A Comparison of Portuguese-Lexified Creoles of Southeast Asia: Kristang, Makista, Batavia, and Tugu

Robert Laub

School of Languages, Cultures and Linguistics

SOAS, University of London

2022

Abstract

This thesis is a comparison of four Portuguese-lexified creoles in Southeast Asia: the endangered Kristang and Makista, and the extinct Batavia and Tugu Creoles. Kristang is spoken by about 2000 people, mostly in the city of Malacca, Malaysia. Originating in the 16th Century from intermarriage between Portuguese sailors and Malay-speaking locals, Kristang is still spoken today with Malay influence in its structure. Makista, spoken in the Macau Special Administrative Region of China, is far more endangered, with fewer than 100 native speakers. The speakers of this language are descended from the Eurasians who settled in Macau shortly after Portuguese colonization in 1557. While Makista has retained Malayo-Portuguese features from Kristang, it has diverged in certain ways.

Tugu and Batavia Creoles both have their origins in the location of modern-day Jakarta, Indonesia, where Dutch colonists brought Eurasians from Malacca over as slaves in the 17th Century. Batavia Creole, formerly spoken in central Jakarta retained more Portuguese influence, while the less central Tugu Creole shows more Malayo-Indonesian influence.

As creoles, Kristang, Makista, Batavia, and Tugu have their origins from more than one language and their particular background languages are taken into account in terms of their histories and interactions with other languages in their environments. Kristang, with its Malay substrate, continues to exist in a majority Malay-speaking city, while the lexifier Portuguese lost influence after about a century in Malacca. Makista, on the other hand, was removed from the Malay-speaking world and came to be spoken in a city where a majority of residents speak Cantonese, and that to this day has Portuguese as an official language. The difference in structures we can see between Kristang and Makista can be attributed to these sociopolitical differences as well. Batavia and Tugu Creoles were spoken in a city whose only European colonial power was the Dutch.

This thesis addresses the more general question of the influence of substrates on the development of a creole language, and aims to show that Kristang Makista, Batavia Creole and Tugu Creole exhibit features that are representative of their respective histories, ecologies, and languages in contact.

Contents

Acknowledgements

I would like to thank my supervisor, Prof Peter Austin, for everything that he has done along my PhD journey. Not only am I grateful for his role as convenor in the first year of my studies, but for his expertise and advice on all things linguistic. I value his enthusiasm for the topic of Luso-Asian creoles and the connections he has helped me make in Malaysia and elsewhere. Even during retirement, he went out of his way to make sure that I saw this project through, and I cannot thank him enough.

I also acknowledge those in the SOAS community who have helped me on my way: Lutz Marten, Julia Sallabank, Miho Zlazli, Martha Tsutsui Billins, Sarah Dopierala, and Andriana Koumbarou.

This thesis would also not have been possible without the help and support of Stefanie Pillai. For helping me settle into Malaysia during my field trips, and for providing me with her valuable data I am deeply thankful.

Muitu grande merseh to all those in Malaysia who made another home for me in Malacca, especially Adri Phillip, Philo Singo, Anne De Mello, Gerard De Costa, Martin Theseira, Jenny Savage, Kenny Hendricks, Elizabeth Perera, and all of the children of the Portuguese Settlement who helped me meet everyone.

I am especially grateful to Sara Santa Maria and her wonderful family for welcoming me into their lives and their help personally and professionally.

I would like to acknowledge Alan Baxter, Mário Pinharanda Nunes, Silvio Moreira de Sousa, Giorgio Arcodia, and Hans-Jörg Döhla, who have welcomed me into the community of Luso-Asian creolists. I look forward to working together in the future.

Special thanks to my friends who have been my emotional and intellectual support network through the past years: Ana Krajinović, Madoka Hammine, Atsuko Oana, and Hillary Swiss MacKay.

To Mom, Dad, and Rex, your support has meant everything to me.

List of Abbreviations Used

1PL	first person plural
1SG	first person singular
2SG	second person singular
3SG	third person singular
ACC	accusative
ADV	adverb
CLF	classifier
COMP	complementizer
COND	conditional
COP	copula
DECL	declarative
DEM	demonstrative
EMPH	emphatic
FUT	future
LOC	locative
GEN	genitive
GER	gerund
HAB	habitual
INF	infinitive
M	masculine
NEG	negative
PASS	passive
PAST	past tense
PERF	perfect
PFV	perfective
PROG	progressive
PRS	present
PL	plural
PRT	sentence particle
PURP	purposive
Q	question particle
V	verb

Part One: Introduction and Background

Chapter 1: Introduction

1.1 Introduction

There are four languages presently and formerly spoken in Southeast and East Asia which provide a window on the structure of Portuguese-lexified creoles. These are Kristang from Malacca, Malaysia, Makista from Macau, China, and the extinct Batavia and Tugu Creoles of what is now Jakarta, Indonesia. In this thesis I will examine and compare these four creoles, exploring how their structures differ, and how these differences can be attributed to each of their histories. While some scholars argue that creoles form a distinct typological category of languages with little input from other contact and substrate varieties, I will argue that the structures of these four creoles show evidence to the contrary. While they do exhibit features that may be deemed typical of creoles, I will show that these features can be attributed to the languages that have played a role in their creation. This supports the notion of a language ecology, which posits that languages are shaped by the socio-cultural environment in which they were created and spoken.

The reason I chose these four languages to make this argument is that they are: (a) related, (b) distinct in morphosyntactic and historical senses, and (c) spoken in cities with well-recorded histories. The diachronic relation between these languages is known, and the separation point can be found shortly after the Portuguese conquest in 1557 of Macau for Makista, and after the Dutch takeover of Malacca and Java in the early 17th Century for Batavia and Tugu Creoles. We can also easily see the political and demographic distinctions between Malacca, Macau, and Jakarta in their past and present situations, just as we can see the structural differences between Kristang, Makista, Batavia, and Tugu Creoles. While the languages themselves may not have been well-researched until fairly recently, the cities in which they are spoken have well-recorded histories which can help to delineate patterns of regular contact with speakers of these languages throughout the past half-millennium.

1.2 Outline of Sections

This thesis is divided into three main sections, with each section being divided into several chapters. Part One focuses on background information on the languages themselves as well as a discussion of the published literature, research methodology and the theoretical framework for this thesis.

In this first chapter I give a brief introduction on the four languages, such as where they come from, their basic syntactic structure and what other languages were and are spoken near them. I will also discuss certain key terms and concepts that are foundational for this thesis, especially those related to creole studies, or creolistics. The literature is inconsistent in regards to the names of the four languages,, so I will explain which names I have chosen and why. I will also explain the orthography used for each, as well as my rationale for the particular ones I have chosen to represent these languages.

The second chapter consists of a historical background on these languages. It explains their origin within the context of Portuguese colonialism in Asia, their continuing histories within the colonial and post-colonial world, and their status today. While Malacca, Macau, and Jakarta all have history of European colonialism going back centuries, there are differences in each of their colonial stories, ranging from who the colonial power was, to who the main inhabitants were, to how long each city was under European control. The different styles of administration of these places where they were spoken continue today, with Malacca and Jakarta being cities integral to Malaysia and Indonesia respectively, and Macau having a special status as a Special Administrative Region of the People's Republic of China. It is important to understand these histories as they give us clues about how the languages that came from a common ancestor have changed over the centuries into the languages that are spoken today and in recent history.

The third chapter is a literature review looking at previous works in language contact, creolistics, and the debate over creole universals. Ever since creoles were identified as languages in their own right there have been theories that aim to explain why these languages have certain features. While some argue that creole languages form a distinct typological category and exhibit these features due to processes of rapid language formation (Bickerton 1981, 1984; McWhorter 2001a, 2005; Parkvall 2008; Bakker et al 2011), others believe that the typological categorization arises from a prejudiced mindset that separated creoles from non-creoles from a sociological perspective (Youssef 1988, 1990; Siegel 2007; Ansaldo & Matthews 2007; Ansaldo 2009; DeGraff 2003, 2005). The latter scholars additionally argue that features found in creole languages can also be found in the substrate languages and other languages in contact spoken in and around the community in which a creole emerges. This chapter provides an overview of previous literature for creole studies as a whole, and some of the reasons why it is argued to be a distinct sub-field.

The fourth chapter is a literature review that zooms into creolistics and looks more specifically at work on Luso-Asian creoles. The Portuguese Empire in Asia included colonies

outside of Malacca and Macau where other creoles that have arisen, such as in the Indian Subcontinent. This chapter also discusses previous studies of Kristang, Makista, Batavia Creole and Tugu Creole that have helped guide the methods and argumentation in this thesis.

The fifth chapter is a discussion of corpora and data sources, my research methods and methodologies, and a description of the fieldwork I undertook. I have been able to utilize Kristang data from the corpus collected by Professor Stefanie Pillai of the University of Malaya in the Portuguese Settlement of Malacca available in the Endangered Languages Archive[1], and those she provided to me. In addition, I made some recordings myself in Malacca during two field trips conducted in March and October 2017. I also discuss some of the challenges I faced in compiling the data sets, especially because Makista, Batavia Creole, and Tugu Creole do not have easily available corpora. I explain my methods in collecting data in Malacca, including the types of elicitation exercises I made in the field.

Part two contains the sixth chapter, which investigates various features from the four creoles in detail. The aim of this chapter is to explore certain structures that can be found in the languages and the possible sources or influences of these features. Therefore, there will not only be examples from the four creoles, but also from other languages that have potentially exerted influences on them. In addition to the lexifier Portuguese, this includes both standard and Bazaar Malay, spoken in Malacca and across Southeast Asia; Cantonese and Mandarin, spoken in Macau; as well as English, spoken in Malaysia and to a lesser extent in Macau. I will use a mixed analytical methodology for each feature depending on what is appropriate in each context. The features discussed are: serial verb constructions, relative clauses and cleft constructions, short and long pronouns, verbal ellipsis after a TAM particle, genitive patterns, passives, reduplication, and constituent order with a focus on object marking. There are limitations in the data available which means I am not be able to fully describe some features as well as others. Due to the nature of the data sources and the number of present-day speakers, I am able to describe Kristang features in more depth than I can with the other languages.

Part three returns to wider issues within creolistics and how the discussion in part two relates to them. The seventh chapter takes the syntactic description in the previous chapter and places it within the context of the dispute creole exceptionalism (which argues that creoles form a distinct typological group of languages) and substratism (which emphasizes the role of substrate and contacting languages within the formation of creole grammars). In

---

[1] https://elar.soas.ac.uk/Collection/MPI130545

this chapter I argue that the four languages under study show structural differences due to their histories and language ecologies, and that it is the superstrate and substrate languages that have had an impact on the languages' structures, rather than supposed creole universals. I address various topics that have been discussed in the existing literature arguing for and against creole exceptionalism. Among these are language ecology (in the work of Ansaldo), an approach which states that the way speakers interact with each other socially and culturally can have an effect on the languages, including the morphosyntactic structure. I also look at how the data analysis fits within the wider field and theory. We will see how the language ecologies of each of the four creoles discussed have impacted their structures.

This chapter also takes a critical look at comparative creole literature. Creole studies that compare one or more creoles with another do so on the basis that the comparands being creole languages is enough to warrant such research. What makes this study significant is that it involves comparison of four Luso-Asian creoles that have come from a single source, namely the Eurasian community of Malacca.

The notion of simplicity is another issue that often arises in creolistics (McWhorter 2001b, Faraclas 2009, Szmrecsanyi & Kortmann 2012). McWhorter, especially, argues in favor of the idea that creole grammars are the simplest of all languages, due to their origins as pidgins and subsequent creolization. With this in mind, I review some of the features discussed in Chapter 6 and look at their differences in terms of measures of complexity and simplicity, as well as how measuring complexity can be problematic and imprecise.

I then look at the idea that creoles are a social construct, rather than a distinct typological group of languages (Mufwene 2000). This expands upon the idea that languages are a product of their socio-cultural environment and that creoles pick up features from the languages with which they are in contact. This links together the sociopolitical histories of Malacca, Jakarta, and Macau with the morphosyntactic structures that we can see in the creoles spoken in these cities. In agreement with Mufwene's proposal, I argue that creolization is a social process, and that the differing structures of the four languages can be traced back to their particular histories.

