving rise (as we shall see) to distinct varieties of the *Kriolu* language. The variation between these islands will be explored through the lens of a Swadesh list in section 5 but first, in the next section, I discuss the notion of variation and highlight which dimensions of variation I privilege in this paper and why.

Table 1: Dates of settlement and population origins

Dates and population origins	Santiago	Fogo	Brava	Santo Antão	São Vicente
Discovery date	1460	1460	1462	1462	1462
Date of first major settlement	1460	1460	1573	1462	1838
Founding Populations	Portuguese Mandingas, Balantas, Bijagos, Feloupes, Beafadas, Pepels, Quissis, Brames, Banhuns, Peuls, Jalofos, Bambaras, Bololas Manjakus	Portuguese Mandingas, Balantas, Bijagos, Feloupes, Beafadas, Pepels, Quissis, Brames, Banhuns, Peuls, Jalofos, Bambaras, Bololas Manjakus	Portu-guese and Slaves	Portuguese from Algarve, Alentejano and Madeira Slaves Spaniards Italians French Portuguese Jews	Portuguese English Italians Neighboring islands

3. Unpacking the notion of variation: a focus on diatopic, basilectal, acrolectal and idiolectal variation

Not all types of variation are created equal, particularly when examining oral languages like creoles that often times escape the subjugation of language standardization. The motivation behind the use of some variants over others can differ from one speech community to the next, and from one speaker to the next. Some variation can be viewed as optional with no particular significance attached to such optionality; in contrast, other types of variation are viewed as indexing social meaning or as carriers of social messages that have been represented in the literature as "acts of identity". LePage & Tabouret Keller (1985) are among the first scholars to argue that variation can index social meaning; in their view, "language is the focal centre of our acts of identity" (Le Page & Tabouret-Keller 1985: 248). In their pioneering study, they use a multidimensial lens to account for language variation, arguing that all linguistic tokens are socially marked in the sense that an individual uses them because such tokens are felt to convey a particular social and semantic meaning pertinent to how the individual wishes to position themselves in the world around them (Le Page & Tabouret-Keller 1985: 248). Le Page & Tabouret-Keller conclude that national, ethnic, racial, cultural, religious, age, sex, social class, caste,

educational, economic, geographical, occupational and other groups are all associated with specific linguistic connotations. In other words, the way a speaker talks may index a particular identity that intersects with all such groupings, making language variation multilayered and deeply complex. Le Page & Tabouret-Keller were also among the first to correlate acts of identity with creoles, proposing that such languages do not evolve along a linear continuum punctuated by the three major lects traditionally labeled basilect, acrolect and mesolect. Their definition of the basilect is a lect that is characterized by features that are linguistically closer to those of the founding populations' original native languages; these native languages typically (but not always) disappear from the linguistic ecology in which creoles evolve but leave their imprint on the creole grammar via substratal transfer and other processes like convergence/ conflation (Kihm 1994; Kouwenberg 2000; Baptista 2006). The acrolect is typically used by speakers with access to the lexifier language via education. Such exposure is believed to imbibe their lect with more lexifier-like features. Instead of a linear continuum, Le Page & Tabouret-Keller propose a multidimensional model highlighting complex social rules that enable speakers to switch and mix items from two or more codes. We return to this point in section 5 where we show that the analysis of the Swadesh list supports Le Page & Tabouret-Keller's claim that individuals tend to mix items from two or more codes (lects).

The notions of basilect and acrolect are particularly relevant to the ways in which the Cabo Verdean language has been described in past works. The early settlement of a Sotavento, leeward island like Santiago and the late settlement of Barlavento windward islands like São Vicente have created the enduring perception that leeward (Sotavento) varieties are basilectal, hence closer to the original (African) source languages whereas the windward (Barlavento) islands are more acrolectal, hence closer to the lexifier. Their distinct histories have put them on opposite poles of the creole continuum. Indeed, the literature on Cabo Verdean Creole has consistently argued for instance that substratal input can be detected on the grammatical make up of Sotavento varieties whereas for Barlavento varieties, the influence of Portuguese has been emphasized. Historians, anthropologists and linguists have all contributed to perpetuating such perceptions. For instance, the historian Elise Andrade writes:

> Dans les îles du groupe Sotavento (Santiago, Fogo, Brava et Maio), de colonisation plus ancienne, l'influence des langues africaines, dans la constitution du créole, est plus marquante qu'à Barlavento (São Vicente, Santo Antão, São Nicolau, Boavista and Sal), de colonisation postérieure. Cela se manifeste tant dans le domaine phonique… que dans

> le domaine lexical où on retrouve un plus grand nombre de mots d'origine africaine.[6] (Andrade 1996: 56)

The same characterization of Sotavento islands as being basilectal and of Barlavento islands as acrolectal is found in linguists' writings, as attested by the following paragraph:

> Due to settlement history, the southern Sotavento varieties not only crystallized earlier but they also show more substrate influence and are more basilectal on a hypothetic Creole continuum than the northern Barlavento varieties. (Bartens and Sandström 2008: 307)

However, when we analyze the Swadesh list we compiled in section 5, we will see that the traditional characterization of Sotavento versus Barlavento corresponding to basilectal and acrolectal varieties does not hold up upon close scrutiny of the elicited lexical items. We show that Le Page & Tabouret-Keller's proposal that speakers consistently mix lects rather than confining themselves to one point of the creole continuum is supported by the empirical evidence found analyzed in section 5. Another linguist's framework that proves quite useful to the analysis offered in section 5 is Coseriu's[7] three-dimensional model of diasystematic variation: the three dimensions involve diatopic, diastratic and diaphasic variation. Indeed, according to Coseriu, the development of any language can be best described by taking into account the fundamental distinction between written and spoken language which cannot be reduced to diasystematic differences. Thus, varieties with stronger diatopic marking (as in the case of dialects and regional varieties), as well as diastratically 'low' varieties (vernaculars for instance), are closely related to spoken language. From a diaphasic point of view, 'low' registers (familiar) have been associated and identified with spoken language and colloquial speech.

This paper wishes to emphasize three main dimensions of variation: 1) variation across social lects encompassing basilects and acrolects (Le Page & Tabouret-Keller 1985), 2) diatopic variation (dialectal regional variation, as defined in Coseriu 1981) and 3) idiolectal variation. The relevance of this third dimension did not escape linguists like Martinet who in his preface of Weinreich's (1964) volume, emphasized the importance of variation in the

[6] An approximate translation would be: "In the Sotavento islands (Santiago, Fogo, Brava and Maio), settled much earlier, the influences of African languages in the genesis of the creole are more dominant than in Barlavento (São Vicente, Santo Antão, São Nicolau, Boavista and Sal), settled later. Such influences are felt not only in the phonological domain… but also in the lexical domain where one finds a great number of words with an African etymology".