Chapter 8 is the conclusion, where I explain how this thesis is an important case study of four related creole languages, each with their own unique histories, but with a common origin. I also address the research questions posed in Section 1.8, and explain how my research has helped to answer them.

1.3 The Names of the Languages and their Speakers

As these creoles are endangered and extinct with no official status, there is not a consistent naming convention for the languages compared in this thesis. There are various names used by speakers as well as researchers (Singho et al 2017, Pillai, Baxter 1988, Pinharanda Nunes 2012).

I refer to the Portuguese-lexified creole of Malacca as **Kristang**. This name has its origins in the phrase *Papiah Kristang*, or "Christian speech" and comes from the Portuguese word *cristão* (Christian), referring to the Catholic faith of the community. It is used by members of the community, as well as Baxter in his 1988 grammar, and the revitalization group *Kodrah Kristang* based in Singapore. It is alternatively spelled as *Cristang*, using orthography that more closely resembles Portuguese. A related term for the same language is *Serani*, or *Bahasa Serani*, also referring to their religion (*Serani* comes from Malay, via Arabic, and refers to followers of Jesus of Nazareth). Serani can be used to refer to both the people and their language (Fernandis 2000).

As the only Portuguese-descended population of any substantial number in Malaysia, the Malacca community often refer to themselves, and their language, simply as Portuguese. This is reflected in the neighborhood of Malacca which has the highest concentration of Kristang speakers, called the Portuguese Settlement. It is believed among some speakers that Kristang is not a separate language, but rather a distinct Malaysian dialect of Portuguese, stemming from the language that was brought there in the 16th Century. They also reject Kristang as a name for their language, arguing that Kristang refers to the Christian religion, and so any practicing Christian would be called Kristang. An example of this attitude can be found in one of Pillai's 2011 recordings: *"Mas relijang Kristang. Kung yo falah Portugis."* ("But [my] religion is Christian. And I speak Portuguese.") Because of the controversy within the community concerning the designation as Portuguese versus Kristang, some publications use the name Malacca Portuguese (Singho et al 2017) or Malacca Portuguese Creole (Pillai 2011). These terms reflect the Portuguese origin of the language, while acknowledging its distinctness from standard Portuguese. I choose to use Kristang because there are previous works using this name (Baxter 1988), it is distinct and brief, and it is an ethnonym used by speakers of the language and other members of the Malacca community. I do not intend to dispute the names other speakers or researchers use for this language.

I refer to the creole of Macau as **Makista** a name also used by Pinharanda Nunes (2008) and Elisabela Larrea (in her blog belamaquista.wordpress.com). It can also be spelled

as *Maquista*, reflecting a more traditional Portuguese orthography (similar to **K**ristang versus *Cristang*). Another name for the language is *Patuá* (from the French *patois* meaning "spoken dialect"), referring to the creole origin of Makista. Patuá is also preferred by the musical theatre group *Dóci Papiaçám di Macau*, who perform skits and plays in Makista at festivals and for an online audience. These performances are referred to as Patuá Theatre, and are listed as an intangible cultural heritage by the Cultural Heritage Department of Macau. Another term used for the language is the descriptive Macau Creole Portuguese (Avram 2015). I have chosen to use Makista because it is an unambiguous and brief name referring to the creole of Macau; no statement is intended about which name should be used for this language. Note that I also want to avoid any confusion between the terms Malacca Creole Portuguese and Macau Creole Portuguese, both of which are abbreviated as MCP.

The names of the Portuguese-based creoles formerly spoken in present-day Indonesia come from the places where they were once spoken, Batavia and Tugu. Batavia, the former name for Jakarta, was used to refer to the creole speakers who were more centrally-based within the city, while Tugu derives from an outer area of Jakarta where the language was spoken. I will refer to these languages as **Batavia Creole** (or **BC**) and **Tugu Creole** (or **TC**) throughout the thesis. In order to avoid confusion with the language I call Batavia Creole, I will refer to the city in Java as Jakarta, both in modern and historical contexts.

Another important general term is **Eurasian**. The Kristang in Malaysia and the Makista in Macau are included in this ethnic designation, but it also includes any people of mixed European and Asian descent, as is also used in Singapore (Pereira 2006). Eurasians can trace descent from a range of people including the Portuguese, Dutch, and British on the European side, and the Malays, Indians, and Chinese on the Asian side.

1.4 Phonology and Orthography

As they are non-standardized languages, the Portuguese-lexified creoles of Asia do not have any official orthograpies. Representations in written work on these creoles tend to be influenced by the language of the author. Recent texts (Baxter 1988 for Kristang, Maurer 2011 for Batavia and Tugu creoles) use a phonemic orthography, based on the pronunciation of the languages. Even more recent is an orthography for Kristang following the expertise of researchers combined with consultation with the local community (Singho et al 2017).

Kristang has picked up many morphosyntactic features under the influence of Malay, which is the main focus of this study (see Chapter 6). It additionally adopted phonemic features from Malay, resulting in pronunciation that is quite different from standard

Portuguese. One important characteristic is the absence of nasal vowels, of which standard Portuguese has five (in addition to five nasal diphthongs) (Cruz-Ferreira 1999). When a word in Portuguese has a nasal vowel or diphthong, the Kristang word will generally have the same (non-nasal) vowel followed by a velar nasal, which is a common sound in Malay, represented by the digraph <ng>. This is evident in the name of the language itself: compare the Portuguese *cristão* (Christian) pronounced [kɾiʃˈtẽw̃], with Kristang, pronounced [kɾistaŋ]. In Portuguese orthography, a nasal vowel can be marked in a number of ways, including the tilde, as in *cristão*, or word-finally, as <m>, such as the word *tem* (have.3SG), pronounced as [tẽj̃], in Kristang, this appears as [teŋ] represented as *teng*.

Because of the Malay phonological influence in Kristang, a Malay-based orthography is well-suited to it. Baxter (1988) uses such an orthography with some modifications made for sounds not found in Malay. While this orthography faithfully represents the sounds of the language, it obscures the Portuguese etymology of much of the vocabulary. Portuguese orthography is largely phonemic, however there can be multiple ways to represent the same sounds, as seen above. Because of this, some Kristang words using the Malay-based orthography do closely resemble the original Portuguese (such as *lembransa*, cf. *lembrança*), while others bear little resemblance (such as *ki*, cf. *que*). Although this style of transcription is useful to researchers, community members undertaking revitalization have shown a preference for a more Portuguese-like spelling system.

Stefanie Pillai of the University of Malaya has conducted extensive documentary research in the Portuguese Settlement of Malacca (a more detailed account of her methodology and methods will follow in Chapter 5). One of her goals was to produce an accurate representation of the phonology of Kristang, in particular its vowels, and she determined that there is a phonemic distinction in Kristang between the vowels /e/ and /ɛ/. While this is a distinction that may exist in Malay, depending on the analysis, it is not reflected in the orthography. In fact, Malay orthography uses <e> to represent three sounds: /e/, /ɛ/, and /ə/ (Yunus 1980). Pillai wanted to represent distinction between the two mid-front vowels in the Kristang orthography, and decided to use accent marks to distinguish these two sounds, with <e> representing /e/, and <é> representing /ɛ/. This does bear resemblance to Portuguese orthography, where <é> is also used for /ɛ/.

Another important distinction is /a/ versus /ə/, both of which can be represented with <a> in Malay. In Kristang, word-finally, <a> represents /ə/, while <á> represents /a/. Note that schwa is not stressed and hence we find pairs such as the following:

Kristang:	caza	cazá
IPA:	[ˈkazə]	[kaˈza]
English translation:	house	to marry
Portuguese translation:	casa	casar

Note that the diacritic reflects both the vowel quality and the stress pattern, and the use of a diacritic is a nod to Portuguese, as its orthography also uses diacritics to mark stress and vowel quality.

Pillai (personal correspondence) noted that it was important to this Kristang community to recognize their Portuguese heritage, and that this influenced their orthographic choices. One particularly important choice is the use of the letter <c>. While [k] in Portuguese is represented with both <c> and <q>, in Malay it is represented with <k>. Previous publications, including Baxter (1988), used the Malay-based orthography and have thus employed <k> to represent this sound. In Portuguese orthography the letter <c> represents a voiceless velar stop [k] before consonants and the vowels <a>, <o>, and <u>, but a voicelsss apico-dental fricative [s] before <e> and <i>. In Portuguese, [k] before <e> and <i> is represented with <qu>, such as in *queijo* "cheese" (Bergström & Neves 2011: 16). To reflect Portuguese orthography, Pillai and her team decided to use <c> to for the phoneme [k] before consonants and <a>, <o>, and <u>, and use <k> elsewhere. This is a departure from standard Malay orthography, in which <c> represents the voiceless postalveolar affricate [tʃ]. Kristang also has this sound, but instead uses the digraph <ch>, another resemblance to Portuguese orthography. Currently, the digraph <ch> is the voiceless postalveolar fricative [ʃ], but at the time of creolization in the 16th Century it represented the affricate [tʃ] (Azevedo 2005: 156). The orthography for Kristang that is used in this thesis is the Malay-based one that Baxter proposed, as it is more strictly phonemic than the recently developed spelling preferred by the community. The same orthography is also used in other sources, such as the written works by Joan Marbeck (1995, 2004), and *Kodrah Kristang*. While the distinction between vowels is better represented in the Pillai orthography, it is not relevant to the present study which has a largely morphosyntactic focus.

All Makista data used in this thesis is sourced from written corpora, so I present it in this work as it is in the original documents. The orthography used in the Makista texts is very clearly heavily influenced by Portuguese spelling, albeit with some usage that reflects the sounds of the creole. While Makista phonology is not very different from Kristang (Batalha 1988), it developed in a city where Portuguese was the main language of administration and

education, and where there was little to no Malay presence, especially in writing. One main reason for choosing the original spelling is that it is hard to know the exact pronunciation of source forms and hence it is difficult to provide an accurate phonemic transcription. As the spellings are based on Portuguese orthography, their etymologies and relations to Kristang words are generally clear; where not, differences will be explained. One important distinction is the vowel [u] word-finally and unstressed. Take the following example:

Makista:	lôgo
Kristang:	logu
Portuguese	logo

This represents the pre-verbal future-irrealis marker in Makista and Kristang, and the adverb "later" in Portuguese, from which they derived. The actual pronunciation of these three words is roughly the same in the three languages, namely [logu]. While Kristang represents this vowel phonemically, Makista uses the final <o> as in Portuguese. The Makista corpora come from a number of original writers, who sometimes have inconsistent or idiosyncratic spellings of words. My choice to keep the original spellings, even if the same word is represented in different ways, is to ensure an accurate representation of the language. All examples in this thesis are glossed so there should be no confusion as to what each sentence, word, or morpheme means.

The TC texts from Maurer (2011) make use of a Dutch-influenced orthography, owing to the documentation taking place during the Dutch control of present-day Indonesia (1782-1949). The following example text is from Schuchardt (1891):

Loegar santoe nosoter o djenti djenti soeä. Nosoter fika denter oengah kampong ki piklinoe ki nos loemija Toegoe ò thing denter konta pègadoe Becassie podeer Meester Cornelis.

The place where we live and its people. We live in a small village which we call Tugu and it is in the district Becassie, in the Meester Cornelis division.

Maurer (2011) provides a phonetic representation for the same text:

Lugar santu nosotër o djenti djenti sua. Nosotër fika dentër unga kampong ki piklinu ki nos lumia Tugu o ting dentër konta pegadu Becassie poder Mestër Cornelis.

As I did for the Makista data, all of the source materials I have available on BC and TC are written, so I follow Maurer's representation.

1.5 Pidgins and Creoles and Key Concepts

This section introduces pidgin and creole studies, and gives a brief overview of the key concepts of the field. I draw on established literature within creolistics to give definitions of terms that I will be using throughout this thesis. I also include discussion of other pidgins and creoles to help illustrate the definitions.

The period of European colonialism in the 17th to 19th Centuries resulted in the mass movement of peoples across continents, and interactions between peoples from differing backgrounds within the colonies. Slavery and oppression of human rights often led to native languages being suppresssed, and groups of people from different language backgrounds often coming together with the only common tongue being that of the colonizer. Trade was another context in which people from different backgrounds had a need to communicate with each other without access to the other's language.

As merchants from one region work with those from another, the communication necessary between them is limited to their professions, so a rudimentary language system arises, called a **pidgin**. A pidgin is a simplified language that has no native speakers (Holm 2000), as it is developed by peoples of disparate backgrounds coming into contact with limited necessity of communication. It is simplified in that it lacks ways to communicate concepts and thoughts that a natural language could, and is lexically limited to the trade or business conducted by its speakers. It is not a stable system, which means that the structure can change very quickly, much more than a natural language would (Hymes 1971, Whinnom 1971, Müllhäusler 1986, Arends et al 1994).

The contact situations throughout the colonial period lasted long enough for Europeans to settle in the areas in which they were conducting trade, to marry locals and form a mixed community, as was the case in Malacca. As members of different language backgrounds have children, they may continue to use a pidgin amongst themselves. As their children acquire this language they become native speakers, and the language expands to a full stable system called a **creole**. In contrast to a pidgin, a creole is a full language with the communicative ability of any other natural language. Its syntactic structure is stable compared to that of a pidgin (Bickerton 1984; Thomason & Kaufman 1988; Holm 1988, 2000)

Because pidgins and creoles arise from contact between two or more languages, the languages spoken by pre-creolized populations play a role in their creation. The interaction between these input languages often has much to do with power dynamics, given the role that

Europeans typically played in colonialism. One of the input languages typically provides the lexicon of the resulting creole, and is called the **superstrate**, or **lexifier**. In the creole languages created by the European period of colonialism, the lexifiers are the languages of the European powers. This includes Portuguese in regards to Kristang and Makista, as well as Cape Verde Creole; French for Haitian Creole, Seychelles Creole, and other creoles from former French colonies; and English for Jamaican Patois as well as Pacific Islands pidgins such as Tok Pisin and Bislama (Michaelis et al 2013).

The other languages that provide input into the creation of a creole are those which are spoken by the people with less power. While not necessarily the case in every creole, it is often these languages that provide the syntactic structure, and are called the **substrate**. An example of a substrate in the case of Kristang is Malay, the dominant language of the people living in Malacca prior to European intervention. For creoles spoken in the Caribbean, which arose out of enslaved populations from Africa, various Niger-Congo languages of West Africa comprise the substrate of the creoles spoken in Haiti, Jamaica, and other Caribbean islands (Lefebvre 2015, Lalla & D'Costa 2009).

Restructuring is another important term in pidgin and creole studies (Bruyn 2008), referring to the process in which a creole language develops and acquires features from substrates or internal innovations. Creoles are distinct in that they have typically gone through rapid change and creation of features novel to that particular language. Thus, the change from, for example, classical Latin to Portuguese took place over many centuries (Williams 1962), while the transition from Portuguese to Kristang took place within a few generations.

Grammaticalization is another important term which not only applies to pidgins and creoles, but has occurred in other non-creole languages as well, where a lexical item comes to be used as a grammatical function (Hopper & Traugott 2003). One notable example is the expression of future tense suffixes in the Romance languages. Today the future is expressed by suffixes that attach to the finite verb, but ultimately these derive from the Latin form of *to have* (Maiden 2010: 264). Another example is the English modal verb *will*, which in the past referred to wanting, but today expresses a future modality (Bybee & Dahl 1989: 58). This is relevant to the four creoles in the present study, as there are many morphemes that today are grammatical particles but have their origins in words. Such examples include the tense-aspect-mood (TAM) markers, such as the Kristang future-irrealis marker *logu*, which derives from the Portuguese *logo* "later", and the genitive particle *sa*, originally from Portugues *sua*, a possessive pronoun (Baxter 1988).

Creole languages are spoken around the world, with some of the more well-known being found in the Caribbean, due to the fact that they have large speaker populations. Thus, Haitian Creole and Jamaican Patois, mentioned above have speakers numbering in the millions (Fattier, 2013, Farquharson 2013). Other nations have creoles as an official or national language, such as Vanuatu, a former British-French colony, where the creole language Bislama is spoken by the majority of the population and is recognized as a national language (Meyerhoff 2013).

Kristang fits well into these concepts of creole languages; in Malacca after the Portuguese took control and intermarried with local women, the creole language that arose out of this contact situation had Portuguese as the lexifier, being the language of the colonizing Europeans and those in control of Malacca at the time, and Malay as the substrate, being the language spoken by the locals. Additionally, Portuguese as the lexifier provided most of the words used in Kristang, with comparatively very little vocabulary from Malay or other languages. Malay, as the substrate, provided the grammatical structure of Kristang (examined in Chapter 6), such as the lack of verbal morphology for tense and person which is found in Portuguese, but not in Malay.

After Makista branched off from Kristang, an already creolized language, it continued to be spoken in a Portuguese colony with a lexifier/substrate dynamic between the languages. The language of power in Macau was Portuguese, and for a longer period of time than it was in Malacca. The syntactic structure of the creole with Malay influence remained, but a new language entered the picture for Makista speakers, namely Cantonese, spoken by the majority of those living in the colony.

BC and TC are similar to Makista in that they branched off from Kristang post-creolization. However, they developed in a city (Batavia, modern-day Jakarta) that was never a Portuguese colony. Although Portuguese was historically used as a lingua franca among European traders in Southeast Asia, the only European power to ever have control of Jakarta was the Netherlands. In Jakarta, Dutch did not act as much of a lexifier as the communities continued to use their Portuguese-lexified creole (Maurer 2011: 3). However, being in the Malay-speaking sphere, they were exposed to Malay as a substrate through the centuries, just as Kristang was (Maurer 2011: 6).

1.6 The Origins of Kristang, Makista, Batavia Creole and Tugu Creole

The presence of the Portuguese in Southeast Asia starting from the early 16th Century left an impact on the languages in the region. Long being a center of trade between many nations

and cultures (Heng 2009), the Portuguese arrived in Malacca as part of an eastward expansion of their existing trading networks, already established in Africa and the Indian subcontinent. While settlement was not a primary goal in the establishment of colonies throughout Asia, there was a mixture of people from Portugal and the local population during the period of colonialism. This led to the creation of creole languages that are still spoken today, although currently endangered (Arana-Ward 1977; Baxter 1988; Clements 1996, 2009; Hancock 1969; Maurer 2011).

The intermarriage of Portuguese sailors and Malay women in Malacca led to the mixing of Portuguese and Malay into what is now the creole language spoken by Portuguese Eurasians in Malacca (see section 2.5 for a more details account). This language has a largely Portuguese-derived vocabulary, with Malay influence on its syntactic structure (Baxter 1988: 3). Kristang continued to be spoken in Malacca through the Portuguese period (from 1511 to 1641) into the subsequent Dutch (1641-1795, 1818-1825) and British (1795-1818, 1825-1957) colonial periods and until today in independent Malaysia (1957-present) (see Chapter 2: Historical Background). Currently the language is endangered with major threats being the increased usage of Malay and English in education and business, with younger generations increasingly speaking English at home (Baxter 1988: 13). After the Dutch ousted the Portuguese from Malacca in 1641, there was little ongoing influence from standard Portuguese, with the exception of its use by priests who visited Malacca or were stationed there by the Catholic Church. As a result, the language was essentially isolated from standard Portuguese. Evidence of Malay influence is obvious in Kristang and will be discussed in detail in Part 2.

As Portugal expanded its colonies eastward, a colony was founded in China in 1557 as a waypoint to Nagasaki, Japan, from Southeast Asia. China granted land to Portugal in the south of China, in what became Macau. The Chinese were not initially allowed to settle in the colony, so Portugal relied on its own people to populate and conduct trade in the city, including Eurasian creole speakers from Malacca. Portuguese control of Macau continued, and the restriction on Chinese inhabiting the city became less and less strict, leading to a city with a majority-Chinese population, mostly speaking Cantonese (Sit et al 2012). The creole-speaking population in Macau have always been a minority in their city. The language that they spoke, derived from the creole in Malacca, evolved separately and became its own creole variety. Like Kristang, it is known by various names including Macanese or Patuá; as noted above, in this thesis I will refer to it as Makista. Unlike in Malacca, the Portuguese presence in Macau lasted for over four centuries, up until the 1999 handover to China.

Similar to Hong Kong, the city is a Special Administrative Region (SAR) of the People's Republic of China, with its own local laws under the One Country, Two Systems rules (Hao 2011: 45). Part of the separate rules is that Portuguese remains as an official language in Macau, despite Cantonese being spoken by the vast majority of Macau residents, according to the 2011 Macau Census.

The Dutch took over Malacca about a century after the Portuguese first took control. In addition to territory on the Malay Peninsula, they took over the islands that are today the country of Indonesia, including its most populous, Java. The biggest settlement on the island, Batavia, became a hub for Dutch trade, and they took slaves from their other colonies with them (Maurer 2011). This included some of the Eurasians living in Malacca at the time. The Portuguese creole-speaking community continued on, but were split into two main groups. The one which remained in the center of Jakarta spoke what I refer to as Batavia Creole (BC) in this thesis, while the second group settled in the area of Tugu on the outskirts of the settlement, from which the language's name Tugu Creole (TC) comes.

1.7 The Situation Today

Portuguese-descended people still live in Malacca and throughout Malaysia, although they remain a very small minority. The 2010 Malaysian Census reported Malacca City to have a population of 484,885, of which 3,240, or 0.7% of the population, reported themselves as "other" race (*lain-lain*), i.e. not Bumiputera (Malays and other groups native to Malaysia), Chinese, Indian, or non-citizens. As the census does not provide information beyond "other", it is difficult to know exactly how many of these are Kristang. Note that this number refers only to the self-reported race of the residents of Malacca. Out of the members of the Kristang community, the number of those who speak Kristang is declining. Kristang is rarely used by the youth in the Portuguese Settlement, and most native speakers are over the age of 40. Education and economic opportunities have introduced English as the main language spoken at home rather than Kristang (Baxter 1988). Further discussion on present and past demographics of Malacca can be found in Section 2.7.

Despite being a language in decline, Kristang culture is still vibrant. The community lives in a densely-populated neighborhood where their culture and identity are fully on display, and is promoted in many ways. They take advantage of their location in a city which is a major tourist destination to invite visitors into their neighborhood to experience the Portuguese Malaysian culture. In a reflection of their faith, their major cultural events revolve around the Catholic calendar, including the Intrudu Festival marking the beginning of Lent,

which involves splashing water on each other and celebrating outside. The water symbolizes the purification of baptism, although all are invited to attend, regardless of religion. As one of the largest Christian neighborhoods in Malacca, they celebrate Christmas in the streets of the Portuguese Settlement as well, displaying their Christmas lights and inviting others inside their homes for food. The biggest festivals of the year are the Feasts of St. John and St. Peter (Festa San Juang, Festa San Pedro). As St. Peter is the patron saint of fishers, his feast day is of great importance to this seaside community. During this festival Catholic priests bless a boat with a statue of St. Peter, which is then brought out to sea. The festival also includes singing, dancing, and eating (Koh 2017).

Apart from the traditional Catholic holidays, the Portuguese Settlement hosts a number of restaurants for tourists serving their local cuisine. This is another way the Kristang can share their culture with others while taking part in the tourist industry of Malacca. These cultural events show that despite their small size as a community, they are strong in their faith and in their culture, and are enthusiastic about sharing it with others. It is from this enthusiasm and a recognition of the decline in number of Kristang speakers that there has been an effort to revitalize the language. Singho et al (2016) is a language the study book *Beng Prende Portugues Malaká (Papiá Cristang): Come, Let's Learn Portugues Malaká (Papia Cristang)*. Created in collaboration with the Kristang community, it is part of an ongoing effort to revitalize the language and celebrate it as part of Malaysian culture.

Other Kristang writers include Joan Margaret Marbeck, from whose volumes *Ungua Adanza* (1995) and *Linggu Mai* (2004) I include examples (see Section 5.4). She is a Eurasian living in Malacca who wanted to preserve her language, and so she wrote songs and poems and retold stories in Kristang that she remembered from her childhood growing up in a Malacca Portuguese family.

Similar to the situation in Malacca, the Makista make up a very small minority in Macau. Another parallel is how demographics are measured: there is no answer for "Macanese" or "Eurasian" in the Macau census. According to the most recent census in 2011, Macau had a population of 552,503, of which there were 4,019 reported as being "Chinese & Portuguese", and 3,485 as being "Portuguese". Added together these make up 1.3% of the population of Macau, although it includes those who may not be ethnically Makista. Unlike the Kristang, the Makista do not have a neighborhood like the Portuguese Settlement that is a center of their cultural activities. Further information on the demographics of Macau can be found in Section 2.7.