[7] I thank an anonymous reviewer for providing me with such a valuable reference.

individual by stating: "linguistic diversity begins […] within one and the same man. It is not enough to point out that each individual is a battle-field for conflicting linguistic types and habits and, at the same time a permanent source of linguistic interference. […] 'A language' is the aggregate of millions of such microcosms many of which evince such aberrant linguistic comportment that the question arises whether they should not be grouped into other 'languages'. (Martinet's preface, in Weinreich 1964: vii). The importance of the idiolect as a primordial site of variation is summarized eloquently in Mufwene (2001):

> "The variation that **matters** [my emphasis] to evolution really begins at that interidiolectal level, before reaching the next higher level of cross-dialect and/or cross-language differences. Contact, which has been dealt with primarily at the level of dialects or languages, really begins at this level of idiolects. Since the locus of dialect or language contact is the mind of the individual speaker, the difference between idiolect contact and language contact or dialect contact is more quantitative than qualitative" (Mufwene 2001: 149-150).

While Mufwene views the idiolect as the initial point of variation, other scholars view it as the basic atom of language. For instance, Oksaar (2000) argues that languages are ultimately aggregates of idiolects in the sense that all forms of language change originate from individual realizations of language systems. As such, he views idiolects as the basic unit of the continuum of variation along which sociolects and dialects are collective entities. Oksaar's dynamic synchronic methodology approaches idiolects in the linguistic dimension of social variation and the social dimension of linguistic variation; the unit of observation is the communicative act (Oksaar 2000).

In section 5, I examine a Cabo Verdean Swadesh list and propose that Swadesh lists, when elicited from multiple individuals, can be used as optimal tools to study diatopic/regional dialectal (between leeward and windward islands), sociolectal (basilect versus dialectal) and idiolectal variation. I emphasize these three dimensions of variation in my analysis of a Swadesh list collected in Cabo Verde (section 5) but first define in the next section what a Swadesh list is and the methodology I used to compile the Swadesh list under study in this paper.

4. On the purpose of Swadesh lists and methodological approach

4.1. The purpose of Swadesh lists

The Swadesh list is a compilation of basic every day common words that are

found in most natural languages. The Swadesh list has proved a useful tool for historical comparative linguistics by revealing to what extent the languages being compared relate to each other. This list, originally designed by historical linguist Morris Swadesh, has been applied to lexicostatistics which quantitatively assess whether a set of languages are genealogically related and (within lexicostatistics) to glottochronology which uses statistical data to gauge when a set of languages may have diverged from their source languages. Over time, several different Swadesh lists have emerged that differ in length and contents, as a result, some scholars refer to "Swadesh lists" in the plural instead of in the singular.

The Swadesh list under study does not apply to distinct languages but instead to distinct varieties of the same language. As such, the purpose of this Swadesh list is not about establishing whether the varieties spoken in the leeward and windward varieties are inter-related because we already know that they are. Instead, the objective is to bring forth the acute linguistic variation known to exist between the two groups of islands, due in part to the different periods of settlement in which each island was settled and in part due to the different populations who were part of such settlements, as discussed in section 2. As mentioned earlier, the three leeward varieties are Santiago, Brava and Fogo and the two windward varieties are São Vicente and Santo Antão. The distinctive histories and periods of settlement of the leeward (i.e., 1460 for Santiago) and the windward islands (i.e., 1838 for São Vicente) place these two clusters at two opposite poles of the linguistic continuum. The leeward islands are believed to represent the basilectal side and the windward islands the acrolectal side of the continuum. However, as we see below, the Swadesh list under study does not support this divide in all cases (also see Baptista 2015).

4.2. Method of collection

The Swadesh list under study was developed on the basis of chapter 3 "Basic Lexicography" in Bert Vaux & Justin Cooper's (1999) book entitled *Introduction to Linguistic Field Methods.* The full Swadesh list can be found in the appendix of this paper but for reasons of space, we are limiting ourselves to the examination of a list of 50 words.

The Swadesh list of 50 words in Table 2 below was collected from Cabo Verdean students who were all native speakers of Cabo Verdean Creole and who were representative of the five islands mentioned above and listed in Table 1. The other four islands of Cabo Verde, namely Boa Vista, Sal, São Nicolau

and Maio, are not represented in the Swadesh list due to the absence of native speakers of these varieties in the classroom. The registered students were in the first cohort of the Masters' in Creolistics which was offered at the University of Cabo Verde in 2010. They were all enrolled in the course on Field Methods for which I was the instructor. One of the course objectives was to study methods of data collection that would allow the class to document variation in the Cabo Verdean language. For each lexical entry of the Swadesh list, students who were native speakers of a given variety were invited to provide the word that in their view, corresponded the most closely to the concept being elicited. All students who were speakers of the targeted variety were encouraged to contribute to the Swadesh list and to vet each other's response. This resulted in the fact that some concepts only have one lexical entry whereas other concepts have several entries, as students speaking the same variety could conjure up different lexical entries that in their view corresponded to the same concept. As students vetted each other's responses, the suggestions that went unopposed were listed under the same lexical entry. The various suggestions were viewed as the output of different idiolects present in the classroom. All the words in Table 2 below have been numbered for ease of reference.

5. Variation in the lexicon: what the study of Swadesh lists can reveal

The objective of this section is to consider three types of variation – diatopic /regional, basilectal vs. acrolectal and idiolectal variation – through the lens of a Swadesh list elicited from native speakers of Cabo Verdean.

The original intent behind the compilation of a Swadesh list as a classroom assignment was to bring forth the dramatic dialectal variation known to exist in Cabo Verdean creole. Notable grammatical descriptions have revealed the extent of the variation at the phonological, morphological, lexical and syntactic levels between the Sotavento and Barlavento islands, and within each group (Veiga, 2000; Quint, 2000; Lang, 2009; Swolkien, 2015; Delgado, 2008, among many others).

In this section, I only go over some of the most salient phonological features differentiating the two clusters of islands. In this respect, Veiga (2000) and Swolkien (2015) are among the first studies to compare Santiago and São Vicente, as representative of the Sotavento and Barlavento clusters, respectively. With respect to phonology, Swolkien (2015) makes important observations about distinctive features between Santiago and São Vicente. For instance, Swolkien (2015: 30) notes (as Veiga 2000: 69) that unstressed final

vowels /u/ and /i/ are often reduced or deleted in the São Vicente variety but maintained in the Santiago variety, causing significant differences in phonological and syllable structures between the two varieties. To be more specific, Swolkien observes that when word-final unstressed vowels like /u/ and /i/ in Santiago gat**u** [ˈgatu] 'cat', póbr**i** [ˈpɔbri] 'poor', xatiad**u** [ʃɐˈtjadu] 'irritated' are deleted in the São Vicente variety, such a deletion makes the preceding stressed vowels in the São Vicente variety higher and more rounded, turning into gót [ˈgɔt] 'cat', pobr [ˈpobɾ] 'poor', xatiód [ʃɐˈtjɔd] 'irritated'. In this deletion resides one the most striking differences between the Sotavento and the Barlavento varieties.

Other distinctive features include the voiced palatal fricative /ʒ/ <j> in the São Vicente variety that finds its counterpart in the Sotavento affricate /dʒ/, as in djánta [ˈdʒantɐ] for Santiago and jantá [ʒɐ̃ˈta] 'to have supper' in São Vicente (Swolkien 2015: 35). With respect to affricates /tʃ/ <tx> and /dʒ/ <dj>, Quint (2000: 110) postulates that such affricates in Santiago have been inherited from Old Portuguese voiceless post-alveolar affricate /tʃ/ (we return to this point below). Swolkien observes that it is, however, possible to find this Sotavento feature in São Vicente where both the voiceless palatal fricative /ʃ/ as in xuva [ˈʃu.vɐ] and the affricate txuva [ˈtʃu.vɐ], 'rain' can occur in free variation. Swolkien also notes that the voiced post-alveolar affricate /dʒ/ that is found in words like pádja [ˈpadʒɐ] 'straw' in Santiago is iotacized as páia [ˈpajɐ] in São Vicente. Interestingly, she remarks that the affricate is kept in São Vicente core vocabulary for words like am**dj**er [ɐmˈdʒeɾ] 'woman', be**dj**a [ˈbedʒɐ] old (feminine), mi**dj** [ˈmidʒ] 'corn', fi**dj** [ˈfidʒ] 'son, child'. Note that the São Vicente word a**mdj**er 'woman' allows for a consonantal cluster where the Santiago variety would not have one, mu**dj**er. A final phoneme that I highlight here, based on Swolkien's comparative analysis, is the preservation of the Portuguese etymological plosives such as no**b**u 'young' or **b**izinhu 'neighbor' in Santiago whereas São Vicente adopts the modern standard European Portuguese voiced labio-dental fricative as in no**v** [ˈnov] 'young' and **v**zinh [ˈvzi ɲ] 'neighbor'(Swolkien 2015: 35; see Santos Baptista 2014: 212 for Santo Antão). On this very topic, Quint (2000: 113) suggests that Santiago plosives could also be possibly inherited from African substrates. Although Quint does not deny a possible transfer from north-Portuguese dialects, he weighs in favor of the African substrates because substrates like Mandinga and Wolof do not have fricatives which would have prompted the early slaves to use plosives in their place.

The aforementioned distinctive features differentiating Sotavento from Barlavento can of course be found in the Swadesh list under study but not always in the expected variety, as we shall see below.

Although the intent behind the compilation of the Swadesh list was to reveal primarily diatopic/regional variation between the five islands, the analysis of the lexical items that were elicited in the classroom yielded several observations.[8]

First, there are words that show no variation and remain constant across the varieties and crucially, such words only have a single form and no competitors. They belong to a very small minority of cases and in the entire Swadesh list of 50 words, only 4 lexical items pertain to that category: (23) *konta* 'to tell/narrate', (25) *dia* 'day', (16) *sopra* 'to blow', and (44) *pena* 'feather'.

Second, and this is important to our discussion about basilectal and acrolectal variation, there are words that do not support the traditional regional divide between leeward and windward islands, as being reflective of the basilectal and acrolectal varieties, respectively. For instance, *mre* 'to die' from Portuguese *morrer* in (26) displays a consonantal cluster typical of windward islands (as reported in Swolkien 2015) but is only present in the windward island of SA and not in its windward neighbor SV which instead shares the form *more* with the leeward islands BR, FO and ST. With respect to affricates, as stated by Swolkien (2015), it is often the case that the Santiago affricates are iotacized in the Barlavento varieties except in high frequency words like *amdjer* 'woman' in SV, as seen in (50). In the case of (50), SV aligns with the Sotavento varieties ST, BR and FO but diverges from its Barlavento neighbor SA which uses the expected iotacization with *miêr*. The same discrepancy with respect to the traditional windward and leeward divide can be observed with *bebe* 'to drink' from Portuguese *beber* in (30). In this particular instance, the windward island of SA and the leeward island of BR share the same form *bebe* whereas the windward island of SV and the leeward island of FO share the form *bibe*. One should also note that in this case, each of the leeward islands has its own variant: *bebi* in ST (also see Quint 2000: 27), *bebe* in BR and *bibe* in FO. As for the *v* vs. *b* alternation, as discussed above, Swolkien (2015), Veiga (2000) and Quint (2000) have noted that *b* is traditionally associated with the leeward islands and may have been preserved in ST due to African substrates that lack the fricative *v*, (see Quint 2000) while *v* is associated with the windward islands. As a result,

[8] The island varieties presented in the Swadesh list carry the following abbreviations: Santiago (ST), Fogo (FO), Brava (BR), Santo Antão (SA) and São Vicente (SV).

the use of the plosive is typically viewed as basilectal whereas the use of the fricative is viewed as acrolectal. However, in this case as well, the divide between basilectal (associated with leeward islands) and acrolectal (associated with windward islands) is not maintained. For instance, *v* in (20) is preserved in both windward *nu**v**en* (from Portuguese *nuvem*) in SV and *ne**v**a* in SA and leeward ST and FO *nu**v**en*. In contrast, one of the ST variants *nu**b**ri* and BR *nu**b**en* use the plosive as expected. Similar to the previous case, in (37), *o**b**u* and *o**v**u* 'egg' compete in ST whereas only *o**b**u* is present in FO and BR.

In contrast, in (27), the divide between *b* for leeward and *v* for windward is maintained with *ko**b**a* 'to dig' being used in all leeward islands as expected, and *ko**v**a* in all windward islands.

Third, there are in addition to (27) other words for which the leeward and windward divide is clearly maintained. Such is the case of (34) where the affricate in the word *ore**dj**a* 'ear' is shared by all leeward islands and the iotacized *or**ea*** by the windward islands. The same goes for (36) where the consonantal cluster in ***km**e* is shared by windward SA and SV vs. *kume* (lack of cluster) by leeward ST, BR and BR. The same clean divide with respect to the maintenance of unstressed final vowels in Sotavento while they are dropped in Barlavento is displayed in the word list. This feature is illustrated in (31) by *sek* 'dry' for leeward islands vs. *sek**u*** for windward, in (32) by *seg* 'blind' for leeward islands vs. *seg**u*** for windward islands, in (37) by *o**v*** 'egg' for leeward vs. *o**bu*** for windward, in (38) by *oi* 'eye' for windward islands vs. *odj**u*** for leeward islands, in (12) by *pos* 'bird' for windward islands vs. *pas**u*** for leeward islands and finally in (19) by *mnin* 'child' for windward islands vs. *minin**u*** for leeward islands.