Despite the lack of a single area with a high concentration of Makista speakers, they do have a presence in the cultural landscape of Macau. Chief among these is the Patuá theatre performed by the group *Dóci Papiaçám di Macau*, which performs yearly at the Macau Arts Festival. This group, while not consisting totally of native Makista speakers, performs music and comedy skits in Makista. They also have a presence on the Web and have posted multiple videos on YouTube[2].

Makista is even more seriously endangered than Kristang. According to Lewis et al (2019), there were 50 native speakers in 2007, and that number is likely to be smaller as of the time of this writing.

BC and TC are no longer spoken today. BC ceased to be spoken by the end of the 19th Century (Schuchardt 1891), and the last speaker of TC died in 1978 (Baxter 1996), but their descendants still live in Jakarta.

1.8 Research Questions

The four creoles in this study have the same origin, but have since diverged into different varieties spoken by distinct communities. Kristang was cut off from the lexifier Portuguese, while Makista was cut off from the substrate of original contact, Malay, and remained in contact with Portuguese. Additionally, Makista-speaking people came into contact with the Cantonese-speaking majority of Macau. BC and TC, meanwhile, developed in a city that was never a colony of the Portuguese. These factors are likely to have led to the divergence of structures in the four languages (detailed in Chapter 6 below). My thesis is informed by the following research questions:

1. In what major structural ways are Kristang, Makista, BC, and TC similar, and in what major ways are they different?
2. How did the substrates (Malay for Kristang; Malay and Cantonese for Makista; Malay and Indonesian for BC and TC) affect these differences?
3. How did the lexifiers (mostly Portuguese, some Malay and English for Kristang; mostly Portuguese and Cantonese, some Mandarin and English for Makista; Malay and Indonesian for BC and TC) affect these differences?

In addition, my research intends to place these questions within the context of comparative creole studies. Although not new in terms of linguistic description, the present study aims to look at a particular, and rather unusual, case study of four creoles that have a historical

---

[2]See https://www.youtube.com/channel/UCxRaAKi5QGPlfGi3Vbr8Ngg for *Doci Papiaçám*'s YouTube Channel

relation but rather different sociopolitical histories in Malacca, Macau, and Jakarta. To achieve this aim I required a comparison of each language to the lexifier and substrates, and how they could have had an influence on the structure of the four languages.

As we will see in Part 2, these questions will be addressed with examples from the creoles that show how they differ. This will then be placed within the context of their lexifier and substrate languages, and how the differences we have seen can be found to have origins in the other languages.

1.9 Theoretical Framework

Creoles were first identified as a separate language group and began to be studied seriously in the 19th Century. As Schuchardt (1883, 1889, 1890) noted, while their lexicons were derived from European languages, their grammars differed significantly. Hypotheses have been put forward to account for this, such as the influence of the mother tongues of the original pidgin speakers, as well as a resetting to the parameters of universal grammar (Bickerton 1981, 1984). Up to today the largest communities of creole speakers are in the Atlantic, and in particular the Caribbean, with nations such as Cape Verde, Haiti, and Jamaica having creoles as their majority languages (Michaelis et al 2013). Due to this, a large part of creole language studies (creolistics) is about the Atlantic creoles, which arose due to the European colonization of the Americas and the slave trade the European nations conducted between Africa and the Western Hemisphere (Holm & Patrick 2007).

The creole languages under investigation in this thesis are and were spoken in Southeast Asia, a region where creole studies have not been undertaken as extensively, perhaps partly due to the smaller number of creole speakers in this region along with South Asia. Kristang, Makista, BC, TC and the other Asian creoles came about due to European colonization, as the Atlantic ones did. However, they arose in colonies that were more concerned with trade than settlement, plantations, or slavery (Arends 1994). In the Caribbean, entire indigenous populations were eradicated and replaced with slaves from Africa. These slave societies were populated with people from various groups who spoke different languages, and were far removed from their homelands. The political and social domination of European culture both before and after independence resulted in these languages becoming the tongue of the majority, or a very significant language in the society (such as in Haiti, see Muysken & Veenstra 1994).

The Asian creoles, on the other hand, were largely created out of intermarriage between the colonizing Europeans and local women, creating a mixed society that existed as

part of a larger one dominated by locals, if not politically then demographically (Cardoso et al 2012). They were therefore not far removed from the substrate language, as it remained the language of the majority in the place where the creole was spoken, for instance Malay in Malacca. While populations did change in this region due to European colonization, it was not the only cause of population shift, as there was movement of people between Asian nations, such as Chinese settlers in Malaya (Heng 2009). Additionally, the scale of shift was not as massive as it was in the Americas, with almost complete demographic changes on some Caribbean islands where the indigenous population was wiped out and replaced by slaves from Africa. This smaller demographic shift that occurred in Asia is taken into account within the larger scope of creolistics. Due to both geographical location and history studies of Asian creoles are a distinct branch of the field, separate from studies of Atlantic creoles. Works such as Ansaldo (2009), Ansaldo & Matthews (2004, 2007), Pinharanda Nunes (2004, 2008), Cardoso (2009) and Clements (2012) fall into this branch.

Another important point of debate is the issue of Creole Exceptionalism, as advocated by Bickerton (1981, 1984), Bakker (2011, 2014), and McWhorter (2001, 2005, 2018). This is the notion that as creole languages are created, they take on parameters that are in a neutral state with regards to universal grammar, and thus features found in creole languages develop out of this, rather than out of features originating in a lexifier or substrate variety. Supporters of this view argue that creoles have relatively simple syntactic and phonological systems, in comparison with both their lexifiers and substrates. They argue that creoles tend to be more analytic than their contact varieties, and that this is evidence of a reset to universal grammar parameters. One example of such is the tendency of creoles to not use verbal morphology in the way that European languages such as French or Portuguese do. This also applies to Kristang and Makista, which use pre-verbal markers for aspect instead of suffixes to mark tense, as is the case in Portuguese.

These proposals about universals, however, have been challenged by Mufwene (2001), DeGraff (2003), and Ansaldo (2009) among others, arguing that creoles are considered as a distinct subgroup of languages largely due to their histories as the languages of the colonized and enslaved. This can have the implication that the structures of such languages are less developed than those of the lexifier languages, despite creoles being fully-fledged, complete languages rather than an inferior dialect of a standard European language. This alternative viewpoint argues that the substrate languages have a significant effect on the structures of the resulting creole languages, a position known as substratism. Substratists propose that the substrate language has an effect on the structure of the resulting creoles, as

the transition to a creole, while sudden, does not necessarily involve a complete break from the substrate language. This is especially true with Kristang, where its main substrate, Malay, has remained the majority language in Malacca. The phonological system of Kristang also is very similar to that of Malay (Baxter 1988).

As for the above two points of view, I take a more substratist position in the description and analysis of the morphosyntactic structures of these Asian creoles. It is true that many of the features in these creoles have parallels with their substrates, and when compared between Portuguese and Malay, or Portuguese and Cantonese, the creole features tend to align more with the substrates than they do with Portuguese. However, this is not the case with every single feature, and these creoles are not simply calques of the substrate languages. The situation in Southeast Asia thus provides us with a good opportunity to look at how these differing structures developed. We have instances of the same substrate (Malay), but with different contact histories. We also have an instance of a completely new substrate (Cantonese in the case of Macau) that does not exist in the other three contexts. There is also variation in the degree of contact with the lexifier, which we will see has likely had an impact on the resultant structures of these languages.

Another important theoretical framework is the Feature Pool hypothesis (FP). It states that the languages from which a creole has derived constitute a pool of features from which the creole can incorporate instances. This leaves the creole with the possibility, but not the certainty, of incorporating various features from the different substrates. As we will see in Part 2, there are some features in the creoles under consideration that highly resemble one from the substrate, while other features do not have such a strong resemblance (Mufwene 2002, Siegel 2008a, Aboh 2009).

Mufwene draws parallels between the development and evolution of biological species with that of languages. Just as an organism has features that can be attributed to the genetic background of its ancestors, a language can also contain characteristics that are indicative of the languages or dialects that have provided input to it. Mufwene (2002: 46) explains:

> partners in the reproduction of a species, which I take a language to be, make
> contributions to the structural makeup of new members… which share
> features on the family resemblance model. These new members select their
> features (typically with some modification) from the same pool, although the
> recombinations are never the same, and those which wind up as dominant are
> not always the same.

Creoles are particularly relevant to such a model of the feature pool, as they demonstrate what can happen when two or more very different languages come into contact. This hypothesis informs the importance of looking at the demographics and sociopolitical situations of Malacca, Macau, and Jakarta, as these languages were parts of these large and multicultural cities.

Chapter 2: Historical Background

2.1 Introduction

This chapter explains the historical background of Kristang, Makista, Batavia Creole and Tugu Creole. Because of the relatively recent development of these languages there are well-recorded histories of the cities after their emergence. We look first at the history of Malacca, and why it became an important trading hub and a multicultural city in the first place. We will then see how the Portuguese came to be in the region and why they wanted to set up colonies in this part of the world, and how their presence led to the creation of a community that speaks a Portuguese-lexified creole language. We will then look at Macau, how it came to be a Portuguese colony, and how creole speakers ended up there, and the origins of Makista. We will also explore the later colonial history of the region. The Dutch, who took over Malacca from the Portuguese in 1641, ended the period of substantial Portuguese contact in Malacca, but it also led to the creation of Portuguese-descended Eurasian communities in Jakarta. We will see how some of the creole-speaking Eurasians were taken from Malacca to Jakarta in the early 17th Century, and how they eventually split into two different communities with distinct languages.

The differing histories of these cities can explain how the differences between the four creoles arose, so we will also look at the centuries following the initial Portuguese period. Portuguese control of Malacca lasted from 1511 to 1641, followed by Dutch (1641-1795, 1818-1825) and British (1795-1818, 1825-1957) holding of the country, ending with Malaysian independence in 1957. Macau, on the other hand, was under Portuguese control until 1999, and today is a Special Administrative Region of the People's Republic of China. During the colonial period Jakarta was always under the control of the Dutch, aside from a relatively brief period under the Japanese during the Second World War (from March 1942 to September 1945). Shortly after the war Indonesia declared its independence with Jakarta as its capital.

The histories of each location show us what languages would have been dominant there over time, and give us an insight into what languages have had the potential for the most influence on the creoles' development and evolution.

2.2 Malacca and Monsoon Asia

Since before the beginning of the European age of exploration and colonization in the late 15th Century, maritime Southeast Asia was a major hub for trade, being located along

important sailing routes. Ansaldo (2009) describes the region as Monsoon Asia, since the monsoon, winds that blew across the Indian Ocean and shifted direction with the seasons, were a major factor in aiding maritime trade across Asia (McPherson 1993). Just as the Mediterranean was a center for widespread trade and intercultural contact between Southern Europe, North Africa, and West Asia, so was Monsoon Asia. The Malay Peninsula's physical location between the Indian Ocean and the South China Sea (see Map 1) made it an important locus for the trade and contact between cultures that led to it becoming economically rich and politically powerful. The Strait of Malacca, about 50 km wide at its narrowest point, provides a lane between these two important trading areas, and a link between the civilizations of East and South Asia. It was a combination of the monsoon winds, helping to carry ships between East Asia and the Indian Subcontinent (and vice versa), and the location of Malacca, at the narrowest point along the important sea route, that led to its importance and dominance in the region (Heng 2009: 1).

Map 1: Overland and sea trade routes across Asia (UNESCO)

The trade network that spread across Asia was used by the Chinese, Indians, Malays, as well as Arab traders from further to the west. Malacca was in the center of this network, and was where these groups of people would mingle and trade goods from all across Asia. As people from various backgrounds were working with each other, the resultant linguistic ecology created language mixing and trade pidgins. The most widespread in Malacca and

what is today Indonesia and Malaysia was Bazaar Malay (Bao & Aye 2012), which seems to have been employed in a similar manner to Lingua Franca in the Mediterranean. Bazaar Malay was a variety of Malay that was most commonly found in Malacca and other trading locations. Structurally, it was much more isolating than literary Malay (Ansaldo 2009: 62) and lacked its prefixal and confixal noun and verb morphology. The port of Malacca in which the Portuguese arrived in the 15th Century thus had a multicultural and multilingual ecology that pre-dated European settlement.

2.3 The Portuguese Empire

The trading network in the Middle Ages that included East and South Asia did extend further west through to the Mediterranean Sea and mainland Europe. As seen in Map 1, a part of the trading network were the overland routes from China to West Asia and Europe, known as the Silk Roads (Frankopan 2015). These routes facilitated trade between East and West, and goods from the Far East, such as silk, jade, precious metals, and ivory, would have traveled across this route before ending up in Europe.

The Silk Roads also passed through different nations and empires, from China to Mongolia, the Kushan Empire of Central Asia, Persia, the Islamic empires of the Middle East, and the Byzantine Empire in the eastern Mediterranean (Miller 2017). As no single political entity dominated the trade routes, they were subject to the instability of politics or war in the region. Rather than being a single route, it was a multi-stranded and ever-changing network that connected the East and West. For goods from the East to reach the Iberian Peninsula on the western periphery of Europe, for example, they would have had to travel a large distance and change hands multiple times.

The Iberian Peninsula has a history of conquest and rule by other powers. It formed a part of the Roman Empire, including the province of Lusitania, which roughly corresponds to present-day Portugal. In the 8th Century CE, Arab and Muslims invaded Iberia and much of present-day Spain and Portugal was conquered. The Muslim state, Al Andalus, lasted until the 15th Century and was part of the larger Islamic world that stretched across the Mediterranean to the Middle East (Diffie & Winius 1977). As the Christian forces of Iberia reclaimed their land during the Reconquista, and formed kingdoms of their own, the importance of maintaining intercontinental trade was paramount. The regions nearest to Spain and Portugal were obvious starting points for establishing colonies and dominance in trade. In 1415 led by King John I, Portugal

conquered Ceuta on the north African coast, which was the beginning of a hunger for further conquest in Africa (Newitt 2005).

The Catholic Iberian kingdoms soon took advantage of their locations on the Atlantic coast of western Europe and started expanding further west and south; islands such as the Canaries and Madeira soon came under their influence. Competition between the kingdoms led to conflict, especially over who would control the newly conquered lands. In 1479 the Treaty of Alcáçovas was signed between Castile and Portugal, which divided the Atlantic into two zones of influence, giving Portugal the right to explore the African coast while assigning the Canaries to Spain (Ferreira 2015: 31). Bartolomeu Dias, a Portuguese nobleman who led an expedition south along the coast Africa in 1488 in search of a sea route to India, made it to the Cape of Good Hope but did not proceed further than the southern tip of Africa, however upon his return to Lisbon the existence of a sea route to India became known to the Portuguese (Crowley 2015: 32).

In the late 15th Century, Genoan sailor Christopher Columbus sailed across the Atlantic to the Americas in the service of Castile and León, claiming the lands he saw in the name of the Spanish crown. Portugal disputed this, arguing that the places Columbus went to were further south and therefore landed in the Portuguese zone of influence. As it became clear that Portugal and Spain were to conquer and seize land previously unknown to them, another treaty was necessary. In 1494 the Treaty of Tordesillas was signed by Portugal and Spain, authored by Pope Alexander VI as mediator. The treaty divided the world along a line of longitude defined at 370 leagues west of the Portuguese Cape Verde islands. Spain was given rights to conquer land west of this line without interference from Portugal, while Portugal was given rights to conquer land to the east without interference from Spain (Kupperman 2012: 15).

Map 2: Treaty of Tordesillas Meridian (Encyclopedia Britannica)

2.4 The Portuguese Arrival in Asia

A result of the 1494 treaty was a lack of interference between the Atlantic seafaring powers, which prompted Portugal to send sailors around the coast of Africa to India without hindrance from other European nations. In 1497 under King John III, Vasco da Gama set sail around Africa and landed in India in 1498, returning to Portugal the following year. This was the longest sea journey known to have been made at that point by a European – longer than the equator – but it provided Portugal with a direct sea route to India and increased opportunities of trade. Despite the length of this route, the Tordesillas treaty gave Portugal a monopoly on the routes around Africa, which were preferable to the dangerous and politically unstable Mediterranean and overland journeys (Burnet 2013: 72).

The Portuguese founded a trading network that started around the coast of Africa, headed east to India, and eventually reached Nagasaki, Japan, the furthest trading post from Lisbon (Sit et al 2012). What resulted was a widespread network, thinly populated by Europeans, aimed at bringing back goods to Portugal. Portugal itself did not have the population to maintain this trade, and relied on intermarriage and people recruited from stops along the way to assist (Ansaldo 2009). This trading empire was "heavily localized, decentralized, and deeply integrated, out of necessity, with local political structures" (Ansalso 2009: 73). Because participants in this network came from various areas with diverse

linguistic backgrounds, they used a pidginized version of Portuguese to communicate with each other. The first non-Europeans the Portuguese worked with were West Africans, due to their geographic proximity. As the network expanded eastward, the Portuguese brought with them people from more western areas of the network. Communities were established in South Asia, modern-day India and Sri Lanka, where Portuguese creoles also developed (Clements 1996, Cardoso et al 2012) By the time they arrived in Malacca, they had already been using a type of Portuguese pidgin amongst themselves for some time (Ostler 2005).

2.5 The Portuguese in Malacca

In Malacca a culturally and linguistically diverse population was encountered by a similarly diverse group of sailors and traders. Therefore, the pidgins and creoles that emerged there were not simply a result of mixing standard Portuguese with standard Malay, since Bazaar Malay was already present as a pidgin for inter-group communication. As mentioned above, Bazaar Malay was more isolating and less morphologically complex than Literary Malay. Kristang and Makista are also relatively isolating, which could be a result of their being descended from admixture between Portuguese (pidgin) and Bazaar Malay, rather than from being creoles in and of themselves. Bazaar Malay was used as the lingua franca among Asian traders in Monsoon Asia, and Portuguese was eventually used as the lingua franca for Europeans (de Silva 2008). As many people in the area had knowledge and access to these languages, this ecology laid the foundations for the formation of new creole languages.

The Portuguese soon became the dominant European power in Southeast Asia, setting up colonies in islands such as Flores (now part of Indonesia) and East Timor (now Timor Leste). Their cultural, if not linguistic, presence is still felt in the region today, with areas of eastern Indonesia being majority Catholic, and Timor Leste being the highest-percentage Catholic country in Asia (Adams & Gillogly 2011), with other members of the Catholic Church present throughout Indonesia, Malaysia, and Singapore. Apart from the creoles themselves, the Portuguese language has provided a large amount of loanwords in the local languages, such as the Malay-Indonesian *gereja* ("church", from Portuguese *igreja*); *keju* ("cheese", from Portuguese *queijo*); and *minggu* ("week", from Portuguese *domingo* "Sunday") (Mintz 1994).

Portugal's control of Malacca ended when traders from the Netherlands took over in 1641. Although the Dutch traders had political control of the territory, the Portuguese-descended creole community continued to speak their language (Maurer 2011: 3). The Dutch language had little significant effect on the structure or vocabulary of the Kristang language, but some loanwords do exist: *kakus* "bathroom", derived from the Dutch *kakhuis*. There are

also today some Eurasian residents in Malacca who have Dutch surnames, although they speak English or Kristang.

When the Netherlands expanded their sphere of influence in Southeast Asia to eventually take over the Indonesian archipelago, they exported labor there from their other colonies. Kristang speakers from Malacca were bought and sold as slaves in Jakarta from the 15th Century. During the years of slavery and eventual freedom they continued to be a coherent community and spoke their own Portuguese-lexified creole. As Jakarta was within the greater sphere of Malay-influenced Southeast Asia, the main Asian language spoken there was Malay. Portugal never set up a colony on Java, so their language never had the same status there that Dutch did. Portuguese, however, continued to be used as a lingua franca between Europeans, but never had any official status, unlike its role in Macau.

Map 3: The Portuguese Trade Routes (Green) and Spanish Trade Routes (Yellow)[3]

2.6 The Portuguese in Macau

Macau was founded as a trading post in 1557 to link Malacca with Nagasaki, due to its convenient location between the two areas. It was also advantageous to the Portuguese who acted as middlemen for otherwise hostile trade relations between China and Japan (Arana-Ward 1977). The Ming leaders of China agreed to lease the territory to Portugal, but did not permit admixture of the people, and did not allow most of its own citizens to set foot in Macau (Sit et al 2012). Coates (1978: 33) points out that non-Christian Chinese were not allowed to stay a night in the city, which was a rule imposed by both Portuguese and Chinese authority. As a result of this policy, the actual presence of Chinese people in Macau was relatively low.

---

[3] https://en.wikipedia.org/wiki/Portuguese_Empire#/media/File:Macau_Trade_Routes.png retrieved 2019/12/15

Although on the coast of what is now Guangdong, a Yue (Cantonese) speaking area, the main Chinese settlers in early Macau were from Fujian, northeast up the coast from Macau. They were already regular fishers in this area when the Portuguese arrived in the 16th ecntury, and worked with the Portuguese in shipbuilding efforts (Coates 1978: 28-29). Ansaldo and Matthews (2004: 2-3) point out that the Fujianese settlers "were mostly fishermen who occupied a special role *within* the settlement of Macao from very early on, and formed a significant proportion of the 16th-century population." Although their language (Hokkien) is not mutually intelligible with Cantonese, it is a related Sinitic language, and its presence in Macau "nevertheless must be considered a potential factor contributing to Sinitic influence in the formation of Macanese." The non-Chinese population following Portuguese settlement included Portuguese people, but also those from other parts of their empire, including Malacca, who brought their Portuguese-based creole with them. Because of this, Makista has many similarities with Kristang, including considerable influence from Malay (Ansaldo & Matthews 2004, Laub 2014).

The general separation between Chinese and non-Chinese populations in Macau until 1887 resulted in relatively small linguistic influence by Cantonese on Makista. While the influence of Cantonese may have been understated by other scholars, according to Arcodia (2017), the influence of Malay on Makista remains dur to its origins in Malacca, and there are many shared features between Makista and Kristang that can be traced back to Malay origins or influence. For example, one of the most prominent features is the use of pre-verbal markers to indicate aspect, also found in Malay, but not in Cantonese or Portuguese. These markers continued to be used in Makista well into the 20th Century, as evidenced in the work by Santos Ferreira (1967, 1973, 1990).

Although the Portuguese lost Malacca to the Dutch, they held onto Macau for over four centuries. There was never a large Portuguese-speaking population in Macau, but Portuguese remained an official language of government and education. In 1999, Macau was transferred to the People's Republic of China and was given the status of Special Administrative Region (SAR). Similar to the status Hong Kong was given in 1997, this allowed Macau to have a separate legal system while being internationally recognized as part of China. The separate legal system of Macau gives it a market-based economy, as well as legalized gambling, which has become a large source of income for the SAR. Macau's casinos and the expansion of the city by means of land reclamation have led to it being a major hub for tourism, attracting many international and Chinese tourists (Sit et al 2012).

As a Special Administrative Region of China, Macau has its own official languages, Portuguese and Chinese. While Portuguese is still not widely spoken, it remains a large part of the linguistic landscape of the city, as street signs and bus announcements can be seen and heard in Portuguese. Portuguese culture also left its mark in the architecture of the city, most notably the ruins of St. Paul's church in central Macau. The cuisine in Macau is recognizably Cantonese, but many of the local dishes have clear Portuguese influence (Cabral 2016).

As for the actual languages that are spoken in Macau, Cantonese continues to be spoken by the majority. Due to increased connections with Mainland China, Putonghua (or Mandarin) has also become an important language, used for education and government, as well as for communication with Mandarin-speaking tourists. Proximity to Hong Kong has led to English being commonly spoken, especially in the tourist industry, and can be seen in menus and signs across the city. If a resident of Macau can speak a European language, it is more likely that it is English than Portuguese, despite the history of colonization and its status as an official language.

Although there is little Portuguese spoken on a daily basis within Macau, it still has a continuing status that it did not have in Malacca after the Dutch takeover. As an ethnic group separate from the Chinese, the Makista would have been more likely to be exposed to the Portuguese language of the government and other officials. There continues to be a presence of Portuguese, with academic opportunities to learn the language, including at the universities as well as at the Instituto Português do Oriente, while no such institute exists in Malacca. Many of the major differences between Makista and Kristang are features where Makista resembles standard Portuguese more so than does Kristang. An example of this can be found in genitive constructions discussed in Section 6.6.