Fourth, there are words that are constant across varieties but that coexist with competitors. For instance *ben* 'to come' in (22) is constant across all five varieties in that exact same form but competes with *txiga/txega* in all five varieties. *Korta* in (24) is constant across all varieties but competes with the form *sapa* in ST, SV, FO, and BR, and with the variant *tsapa* in SA. *Korta* also competes with *pika* in ST, SA and FO. *Tera* in (33) is constant as the same form across all five varieties but competes with *po* in ST, SA, and SV, with *puera* in SA, SV, FO, and BR, and with *poera* in ST. For some speakers, the meaning of *tera*, *puera* and *po* are very distinct from each other whereas for others, they are similar. This is a case that points to idiolectal variation (addressed further below) while also challenging the Barlavento (windward) and Sotavento (leeward) divide discussed above, as *po* was proposed by the students as a variant of leeward ST and windward SA and SV but not of the other leeward islands of FO and BR.

Fifth, one can observe that some varieties have multiple variants for the same word as in (9) where ST displays the competitors *pamo, pamodi, pabia, parabia, parabias, purkauza* and *purkazu* 'because' whereas all the other varieties have a single form for the same word. In this particular case, all the variants are semantically equivalent. In contrast, lexemes in (7) for 'bad' are not semantically equivalent and yet were elicited from the students. For instance *mau* and *malvadu* in (7) can be translated by 'bad' but *kasabi* conveys instead a feeling of sadness or bitterness, despair (Brüser *et al.* 2002: 301). In other cases, students provided lexemes with distinct etymological roots, as in the case of (36) in which *kumi/kume* 'to eat' in ST is derived from Portuguese whereas *nhemi* (closer to the meaning of 'to chew') is derived from Mandinga and Bambara *ñeme* (Brüser *et al.* 2002: 486).

Sixth, the Swadesh list also displays vocalic alternation in the same environment in the sense that distinct vowels can be used in the same phonological environment, depending on the varieties. Observe the following alternations:

- /i/ /u/ alternation in *l**i**mária* or *l**u**mária* 'animal' in (3), attested in ST (also see Brüser *et al.*, 2002: 394). /e/ to /i/ in ST *mord**e*** 'to bite' in SA, FO, BR, and SV, and *mord**i*** in ST in (13) (derived from Portuguese *mord**e**r*).
- /a/ /e/ alternation in *b**a**riga* 'belly' in (10) in the leeward islands contrasting with *b**e**rig* in SA and *b**e**riga* in SV, derived from Portuguese *b**a**rriga*.
- /a/ /o/ alternation in *k**a**va* 'to dig' in SV (27) and *k**o**va* in SA, *k**o**ba* in the leeward islands, derived from Portuguese *c**a**var*.

Finally, it is worth noting that while vocalic and consonantal elision is rampant in the elicited Swadesh list, vowel insertion or epenthesis, otherwise a well-known phenomenon in varieties such as ST (Portuguese *criar* 'to raise' > ST *kiria* (see Quint 2000: 17) was not produced by the students. This could be due to the fact that vowel epenthesis is more likely to be found in the rural areas of ST and in its older speakers but only a sociolinguistic analysis could support this assumption.

The words listed under the third, fourth, fifth and sixth observations above seem to point to idiolectal variation, which reflects both diastractic and diaphasic variation in spoken language (Coseriu 1981) and would account for the presence of distinct phonological variants for the same word or multiple lexical variants for the same concept.

These data elicited from the students in the field methods course point to a few generalizations:

1) There is intra-island variation as well as an acute inter-island variation.
2) In some cases, the observed variation defies the traditional divide between basilectal leeward and acrolectal windward islands as some islands can display forms overlapping with forms outside of their geographic in-group peers.
3) In other cases, the divide between leeward and windward islands is clearly maintained.
4) The observed variation may display competitors that are semantically equivalent or synonymous (as in (9)). Other times, the competitors that were elicited from the students were semantically distinct (as in (7)), pointing to the importance of idiolectal variation.

The analysis of the Swadesh list in Table 2 clearly reveals that the traditional characterization of Sotavento versus Barlavento corresponding to basilectal and acrolectal varieties (see Andrade 1996; Bartens & Sandström 2008), does not hold up upon close scrutiny of the elicited lexical items. Instead, we show that Le Page & Tabouret-Keller's proposal that speakers consistently mix lects rather than confining themselves to one point of the creole continuum is supported by the empirical evidence found in the Swadesh list. Coseriu's (1981) three-dimensional model of diasystematic variation is also validated: the three dimensions involving diatopic (regional), diastratic and diaphasic (spoken, oral language) variation illustrate that the development of any language can be best described by taking into account the fundamental distinction between written and spoken language which cannot be reduced to diasystematic differences. This ultimately points to the importance of the idiolect as a crucial site of variation (Mufwene 2001). It is by taking into account variation at the inter-idiolectal level (Mufwene, 2001) that one can best account for the way distinct vowels can alternate in free variation in the same environment (see (3)), lexemes with *distinct* semantics can be mapped onto the same concept (see (33)) whereas in other cases lexemes with the *same* semantics can be mapped onto the same notion (see (9)). This also accounts for why the most salient features that are supposed to characterize and differentiate Sotavento from Barlavento islands can hardly be confined to the geographical boundaries that they have meant to define.

For all these reasons, this paper has focused on three main dimensions of variation: 1) variation across social lects encompassing basilects and acrolects (Le Page & Tabouret-Keller 1985), 2) diatopic, diastractic and diaphasic variation (as defined in Coseriu 1981) and 3) idiolectal variation (as defined in Weinreich 1964; Mufwene 2001).

Table 2: Swadesh list under analysis

Santiago	Fogo	Brava	Santo Antão	São Vicente
(1) tudu, moku, fepu sima sta “all”	tudu, moku, fepu	tudu, moku, fepu	tud	tud
(2) ku, y “and”	y, ku	y, ku	y, ma	y, ma
(3) limaria, lumaria, bitxu, animal “animal”	limaria	limaria, bitxu	nimal, animal, bitx, elmara	bitx, animal
(4) sinza, sinsa “ash”	sinza	sinza	sinz	sinza
(5) na, pa “in”	na, pa	na, pa	ne, pe	na, pa
(6) kosta “back”	kosta	kosta	koxta, tras	koxta
(7) kasabi, kadretu, mariadu, mau, runhu, malvadu, kabala, mufinu, bidjaku “bad”	mau, kasabi, mariadu, mofinu, fedi	mau, mariadu, run	mariod, mofnod	mau, mariod
(8) kaska di po, laska, fiasa di po “trunk”	kaska di po	kaska di po	rom	rom
(9) pamo, pamodi, pabia, parabia, parabias, purkauza, purkazu “because”	pamodi	pamodi	mod	purke
(10) bariga, ventri “belly”	barriga	bariga	berig	beriga
(11) gordu, grandi, tamanhu, gros “fat”	gordu, grandi, forti	gordu, grandi, forti	grand, tmonhon	grande
(12) pasu, avi, txota “bird”	pasu	pasu	pos	pos
(13) mordi “to bite”	morde	morde	morde	morde, mukna
(14) pretu, negru, sukuru “black”	pretu	pretu	pret	pret
(15) sangi “blood”	sangi	sangi	seng	sang
(16) sopra “blow”	sopra	sopra	sopra	sopra
(17) osu “bone”	osu	osu	os	os
(18) kema, txema “burn”	kema	kema	kema, psi	kema, txumska
(19) mininu, kriansa “child”	mininu	mininu, minizu	mnin	mnin
(20) nubri, nuven “cloud”	nuven	nuben	neva	nuven
(21) friu “cold”	friu	friu	fri	friu
(22) ben, txiga “come/ arrive”	txiga, ben	txiga, ben	ben, txega	ben, txega
(23) konta “to tell/narrate”	konta	konta	konta	konta
(24) sapa, korta, pika “to cut”	korta, sapa, pika	korta, sapa	tsapa, korta, pika	korta, sapa
(25) dia, dia “day”	dia	dia	dia	dia
(26) more “to die”	more	more	mre	more

(27) koba “to dig”	koba	koba	kova	kava
(28) xuxu, suxu, suju, mariadu “dirty”	suju	suju	xuj, kerdid, inkorod	suj, xuj, inkardid
(29) katxor “dog”	katxor	katxor	katxor	kotxor
(30) bebi “to drink”	bibe	bebe	bebe	bibe
(31) seku “dry”	seku	seku	sek, storod	sek
(32) segu “blind”	segu	segu	seg, mok	seg, mok
(33) poera, po, tera “dust”	puera, tera	puera, tera	tera, po, puera	tera, po, puera
(34) oredja “ear”	oredja	oredja	orea	orea
(35) txon, tera “ground”	tera, txon	tera, txon	txon, tera	txon, tera
(36) kumi, kume, nhemi “to eat”	kume	kume	kme	kme
(37) obu, ovu “egg”	obu	obu	ov	ov
(38) odju “eye”	odju	odju	oi	oi
(39) kai, kei “to fall”	kai	kai	kei	kei
(4) lonji, distanti, afastadu “distant”	lonji	lonji	foxtod, lonj	lonj
(41) gordura, gurdura, seti, untu, banha “fat/grease”	gurdura, banha	gurdura, banha	gurdura, banha	gurdura, banha
(42) pai, papa, nha grandi “father”	papa, pai, nha pa	papa, pai	pe	pai
(43) xinti medu, tene medu, meda “to fear”	ten medu	ten medu	treme di med, te k’med, treme d’med	ten med
(44) pena “feather”	pena	pena	pena	pena
(45) poku, pitada, faiska, kusinha “little”	poku, kuzinha, bokadinhu	poku, kuzinha, bokadinhu	pok, kzukuk, kzinha, bkedin	pok, kzinha, bkedin
(46) briga, luta, gera “to fight”	briga	briga	briga, gera	briga
(47) lumi, fogu “fire”	lumi	lumi	lum	lum
(48) pexi “fish”	pexi	pexi	pex	pex
(49) sinku “five”	sinku	sinku	sink	sink
(50) mudjer, mudjei, mudjel “woman”	miêr, mlhêr	amdjêr, mdjêr	mudjê	mudjér

6. Concluding remarks

The purpose of the analysis of this Swadesh list is to bring to light the complexities of studying variation and of accounting for the competing variants in speakers’ idiolects. I show in this section that a close inspection of a Swadesh list reveals that the traditional divide between acrolect and basilect on the one

hand and between leeward and windward varieties islands on the other does not reveal well-defined boundaries across the lects/varieties, except to demonstrate the dramatic variation that occurs within the same oral language.

The original premise of the data collection in this field methods course was to bring to the fore the varieties spoken on the various islands who were represented by students in the class. As already discussed, some of the words only have one variant across the five islands of Cabo Verde whereas others have multiple variants. The variation manifests itself both at the morpho-phonological and semantic levels and for some of the lexemes, some of the variants denote distinct etymologies, attesting to long-term competition between them. As I note in Baptista (under review), it is possible that the observable competition favors the coexistence of multiple variants due to the specific social meaning that some of these variants index in the language.

It is important to note that although the original motivation to collect this Swadesh list was to illustrate dialectal variation between islands (based on the presence in the class of speakers of five varieties), the Swadesh list clearly revealed variation within each island and within the groups of speakers representative of each island. This highlights Mufwene's notion of interidiolects as the loci of language variation and change. Beyond the natural geographic boundaries of each variety, there can be a number of reasons for the observed variation including age of speakers, gender, level of education and sense of identity in the sense that the speech community that a speaker identifies with may lead them to emulate or accommodate to a particular type of pronunciation. This is how each variant can find its "niche".

We hope to have demonstrated in this paper the relevance of Swadesh lists as a tool to unveil complex, multilayered variation that defies clean boundaries between basilectal and acrolectal lects and that certainly defies traditional geographic boundaries. The complexity and multidimensional nature of variation teach linguists a lesson in humility, as in the end individual speakers (and their idiolects) are the ultimate loci of language variation and change.

References

Aboh, Enoch & Michel DeGraff. 2017. A null theory of creole formation based on Universal Grammar. In Ian Roberts (ed.), *The Oxford Handbook of Universal Grammar*, 1-67. Oxford Handbook online.

Andrade, Elise. 1996. *Les Îles du Cap-Vert: De la "Découverte" à l'Indépendance Nationale (1460-1975).* Paris: L'Harmattan.

Baptista, Marlyse (under review). Reflections on Darwin's Natural Selection: A lens on variation, competition and change.

Baptista, Marlyse. 2006. When substrates meet superstrate: The case of Cape Verdean Creole. In Jürgen Lang, John Holm, Jean-Louis Rougé & Maria João Soares (eds.), *Cabo Verde - Origens da sua Sociedade e do seu Crioulo*, 91-116. Tübingen: Gunter Narr Verlag.

Baptista, Marlyse. 2015. Continuum and variation in creole languages: Out of many voices, one language. *Journal of Pidgin and Creole Languages* 30(2). 225-264.

Biberauer, Theresa. 2010. *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge: Cambridge University Press.

Brüser, Martina, André dos Reis Santos, Ekkehard Dengler, Andreas Blum & Jürgen Lang. 2002. *Dicionário do Crioulo da Ilha de Santiago (Cabo Verde)*. Tübingen: Gunter Narr Verlag.

Chomsky, Noam. 2000. *New Horizons in the Study of Language and Mind*. Cambridge: Cambridge University Press.

Coseriu, Eugenio. 1981. La socio- y la etnolingüística: Sus fundamentos y sus tareas. *Anuariu de Letras de la Facultad de Filosofía y Letras. Centro de Lingüística Hispaníca* XIX. 5-30.