Trade and colonization by the Portuguese in Southeast Asia were the reason for the genesis of the creoles in the present study. However, they were not the only European powers to have influence in the region, and we will see how other countries, namely the Dutch and the British, had an impact on the political history in this part of the world.

2.7 The Dutch in Southeast Asia

The initial period of European colonialism in the 16th Century was dominated by the Iberian powers. The Spanish and Portuguese were powerful nations with intercontinental empires, prompting other countries to follow suit, tempted by the wealth Spain and Portugal had attained. As the center of pre- and post-European trade in Southeast Asia, Malacca became a

target of the Netherlands. After about a century of Portuguese rule, the Dutch expelled them in 1641 and established their own colony which lasted until 1825 (Parthesius 2010).

The period of Dutch domination saw even more extensive colonization, where the Netherlands eventually took over the archipelago that today makes up Indonesia. They established a city on the island of Java in 1619 on the ruins of Jayakarta, and named it Batavia; this was the base for the Dutch East India Company and expanded into what is now Jakarta. As part of their colonization, they took Eurasians from Malacca as slaves and transported them to Jakarta. The Malaccans were transported in such numbers that they maintained their language and continued to speak it for centuries afterwards. The slaves were eventually freed, and became known as the Mardijkers. A condition of their freedom was that they convert from Catholicism to Protestantism and adopt Dutch names. In the mid-17th Century they moved to Tugu, a rural area just outside of Jakarta (today a neighborhood within the city) (Schuchardt 1891, Huet 1909, Wallace 1978, Maurer 2011).

The creole speaking communities in the Jakarta area can be split into two different groups. The first, speakers of BC, was the group that lived a central area of the city. Their language shows a comparatively large amount of Portuguese influence on the morphosyntactic structure. The other group are those who lived in Tugu; their language was spoken in a less European-influenced area, as is evident in the language's structure and lexicon. Compared to BC, TC shows a much deeper influence from Malay. Although it is still considered to be a creole, there are many lexical items that have been directly borrowed from Malay, while in the other creoles these corresponding terms would be calqued into a Portuguese-derived term or not exist at all. One such example is the conjunction *tapi* "but", which is a direct borrowing from Malay. Another example is the verbal prefix *di-*, which transforms verbs into the passive voice. Unlike other grammatical markers in TC and the other Luso-Asian creoles, this prefix is a direct borrowing from Malay, without any calquing into a Portuguese equivalent (Maurer 2011). BC and TC are no longer spoken. The last speaker of TC died in 1978 (Baxter 1996).

2.8 Historical and Current Demographics

As mentioned in Chapter 1, ethnic Kristang and Makista people are a very small minority within their respective cities, and the speakers of their languages are a smaller subset of that minority. While they have their minority status in common, the actual demography of each city is quite different.

Since the establishment of a Portuguese colony in Malacca there has been a presence of Eurasians of partial Portuguese descent, and the Kristang language that is still spoken today. The Portuguese arrived in an already-established multiethnic trading city and their descendents never numbered larger than a small minority. Demographic information about pre-British Malacca is difficult to find and few sources exist (Baxter 2018: 253), although sources do point to total numbers of Eurasians being low enough to be a small minority.

MacGregor (1955: 6) estimates the maximum number of Portuguese in Malacca to have been 600 during their period of rule. In 1525 there was 38 *casados* (Portuguese settles married to locals), and just over a century later in 1626 their number only rose to 114, of whom 62 lived outside the city walls of Malacca (MacGregor 1955: 12). Eredia's account of Malacca counts 7,400 Christians in the town and 300 *casados* and Portuguese soldiers (Eredia 1613: 20-21). This estimation is typical of visitors and chroniclers at the time who would have counted Christians, rather than the number of speakers of Portuguese or Kristang, and it would be a reasonable assumption that at least half of the Christians counted by Eredia were speakers of a variety on a Portuguese-Kristang continuum (Baxter 1988: 6).

After the Dutch siege and takeover of the city, the total population of Malacca decreased nearly tenfold, from around 20,000 to 2,150 inhabitants (Leupe 1859: 116). Dutch Governor Balthasar Bort described the demographics in Malacca in 1678, about four decades after the siege. The following numbers are also inclusive of those enslaved (in separate column) by each group.

Table 1: Demographics of Malacca in 1678 (Bort 1678: 39-41)

Ethnic group	Total population	Of whom were enslaved
Dutch Burghers	494	349
Portuguese half castes and blacks	2,020	551
Chinese	716	290
Moors and Gentoos	761	214
Malays	768	180
Bugis	125	23
TOTAL	**4,884**	**1,607**

Baxter (1988: 6-7) notes that of the slaves possessed by the Dutch would have been creole speakers, so at this point in time Kristang would have been at its peak as the largest linguistic group of Malacca.

Another Dutch account of Malacca in 1680 describes the Eurasian community's numbers as follows:

Table 2: Demographics of Malacca in 1680 (Fernando 2004: 175)

Ethnic group	Population
European	40
Mixed	61
Black	537
Slaves	529
TOTAL	1,167

The British period contained more recordings of the population and their demographic breakdown. An 1865 account describes the ethnic composition: "But the Portuguese though the most interesting inhabitants of Malacca are not the most numerous. Chinese, Malays, Klings and other natives of India … each equal if not outnumber them; and in the country, few but Chinese and Malays are to be met with" (Cameron [1865] 1965: 376).

The ethnic division of Malacca in 1865 was speculated to be as follows (Cameron [1865] 1965: 109):

Table 3: Demographics of Malacca in 1865 (Cameron [1865] 1965: 109)

Ethnic group	Population
Aborigines	900
Malays	55,000
Chinese	12,000
Natives of India	1,200
Other Asiatics	2,500
TOTAL	71,600

Cameron ([1865] 1965: 109) also mentions that "Europeans and their immediate and unmixed descendants do not, I think, number 800, nearly two-thirds of whom are stationed at Singapore." These 800 were in the Straits Settlements of Malacca, Singapore, and Penang, and the populations in each city were not shown by the author. Little is written

about the number of Kristang speakers who would have been in Malacca at the time, aside from the quote above. They would most likely have been considered members of the "Other Asiatics" group, which numbered a minority of less than ten percent of the total. McNair (1871), cited in Niew (1969: 142) showed the census results of 1871:

Table 4: Demographics of Malacca in 1871 (Niew 1969: 142)

Ethnic group	Population
Chinese	13,482
Indians	3,278
Europeans	50
Eurasians and others	2,850
Malays	58,096
TOTAL	77,756

These results are from the first modern census taken in the area in 1871 under British rule. To conduct the census, including its demographic breakdown, categories were formulated to purposely be "mutually exclusive and exhaustive" (Hirschman 1987: 559). This census included "Eurasians" as a category, but did not include further subcategories such as which specific European and Asian ancestries these groups of people had. No other ethnic classification from this census indicates any type of Portuguese or Portuguese/Malay ancestry, so the Kristang speakers would most likely have been a subset of this category.

The 1881 and 1891 censuses reported the ethnic breakdown as follows:

Table 5: Demographics of Malacca in 1881 and 1891 (Merewether 1892: 29-30)

Nationality	Population per 10,000 (1881)	Population per 10,000 (1891)
Europeans and Americans	4	15
Eurasians	236	190
Chinese	2,110	1,970
Malays and other Natives of the Archipelago	7,415	7,630
Tamils and other Natives of India	202	179
Other Nationalities	33	16
TOTAL	10,000	10,000

The total population of Malacca was reported as 93,579 in 1881 and 92,170 in 1891. The table above reports the population per 10,000 of each of the ethnic groups. (Merewether 1892: 29-30).

While this census makes no specific mention of the Kristang-speaking Eurasians, it is evident that they would not have been counted as Portuguese, as this was a separate category, and of whom 19 were counted (Merewether 1892: 136). Every census while under British rule contained a "Eurasian" category, and none of them had sub-categories as the others did (such as Hokkiens as a subset of Chinese, or Javanese as a subset of Malays) (Hirschman 1987: 571-577).

More recently, the city of Malacca is typical of modern urban Malaysia, being a multi-ethnic and multi-racial society. Figure 1 shows the racial makeup of Malacca according to the 2010 census, based on the following numbers:

Table 6: Demographics of Malacca in 2010 (Malaysia census 2010)

Racial group	Population
Bumiputera	280,336
Chinese	158,828
Indian	20,310
Others	3,240
Non-Malaysian	22,171
TOTAL	**484,885**

Chart 1: Ethnic groups in Malacca City (Malaysia census 2010)

■ Bumiputera   ■ Chinese   ■ Indian   ■ Others   ■ Non-Malaysian citizens
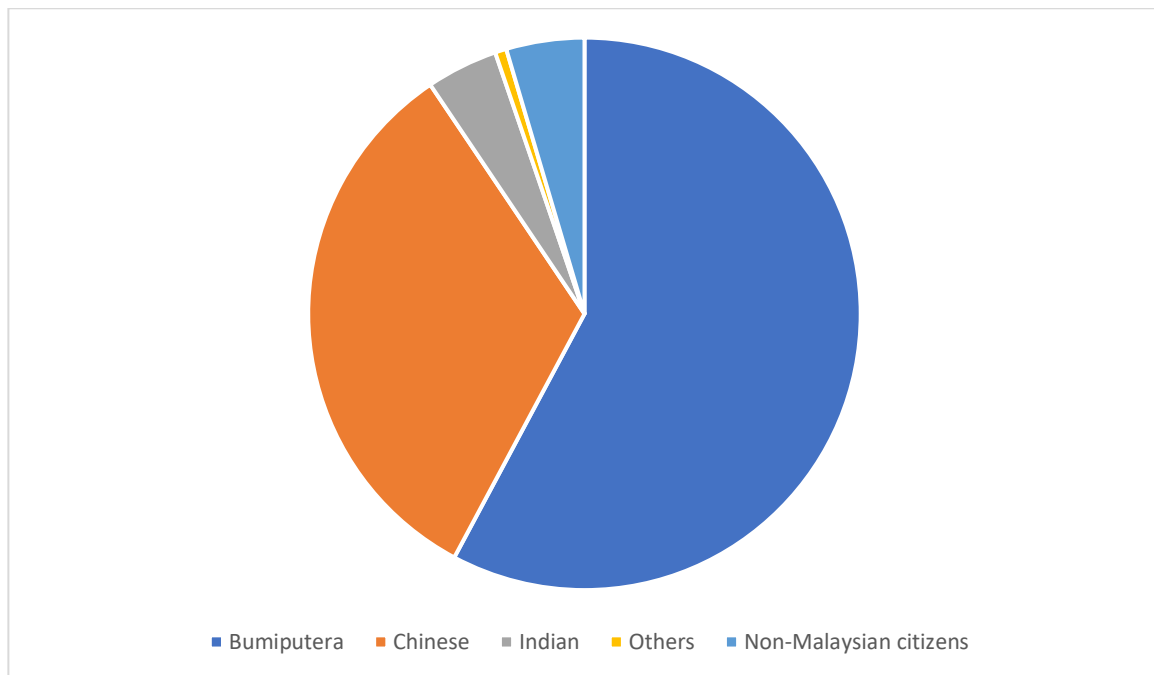
While the Bumiputera (Malays and other groups indigenous to Malaysia) make up the majority at 58%, the other groups comprise a large minority, including the Chinese, who make up one-third of the city's population. The "others" group (which includes Eurasians) is by far the smallest at less than 1% of the total population.

There is no information in the census about which languages are spoken at home, although ethnic data is provided, which can give us an idea about what languages might be in everyday use and by how many people. Figure 1 shows a majority-Malay but diverse city, which can easily be observed by visitors to Malacca. Speaking multiple languages is a part of daily life in the city, not just among Kristang speakers. Ethnic Malays speak Malay at home, while Chinese speak Hokkien or Hakka, and Indians generally speak Tamil. The city also has a population of the Peranakan or Baba Nyonya, ethnic Chinese who settled in Malaysia and Thailand before the European colonial period starting in the 15th Century. They retain elements of their Chinese culture but mostly speak a variety of Malay (Ansaldo et al 2007). Another distinct ethnic group are the Chitty Indians, descended from Indians who arrived before the European period, who also tend to speak Malay (Pillai 2015). The majority of ethnic Chinese and Indian people in Malacca are descendants of those who immigrated during the British colonial period from 1824 to 1957 (Hirschman 1986).