Delgado, Carlos Alberto. 2008. *Crioulo de Cabo Verde: Situação Linguística da Zona do Barlavento*. Praia: Instituto da Biblioteca Nacional e do Livro.

Dittmar, Norbert. 1996. Explorations in 'idiolects'. In Robin Sackmann & Monika Budde (eds.), *Theoretical Linguistics and Grammatical Description: Papers in Honour of Hans-Heinrich Lieb*, 109-128. Amsterdam: John Benjamins.

Kihm, Alain. 1990. Conflation as a directive process in creolization. In Norbert Boretzki, Werner Enninger & Thomas Stoltz (eds.), *Beiträge zum 4. Essener Kolloquium*, 111-137. Bochum: Brockmeyer.

Kouwenberg, Silvia. 2001. Convergence and explanations in creole genesis. In Norval Smith & Tonjes Veenstra (eds.), *Creolization and Contact*, 219-239. Amsterdam: John Benjamins.

Lang, Jürgen. 2009. *Les Langues des Autres dans la Créolisation*. Tübingen: Gunter Narr Verlag.

LePage, Robert B. & Andrée Tabouret-Keller. 1985. *Acts of Identity: Creole-based Approaches to Language and Ethnicity*. Cambridge: Cambridge University Press.

Lipka, Leonhard. 1988. Variety is the spice of life: Language variation and sociolinguistics. In *Das Sprachtheoretische Denken*, 317-325. Tübingen: Gunter Narr Verlag.

Mufwene, Salikoko. 2001. *The Ecology of Language Evolution*. Cambridge: Cambridge University Press.

Oksaar, Els. 2000. Idiolect as the basis of variation-oriented linguistics. *Sociolinguistica* 14. 37-41.

Patrick, Peter. 1999. *Urban Jamaican Creole: Variation in the Mesolect*. Amsterdam: John Benjamins.

Quint, Nicolas. 2000. *Grammaire de la Langue Cap-Verdienne.* Paris: L'Harmattan.

Raumolin-Brunberg, Helena. 1988. Variation and historical linguistics: A survey of methods and concepts. *Neuphilologische Mitteilungen* 89(2). 136-154.

Requedaz, Sabrina. 1999. *Cap-Vert*. Genève: Guide Olizane/Découverte.

Santos Baptista, Maria do Céu dos. 2014. Descrição fonológica da variedade da ilha de Santo Antão. In Jürgen Lang (ed.), *A Variação Geográfica do Crioulo Caboverdiano*, 183-250. Erlangen: FAU University Press.

Stell, Gerard 2018. Representing variation in creole continua: A folk linguistic view of language variation in Trinidad. *Journal of English Linguistics* 46(2). 113-139.

Swolkien, Dominika. 2015. *The Cape Verdean Creole of São Vicente: Its genesis and structure*. Coimbra: Universidade de Coimbra, PhD dissertation.

Veiga, Manuel. 2000. *Le Créole du Cap-Vert.* Paris: L'Harmattan.

Weinreich, Uriel. 1964. *Languages in Contact: Findings and Problems.* London: Mouton &Co.

Winford, Donald. 1984. The linguistic variable and syntactic variation in creole continua. *Lingua* 62. 267-288.

APPENDIX

Full Swadesh list collected in 2011 with the first cohort of students enrolled in the Field Methods course I taught for the Masters' in Creolistics offered by the University of Cape Verde.

Table 3: Full Swadesh list collected in Field Methods course I taught in 2011

Santiago	Santo Antão	São Vicente	Fogo	Brava
tudu, moku, fepu sima sta	tud	tud	tudu, moku, fepu	tudu, moku, fepu
ku, y	y, ma	y, ma	y, ku	y, ku
limaria, lumaria, bitxu, animal	nimal, animal, bitx, elmara	bitx, animal	limaria	limaria, bitxu
sinza, sinsa	sinz	sinza	sinza	sinza
na, pa	ne, pe	na, pa	na, pa	na, pa
kosta	koxta, tras	koxta	kosta	kosta
kasabi, kadretu, mariadu, mau, runhu, malvadu, kabala, mufinu, bidjaku	mariod, mofnod	mau, mariod	mau, kasabi, mariadu, mofinu, fedi	mau, mariadu, run
kaska di po, laska, fiasa di po	rom	rom	kaska di po	kaska di po
pamo, pamodi, pabia, parabia, parabias, purkauza, purkazu	mod	purke	pamodi	pamodi
bariga, ventri	berig	beriga	barriga	bariga
gordu, grandi, tamanhu, gros	grand, tmonhon	grande	gordu, grandi, forti	gordu, grandi, forti
pasu, avi, txota	pos	pos	pasu	pasu
mordi	morde	morde, mukna	morde	morde
pretu, negru, sukuru	pret	pret	pretu	pretu
sangi	seng	sang	sangi	sangi
sopra	sopra	sopra	sopra	sopra
osu	os	os	osu	osu

kema, txema	kema, psi	kema, txumska	kema	kema
mininu, kriansa	mnin	mnin	mininu	mininu, minizu
nubri, nuven	neva	nuven	nuven	nuben
friu	fri	friu	friu	friu
ben, txiga	ben, txega	ben, txega	txiga, ben	txiga, ben
konta	konta	konta	konta	konta
sapa, korta, pika	tsapa, korta, pika	korta, sapa	korta, sapa, pika	korta, sapa
dia, dia	dia	dia	dia	dia
more	mre	more	more	more
koba	kova	kava	koba	koba
xuxu, suxu, suju, mariadu	xuj, kerdid, inkorod	suj, xuj, inkardid	suju	suju
katxor	katxor	kotxor	katxor	katxor
bebi	bebe	bibe	bibe	bebe
seku	sek, storod	sek	seku	seku
segu	seg, mok	seg, mok	segu	segu
poera, po, tera	tera, po, puera	tera, po, puera	puera, tera	puera, tera
oredja	orea	orea	oredja	oredja
txon, tera	txon, tera	txon, tera	tera, txon	tera, txon
kumi, kume, nhemi	kme	kme	kume	kume
obu, ovu	ov	ov	obu	obu
odju	oi	oi	odju	odju
kai, kei	kei	kei	kai	kai
lonji, distanti, afastadu	foxtod, lonj	lonj	lonji	lonji
gordura, gurdura, seti, untu, banha	gurdura, banha	gurdura, banha	gurdura, banha	gurdura, banha
pai, papa, nha grandi	pe	pai	papa, pai, nha pa	papa, pai
xinti medu, tene medu, meda	treme di med, te k'med, treme d'med	ten med	ten medu	ten medu
pena	pena	pena	pena	pena
poku, pitada, faiska, kusinha	pok, kzukuk, kzinha, bkedin	pok, kzinha, bkedin	poku, kuzinha, bokadinhu	poku, kuzinha, bokadinhu
briga, luta, gera	briga, gera	briga	briga	briga
lumi, fogu	lum	lum	lumi	lumi
pexi	pex	pex	pexi	pexi
sinku	sink	sink	sinku	sinku
boia	boiá	boiá	boiâ	boiâ
kore	krê	korê	korê	korê
flor	flor	flor	flor	flor
buâ	vuâ	vuâ	buâ	buâ
nébua, négua	neva, nevuer	nevuer	néba	néba
pê	pê	pê	pê	pê
kuatu	kuat	kuat	kuatu	kuatu
jela, konjela	jelá	jelá, konjelá	jelâ	jelâ
fruta	fruta	fruta	fruta	fruta
da	da	da	da	da
bon, fadjadu, dretu	bon, dret	bon, dret	bon, dretu	bon, dretu