The census reveals that ethnic and linguistic minorities comprise a large percentage of the population of Malacca, resulting in a diverse city. Kristang speakers are a very small

portion of the population and are vastly outnumbered by other ethnic minorities. This contrasts with the demographics of Macau, as shown in Chart 2:

Chart 2: Languages spoken in Macau (Macau census 2011)
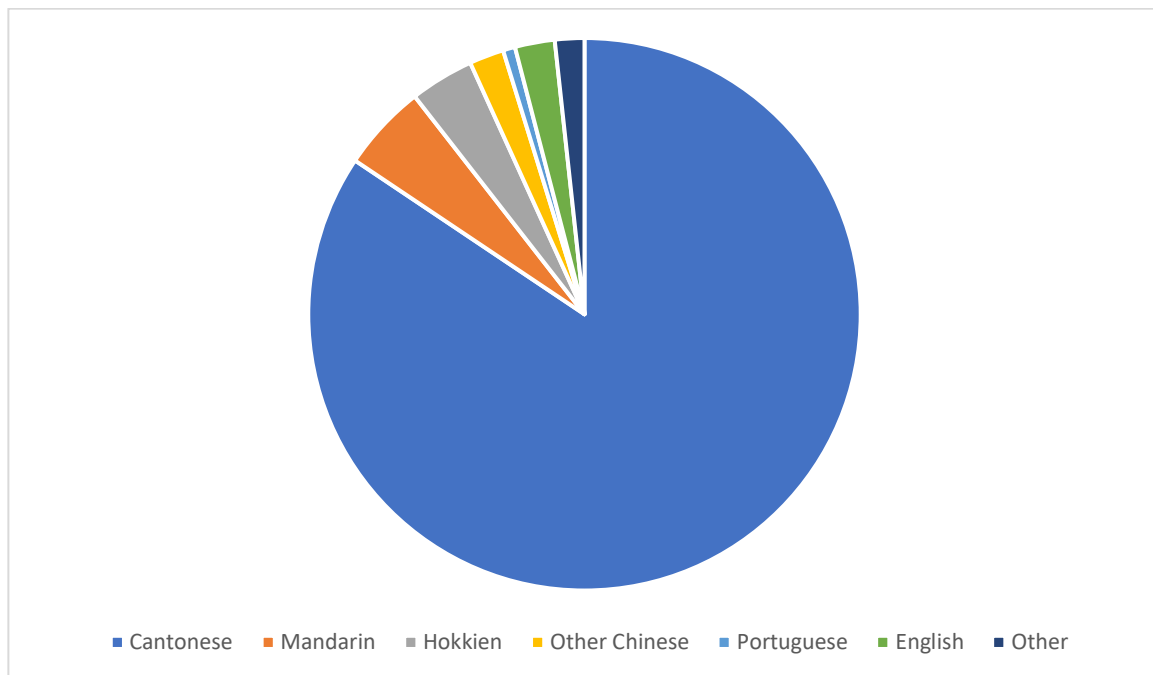
Figure 2 shows the main languages spoken by Macau residents, according to the 2011 census. This reveals a different the picture than in Malacca, since the Cantonese-speaking linguistic majority is much larger proportionally then the ethnic Malay majority in Malacca. If all Chinese languages are combined, 94% of Macau residents speak some Chinese language on a daily basis. The smallest language reported is Portuguese, with 0.7% of Macau residents using it as their usual language. This is less than a third of the number of English speakers, the only other European language represented. With a growing tourism industry attracting visitors from the People's Republic of China and overseas, a knowledge of Mandarin or English is useful for those working in tourism and hospitality. However, the day-to-day language of the vast majority of the population of Macau is Cantonese and other Chinese languages.

The ethnic breakdown of Macau shown in Chart 3 is another reflection of its relative lack of diversity.

The largest ethnic group in Macau by far is Chinese. Compared to Malacca, where the non-Bumiputera population represents 42% of the population, non-Chinese or mixed Chinese are only 6.6% of the population, while all non-Chinese number 7.6%. This proportion, plus the fact that there is no ethnic enclave in Macau, akin to the Portuguese Settlement in Malacca, shows the kinds of interactions Makista people would have in their daily life would be more intercultural, and less intracultural, compared with the Kristang.

The difference in diversity between the populations of Malacca and Macau can largely be attributed to their contrasting histories. As Malacca was a major trading center, it attracted merchants from various nations, many of whom settled in the city. It has also changed hands multiple times through the last half-millennium. Macau's location was less strategic than that of Malacca, and it was politically controlled by the same country for most of the past five centuries. The post-transfer period has seen an increase in development and investment from China, and this has taken place peaceful manner without any major social or political upheaval.

While the presence of the Portuguese in power lasted for centuries in Macau, their actual numbers were always a minority of the population of the settlement. Ansaldo & Matthews (2004: 3), citing Tomás (1988), demonstrated this with demographic information, further shown by census information for the 21st Century.

Table 7: Demographic shift in Macau (Ansaldo & Matthews 2004: 3; Macau census 2011: 11)

	1622	**1878**	**2001**	**2011**
Portuguese	1,000	4,476	2,810	3,485
Chinese	6,000	63,532	416,353	510,383
Others	N/A	122	16,072	37,194

The centuries-long decline in the proportion of Portuguese in Macau can be further illustrated by also considering the 21st-Century data from above:

Chart 4: Proportion of Chinese and Portuguese in Macau, 1622-2011
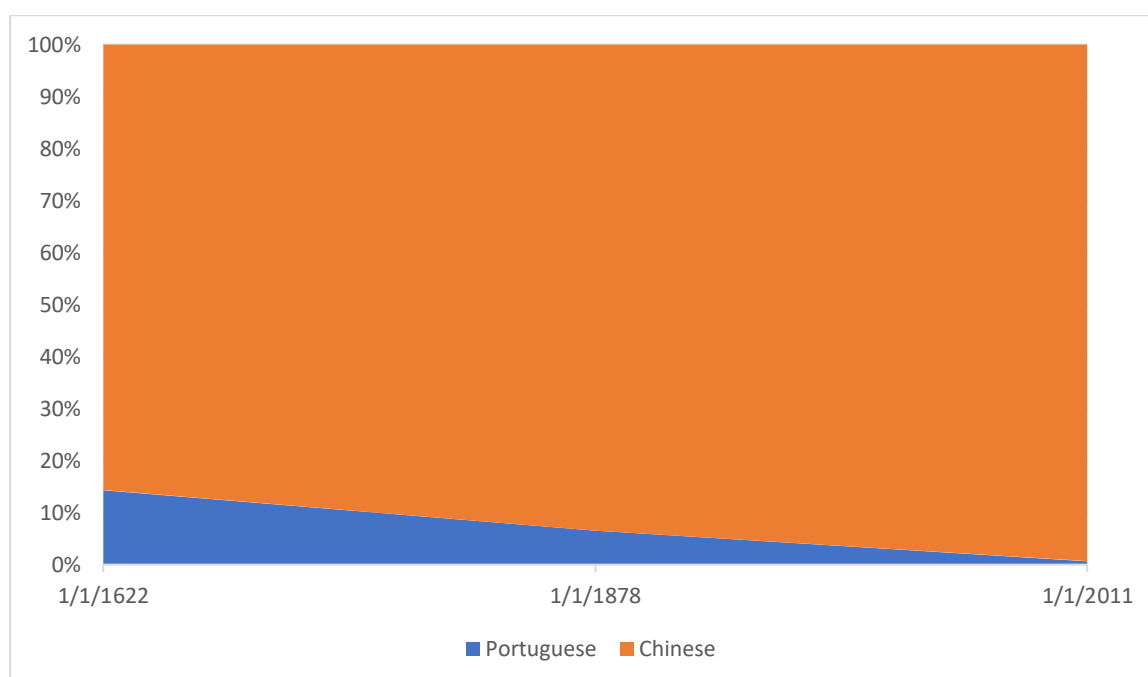
This chart shows the proportion between those listed as "Chinese" and those listed as "Portuguese", using the data from 1622, 1878, and 2011. The 17th- and 19th Century data only listed three categories: Portuguese, Chinese, and other, while the 2011 census data had categories including "Portuguese and Chinese". For the purpose of the chart in Figure 5, this only includes those listed as "Portuguese" and "Chinese" in the 2011 census.

The descendants of BC and TC speakers still live in the Jakarta area. Although their language is no longer spoken, they still use it in song and in greetings, such as *Bong anu nubu* "Happy New Year". Other aspects of their culture remain, such as their Christian faith and their music, called *keroncong* (La Batu 2016, Ulung & Ali 2019). Java was never a colony of the Portuguese and this could be a reason for the declines of the creoles that were spoken there, with the last speaker having died in the mid-20th Century. Jakarta is also by far the biggest city of the three, with a population in 2019 surpassing 10 million.

Malacca eventually became overshadowed by Singapore and Kuala Lumpur, while Macau has the much larger neighbors of Hong Kong and Guangzhou. The BC and TC speaking groups have become a fraction of a percent of the overall population of Jakarta.

2.9 Conclusion

The histories of Malacca, Macau, and Jakarta can reveal to us aspects of the linguistic ecologies of these cities. The creole languages spoken there all arose from contact between different groups speaking their own languages, but their trajectories were markedly different following the initial point of Portuguese contact. In Malacca the descendants of the first Eurasians still speak Kristang, alongside a majority-Malay population in an ethnolinguistically diverse city. This language was still spoken throughout centuries despite colonial powers repeatedly usurping one another, and continues to exist in the modern independent state of Malaysia. Macau is a city that has had centuries of stability and a largely homogeneous (ethnic Chinese) population. It is the only city of the three where Portugal held political power continuously, and hence the one where Portuguese has had the strongest influence. Macau and Timor Leste are the only Asian territories that have Portuguese as an official language. Just as the Macau SAR is dwarfed in size by the rest of China, Eurasians and Makista speakers are a tiny proportion of the population of mostly Cantonese speakers. The speakers of BC and TC were forcibly taken to a city that had never seen Portuguese control and have a history of forming their own community after gaining their freedom, while still a part of the Dutch Empire. On the outskirts of Jakarta, TC speakers had little contact with Europeans, compared to communities in the center of the city. The same factors that led to the heavy Malay influence on TC also resulted in its decline, and both Jakarta creole varieties are no longer spoken at all.

Chapter 3: Literature Review: Creolistics

3.1 Introduction

This chapter is a literature review that describes previous works in the study of creole languages, a field referred to as creolistics. Since the recognition of creoles as a distinct subset of languages, it has been the aim of creolists to describe them. When two (or more) languages come together and a resultant pidgin or creole emerges, there will be some interference in its transfer to new generations of speakers. This interference can result in changes to the language from one generation to the next, up to the point where a new language emerges. However, there will remain syntactic features or lexicon (or descendents thereof) in the emerging creole.

3.2 Language Contact, Pidgins and Creoles

While looking at pidgins and creoles, especially their structures, it is essential to consider the implications of language contact and contact-induced change. The study of language contact is not limited to that which results in the creation of creole languages, and there are indeed phenomena in non-creole language contact that have similarities to the kinds of changes that result in development of structures of a pidgin or creole.

Weinreich (1953) was a pioneer in the study of language contact, who presumed that language contact took place where there were bilingual speakers giving rise to interference that different languages provide in their speakers' interactions. Relative homogeneity of outcomes derives from the social aspect of language contact and change is explained: "[W]hen a group of some size brings two languages into contact, idiosyncrasies in linguistic behavior tend to cancel each other, while socially determined speech habits and processes characteristic of the group as a whole become significant" (Weinreich 1953: 83). Weinriech's views on the role of society in language change would become significant in later studies of pidgins and creoles.

Thomason (2001) provides a thorough introduction to the study of language contact in general, not just in pidgins and creoles. In terms of the situations from which language contact can arise, she lays out four possibilities concerning the social positions of speakers, and how each one can affect the outcome of language change: (1) indigenous superordinates, who will rarely change their language; (2) migrant superordinate; (3) indigenous superordinate; and (4) migrant subordinate (Thomason 2001: 23). This is an interesting typology, although it might not be detailed enough for our purposes; the Portuguese

colonizers, for example, did not arrive alone in Malaysia and Macau, and often brought people speaking non-standard and creolized versions of their language with them.