padja	mot, paia, erva	mote, paia, erva	erva	erva
berdi, verdi	verd	verd	berdi, verdi	verdi
tripa	tripa	tripa	tripa	tripa
kabelu	kebel	kabel	kabelu	kabelu
mo, mon	mon	mon	mon	mon
e, el, ael	el, e, le	el, e	e, el, ael	e, el, ael
kabésa, kaku, kabê, kukuruta	kukruta, kebésa, kok	kabésa	kabésa	kabésa, kaku
obi	uvi	uvi	obí	ubí
kurason	kroson	korason	kurasan	kurason
pisadu	pezod	pezod	pizadu	pizadu
li, ali	ei	li, ei	li	li
bati, sota, dadji, da, lonba, pila	da, sotá, lonbá	da, xpanká, lonbá	dadji, batê	dadji, batê
pega, susti, sugura	suxtê, guentá, sgerá, pegá	suxtê, guentá, sgerá, pegá	pegá	pegá
modi, mó, kuma	kmenera, mené, mod ken	manera, modi ke, mod, ki manera	moda ki, kumá	moda ki, kuma
kasa, pega	kasá, pegá	kasá, pegá	kasá	kasá
maridu, omi, kunpanheru	merid	merid, kunpanher	maridu	maridu
n	n	n	n	n
jelu	jel	jel	jelu	jelu
si	se	se	si	si
na, dentu	dent	dent	na, dentu	na, dentu
tra, mata	matá	matá	matá	matá
konxe, sabe	kunsé, sebé	konxê, sabê	konxê, sabê	konxê, sabê
laguâ	lagoa, pósa	pósa	lagoa	lagoa
ri, ratxa grasa	erí,	erí	ari, ri	ari
fodja, folha	fólha	fólha	fódja	fódja
skérda	skerda (xkerda)	skerda	skérda	skérda
pé	pérna	pérna	pé	pérna
detadu di ladu	detód d´ lód	detód d´lód	detadu di ladu	detadu di ladu
mora, vive, ta sta	morá, vive, ta stód	morá, vive, ta stód	vivê, morâ, sta	vive, morâ, sta
figu, fígadu	fig	fig, fígd	figu	fígadu
kunpridu, tamanhu	kunprid	kunprid, grande	kunpridu, grandi	kunpridu, grandi
piodju, léndia	pioi	pioi	piodju	piodju
ómi	ome	ome	ómi	omi
txeu, un monti, kotxada	bestent, un data, un larada, un ktxada, un mont, un monzada	txeu, bastant, un data, un mund, un larada, un ktxada, un mont, karada	txeu, monti	txeu, un monti
karni	ker	kar, karn	karni	karni

mai, mamai, mama, nha velha	men	mai, mámá,	má, mámá, mai	mai, mámá
monti	mont	mont	monti	monti
bóka	bóka	bóka	bóka	bóka
nomi	nom	nom	nómi	nómi
stretu, pertadu	stret, pertód, rotxód,	stret, pertód, rotxód,	stretu, pertadu	stretu, pertadu
pertu, djuntu	pert, junt	pert, d´junt	pertu, djuntu	pertu, juntu
piskos, katxu	peskos,	peskos	piskos, katxu	piskos, katxu
novu, nobu	nov	nov	nobu	nobu
noti, sukuru	dnot, skur	dnot, skur	noti, sukuru	noti, sukuru
naris	neris	neris	naris	naris
nau, ka	no,	ka, no	nau, ka	nau, ka
bedju, uzadu, antigu	vei, maltrotód, entig, uzód,	bedj, maltratód, antig, uzód	bédju	bédju
un	un	un	un	un
otu, otru	ot	ot	otu, otru	otu, otru
algen, argen, pesoa	kriatura, psoa,	kriatura, psoa	argen, pesoa	algen, pesoa
djuga, juga	jgá	jgá	djuga	juga
puxa	pxá	pxá	puxa	puxa
pintxa, npura	inpurá, pintxá	inpurá, pintxá	pintxâ	pintxâ
txobe	txvé	da txuva	txobê	txobê
burmedju, brumedju	vermei	brumedj, vermelh, inkarnód	brumedju	brumedju
dretu, sértu	dret	dret, kul	dretu	dretu
mo ndreta	mon dreita	dreita, direita	mosdreta	mo dretu
riu	riu	riu	riu	riu
strada	strada,	strada	strada	strada
raís	reís	reis	raís	raís
kórda	kórda	kórda	kórda	kórad
podri, stragadu	podr, strogód, pudresid	podr, stragód	podri, stragadu	podri, stragadu
frega, ferga, siridja	xuriá, sfregá, reá,	xuriá, sfregá, reá,	fregá,	fregá
sal	sal	sal	sal	sal
areia, riâ	erea	área	areia	areia
fla	dzê, mentá	dzê, menta	fra	fra
kosa	kosá	kosá	kosâ	kosá
mar	mar	mar	mar	mar
odja, djobe	oiá, spiá	oiá, spiá, biziá	odjâ, spiâ, djobé	odjá, spiá, djobé
simenti	sment	sment	simenti	simenti
kozi, kosi	kzé	kozé	kozê	kozé
sabi, moladu, infiadu	molód	molód	moladu, fiadu	moladu, fiadu
pikinoti, kurtu, réngu, baxu, rokotó	pknin, kurt	piknin, kurt	pikinoti, kurtu, baxu	pikinoti, kurtu, baxu
kanta	kantá	kantá	kantâ	kantâ
xinta	sentá	sentá	sintâ	sinta
péli	pel	pel	péli	péli

séu	séu	séu	séu	séu
konxe, sabe	kunsé, sebé	konxê, sabê	konxê, sabê	konxê, sabê
laguâ	lagoa, pósa	pósa	lagoa	lagoa
ri, ratxa grasa	erí,	erí	ari, ri	ari
fodja, folha	fólha	fólha	fódja	fódja
skérda	skerda (xkerda)	skerda	skérda	skérda
pé	pérna	pérna	pé	pérna
detadu di ladu	detód d´ lód	detód d´lód	detadu di ladu	detadu di ladu
mora, vive, ta sta	morá, vive, ta stód	morá, vive, ta stód	vivê, morâ, sta	vive, morâ, sta
figu, fígadu	fig	fig, fígd	figu	fígadu
kunpridu, tamanhu	kunprid	kunprid, grande	kunpridu, grandi	kunpridu, grandi
piodju, léndia	pioi	pioi	piodju	piodju
ómi	ome	ome	ómi	omi
txeu, un monti, kotxada	bestent, un data, un larada, un ktxada, un mont, un monzada	txeu, bastant, un data, un mund, un larada, un ktxada, un mont, karada	txeu, monti	txeu, un monti
karni	ker	kar, karn	karni	karni
mai, mamai, mama, nha velha	men	mai, mámá,	má, mámá, mai	mai, mámá
monti	mont	mont	monti	monti
bóka	bóka	bóka	bóka	bóka
nomi	nom	nom	nómi	nómi
stretu, pertadu	stret, pertód, rotxód,	stret, pertód, rotxód,	stretu, pertadu	stretu, pertadu
pertu, djuntu	pert, junt	pert, d´junt	pertu, djuntu	pertu, juntu
piskos, katxu	peskos,	peskos	piskos, katxu	piskos, katxu
novu, nobu	nov	nov	nobu	nobu
noti, sukuru	dnot, skur	dnot, skur	noti, sukuru	noti, sukuru
naris	neris	neris	naris	naris
nau, ka	no,	ka, no	nau, ka	nau, ka
bedju, uzadu, antigu	vei, maltrotód, entig, uzód,	bedj, maltratód, antig, uzód	bédju	bédju
un	un	un	un	un
durmi, drumi	drmi	durmí	durmí	durmí
pikinoti	peknin	pikinin	pikenu	pikinoti
txera, kera	txeré	txeré	txerâ	txerâ
fuma	fmá	fmá	fumâ	fumâ
lizu, finu	lis	lize, fine	lizu	lizu
kobra	kobra, serpenta	kobra	kobra	kobra
nevi, jelu	jêl, nev	néve, jêl	nevi	nevi, jelu

argun, algun, un bokadu, un kusinha	elgun, un kzuk	elgun, un bokód, un kuzitxitxa	un poku, un bokadu, un kuzinha	algun, un poku, un bokadu, un kuzinha
kuspi	kspí	kuspí	kuspí	guspi
ratxa, parti, kebra	ratxá, kebrá	partí, ratxá, kebrá	ratxâ, partí, kebrâ	ratxâ, partí, kebrâ
spremi	spremê	spremê	spremê	spremê
da ku faka, fakia, soti faka, finka faka	sfkiá	dá d faka, sfakia,	da ku faka	da ku faka
sakedu, firma, argi, labanta	poinpé,	po in pé, levantá, ergi, spendé	sakê, labantâ,	sakê, lebantâ
strela	strela	strela	strela	strela
pó	spêt	spêt, pau	pô	pô
pedra	pédra	pédra, blok	pedra	pedra
retu	rêt	rêt, streid	retu	retu
txupa, mama	txpá, mamá	txpá, mamá	txupa, mamâ	txupa, mamâ
sol	sol	sol	sol	sol
intxa, inflama	intxá, inflamá	intxá, inflamá	intxâ, inflamâ	intxâ, inflamâ
nada	nadá	nadá	nadâ	nadá
rabu,	rób	rób	rabu	rabu
kel, kel la	kel	kel	kel	kel
la,	lá	lá	la	la
es, eris, aes, aeris	ês	ês	es, aes	ezi
grós	grôs	grôs	grós	grosu
miodu, finu	miud, fine	miud, fine	miodu, finu	miodu, finu
pensa, matuta, kuda	pensá, matutá	pensá, matutá	pensâ	pensâ, kuda
kel li, es li, es ali	kel ei	es li	kel li	kel li
trés	três	três	três	três
fulia, bota, da ku el	remesá, fuliá, mandá	remesá, mandá, btá	fuliâ, botâ	botâ
mara, prende, ata	mará	mará	marâ	marâ
língua, luinga	linga	linga	lingua	lingua
denti	dent	dent	denti	denti
arvi, pe di po, po	arv, môt, pê d môt	arv, môt, pê d môt, pe d arv	arvi	arvi
vira, voita	vrá, voltá	vrá, voltá	birá	bira
dos	dos (x)	dos (x)	dos	dos
bota, vomita, tra stangu, gomita, disgumita	vmitá, tra d stóm	vumitá, vmitá, tra d stóm	vumitâ	vomitâ
anda	andá, kaminhá	andá	andâ	anda
sta kalor, kalor sta, sta ku kalor	ta k kalor, ta kent	ta k kalor, ta kent	sta kalor	sta kalor
laba	lavá	lavá	labâ	labâ
agu	aga	aga, agua	agu	agua
nos, nu, anos	nôs, nô	nôs, nô	nos, nu	nos, nu
modjadu, suadu	moiôd	moiôd, suôd	modjadu	modjadu
kusê, kuzê, modi	kmenerâ, kzikil, u kié	u ké, kmanera	kuzê, kumodê	kuzê kumodê

ki tenpu, na ki mare, ki dia, kuandu	kondé, kual é dia	kondé	ki tenpu, ki dia,	ki tenpu, kuandu
branku, klaru	brônk, alvinh	brônk, alv	branku	branku
kenha, kenhe, ken	ken kenhê, kenkel	ken kenhê,	kenhé	kenhi
largu	lôrg	lórg	largu	largu
mudjer, mudjei, mudjel	miêr, mlhêr	amdjêr, mdjêr, spoza	mudjê	mudjér
bentu,	vent	vent	bentu	bentu
aza, asa	aza	aza	aza	aza
linpa	linpá	linpá	linpâ	linpâ
ku	ma	ma	ku	ku
mudjer, mudjei, mudjel	miêr, mlhêr	amdjêr, mdjêr	mudjê	mudjér
matu, floresta, mei d po	môt, floresta	môt, floresta	matu	matu
bitxu	lagarta, bitx	lagarta, bitx	bitxu	bitxu
anu	ôn	ón	anu	anu
marelu, amarelu, kor d obu	merel	amarel	marelu	marelu
bo, bu, abo	bo	bo	abo	bo, abo, bu

The students who contributed to the compilation of this Swadesh list are as follows: Diamantino Tavares Freire, Adelica Teixeira Barbosa Goncalves, Amélia Gomes, Gil Vaz, Delcy Sena, Elisangela Spencer Coelho, Karina Moreira, Lurdes Lima, Maria do Céu Santos Baptista, Eliane Vieira, Emanuel de Pina, Risete dos Santos, Raimundo Lopes, Neidina Lopes, Felisberto Veiga, Angela Cidário, Angélica Cardoso, Hulda Costa, Dysena Pereira, Manuel Rosa, Neidina Lopes, Muniz Afonso, Carmen Oliveira Semedo, Ligia Herbert Robalo, Sérgio Costa, Djito Lopes.