Thomason also points out seven mechanisms for language change: code-switching, code alternation, passive familiarity, negotiation, second-language acquisition and deliberate change (Thomason 2001: 129). One major theme running through these mechanisms is bilingualism, or at least attempts to become proficient in another language in the case of second-language acquisition. Matras (2010) goes further and includes a continuum of volition in language change, from completely unconscious interference to deliberate change (Matras 2010: 67). Matras argues that language change is essentially a functional phenomenon, whereby any language change, regardless of whether it is deliberate or not, is made with the ultimate goal being communicative.

These concepts that see language change as a functional phenomenon are especially important in the study of creole languages. As various Luso-Asian varieties have developed out of Portuguese traders and colonists settling and intermarrying with local women, linguistic changes arising from this contact are primarily due to their use for communication between groups of people from different backgrounds. The languages that helped form Kristang, Makista, BC, and TC include the ones that were spoken around the cities of Malacca, Macau, and Jakarta; in Part 2 we will see how that had an effect on the structure of the languages.

3.3 Creole Exceptionalism

The Language Bioprogram Hypothesis (LBH), proposed by Bickerton (1981, 1984), argues that all creoles belong to a typological category distinct from other languages, regardless of superstrarte and substrate input. Citing Chomsky's views positing a language organ present in the brain, Bickerton argues that LBH can explain the structural similarities of creole languages, regardless of their linguistic sources. Using evidence from two English-lexified creoles in Hawai'i and Guyana, he identifies features that are present neither in English nor in their respective substrates. For example, there is a feature termed by Bickerton as *fu*, based on a word in Saramaccan creole from Guyana (this has its lexical origins in English *for*; Portuguese-lexified and French-lexified creoles have similar words derived from *para* and *pour*, respectively). Looking at a variety of creoles around the world, he shows how *fu* is employed to affect the tense, aspect, and mood of verbs, and how in Saramaccan it can be tensed itself. This commonality, he argues, is evidence for the claim that creoles have been created within a framework of universal grammar present in all humans. As these creoles are

spoken in widely separated locations, similarities between them it cannot be a coincidence but must be due to cognitive capacities shared by all humans.

The LBH proposes that rapid creation of a creole language results in features that depend on the language organ, this view has been rejected by other scholars such as McWhorter (2005); some opponents of LBH do believe that creoles have certain features that set them apart from non-creoles, and that these can be predicted. McWhorter (2005) laid out what he called a "creole prototype", which is characterized by: (1) lack of inflectional morphology; (2) lack of tone on monosyllabic words; and (3) lack of semantically opaque word formation. This approach to creolistics has come to be known as **Creole Exceptionalism** (see also McWhorter 2018).

McWhorter (2001, 2005), Parkvall (2008), and Bakker et al (2011) also support creole exceptionalism, agreeing that, creoles are typologically distinct from non-creoles, regardless of the lexifier or substrate. Using a quantitative method, Bakker et al (2011) scored different creoles against each other by degrees of creoleness based on features selected by Holm & Patrick (2007). The different variables compared include lexifier language, geographical area, type of creole (including plantation, trade, maroon, and others), and the age of the language. The resulting intersectional groups were shown to contain members that have rankings ranging from high to low creoleness scores. Bakker et al (2011: 25) point out that "there is no significant relation between the degree of creoleness [ie the amount of typical creole features in a given language] … and the kind of lexifier, the kind of socio-historical situation, the area or the age of a given creole". Thus, they argue the similarities that all creoles have with each other must come from something else. They then compared the creoles' scores to non-creole languages with similar structural complexity to creoles, looking at relatively isolating languages such as English, Mandarin, and Indonesian. Using an algorithm to analyze and compare typological features based on the *World Atlas of Language Structures*[4], they showed that the selected creoles clustered around the same group. Only one non-creole, Hmong, was found to fall among the creoles, and no creoles were found separate from their fellow languages (Bakker at al 2011: 32). Although their selection of creoles is diverse, we can identify are some limitations in their selection. For one, only two of the eighteen come from the Pacific region, and twelve are from the Atlantic region. All but two of the creoles were lexified by languages of European origin, and only one was lexified by a non-Indo-European language (Nubi, by Arabic). In response, Bakker (2014) explains that most known creoles are

---

[4] https://wals.info/

lexified by a European language, so the original sampling does not suffer from problems due to this.

One issue with this methodology is that each language is reduced to a single data point. It does not take into account the variation that a creole may have (or have had) over geographical space or diachronically. It also ignores the fact that some creoles can split off from another, as Makista did from Kristang. Of course, Bakker et al (2011) dealt with massive amounts of data and focused on trends, however in such an approach nuanced differences between the studied languages can be lost. The idea that creoles share features despite geographic and temporal distance has given rise to the idea of a "virtual sprachbund" (Kihm 2011). Kihm (2011) specifically looked at the Portuguese-lexified Kriyol of Guinea-Bissau and the Arabic-lexified Nubi of Uganda and Kenya, two languages whose origins and development were unrelated and which did not have any known contact. Kriyol and Nubi are shown to have similar features, none of which found in their lexifiers. Rather than having origins in the internal language organ that the LBH proposes, the similarities that can be found among creoles are not unlike similarities that occur between languages in a geographically-defined sprachbund. Kihm argues that what gives rise to the typological similarity among creoles is the common origins they have had from pidgins.

3.5 Responses

Youssef (1988, 1990) criticizes Bickerton's argument for LBH, challenging his notion that creoles develop in a situation where speakers use their language organ to create a grammar in the absence of sufficient input, ie out of a pidgin. Yousseff (1988: 452) argues against this point saying that he presumes creoles to form in a "linguistic vacuum, denying the well-established fact that language and its acquisition are social phenomena."

Siegel (2007), while giving credit to Bickerton for increasing interest in creoles within linguistics, also challenges the LBH, drawing evidence from Hawaiian creole. For example, one feature that Bickerton contended was likely to be present in creoles and to originate in the Language Bioprogram was the marking of tense, mood, and aspect (TMA) in verbal phrases. Bickerton (1984) claims that these markers are formed preverbally and always occur in the order T-M-A. Siegel (2007) provides evidence to the contrary, such as the expression of anterior tense in Hawaiian creole, which he shows comes from the preverbal markers *bin* (from English *been*) and *wen* (from English *went*). Siegel (2007) also points out that some verbs can be directly inflected for tense, including *sed* from the irregular English past tense

form *said*. Another example is a shift from the progressive marker *stay*+V to *stay*+V-ing, showing additional verbal morphology that the LBT would not predict.

In addition to the typological arguments that creoles are not distinct, there have been critical looks at the field of creolistics itself. Ansaldo & Matthews (2007: 3) have also sought to "overcome the artificial dichotomy between creole and non-creole languages" by looking at creolistics in a larger context, i.e. that within the studies of language creation and language contact in general. DeGraff (2003, 2005) also argued against the notion of creoles being a group to themselves, suggesting the need for a social and political point of view. Preferring to use the label "creole" as a sociohistorical one outside of language to describe the mixing of European and African languages following colonization and slavery in the Americas, he proposes that the claimed linguistic distinction has its roots in scholars' prejudiced viewpoints of non-Indo-European languages being less evolved and therefore inferior. DeGraff (2008) further argues that creoles are not necessarily distinct from non-creoles, and that the claim of uniqueness can be dangerous, especially where education is concerned. He cites examples such as negation from his native language, Haitian Creole, and compares them to French and English: "Modern English is like Haitian Creole in having its negation marker and certain adverbs occur regularly in the position that *precedes* the verb … Yet, Modern English, like Haitian Creole, has an ancestor where these adverbs and the negation marker regularly occurred to the *right* of the finite verb" (DeGraff 2008: 135). The parallel histories of Haitian Creole and English suggest that it is a misnomer to argue there is something different or special about changes that occur in "creolization" for languages that have emerged from colonial contexts.

Mufwene (2000) also argues that creolization is a social process, rather than a structural one. This is in opposition to the position McWhorter and others that creolization is a particular process that can result in predictable outcomes. Part of Mufwene's argument is that the language acquisition phase in the creation of a creole does not take place in a vacuum, and is quite different from that of formalized language acquisition in an educational or institutional context. Mufwene (2000: 69) notes that "unlike in the scholastic setting, nobody teaches others the system of their language; every learner acquires (some of) the vocabulary to which they have been exposed and infers a system that enables them to communicate." Thus, the output is the result of the communication between different groups of people, and one which depends upon the particular social context. Thus, children of Portuguese sailors and Malay traders will grow up and acquire language in a very different context than West African slaves on a Caribbean plantation.

It is vital to consider these various viewpoints in the investigation of the creole languages within our case study. Investigating and comparing the structures of Kristang, Makista, BC, and TC to each other and to Portuguese, Malay and Cantonese can serve as a test bed for various hypotheses proposed in the literature. Recent typological studies (see Cardoso et al 2012) have shown how the studies of specific languages can fit into a much larger picture, and this thesis aims to address this as well.

The islands, plantations, and colonies where creole languages developed can be seen as a kind of language laboratory, where we can see what happens when a new language is created. However, the reality is that each colony was unique, their foundations depending on many factors, such as the presence or absence of slavery, the nation of origin of the colonizers, the ultimate fate of the indigenous populations, the number of European settlers, and many others. Chaudenson (2000: 362) points out that the creation of a creole is a very complex process, one in which there are "more exceptions than rules"[5]. The particular sociopolitical circumstances in which creoles are born are an integral part of their identity, not only linguistically but culturally as well. Therefore this study pays particular attention to the demographics and political histories of the cities in which Kristang, Makista, BC and TC were spoken. Although it is not possible nor practical to elaborate every individual difference and complexity that existed throughout their histories, a socio-political approach can make it easier to reveal the origins of certain similarities and differences between the languages.

DeGraff (1999: 6-7) proposes there are three main camps of creolists: universalists, substratists, and superstratists. Universalists (such as Bickerton 1981, 1984, 1999), mentioned above, argue that creoles are structurally distinct with little significant influence from either the lexifier or substrates. Substratists, such as Lefebvre & Lumsden (1989) and Koopman (1986) argue that the most important element in the creation of creole syntactic structures is the substrate. Although the superstrate provides the bulk of the lexicon, and often the phonology, in this view it is the substrate that has the most input into the resulting syntactic structure. Superstratists include Chaudenson (1979, 1992), Hall (1966), and Valdman (1978). DeGraff (1999: 7) describes the superstratist position that creoles "should be viewed as having evolved from (nonstandard) versions of their superstrates." It is through successive waves of acquisition of the lexifier language that it evolves into what we now would call a creole.

---

[5] "plus l'exception que la règle", my translation

Although DeGraff makes a three-way contrast here, I note that the substratists and superstratists cluster together in their views that the input languages have a strong effect on the resultant creole, with the universalists standing apart. The approach I take is more of a substratist viewpoint, although I believe that the lexifier (Portuguese in the case of this thesis) does have an effect on the structure of the languages as well. One such instance, which will be further discussed in Chapter 6, is the syntax of the three main TAM markers in the creoles. Two of them are derived from Portuguese adverbs (the perfective marker from Portuguese *já* "already", and the future-irrealis marker from Portuguese *logo* "later"), while one of them is from a Portuguese verb (the progressive marker from Portuguese *está* "is"). In Kristang and Makista, adverb-derived markers may occur in sentence fragments without a verb following them, but in Kristang the verb-derived marker must be followed by a verb (there were no instances of this in the corpus of Makista either). This particular instance is notable in that pre-verbal TAM markers are not something that occur in Portuguese, but they are found in Malay, as well as all four creoles. In this sense, the output structure resembles the substrates more. However, the difference in outcomes for the adverb-derived markers and the verb-derived one does suggest possible Portuguese influence.

Selbach (2008: 30) argues that we should divorce the terms "superstrate" and "lexifier" as the former is more descriptive of social relations, while the latter is more descriptive of language structure. Although I do use the term "lexifier" to refer to Portuguese throughout the thesis, and "substrate" to refer to Malay and the other input languages that contributed to the resulting Luso-Asian Creole structures, I agree that we need to make this distinction and not let "superstrate" and "lexifier" become synonymous. These terms are revealing in the underlying ideology that has long been present in creolistics, that there was a dominant and a subordinate class of people who came together to create a language. It also helps to focus the idea that creoles are viewed as distinct because of their special social circumstances, and structural descriptions or generalizations came after this distinction was made.

The language ecology theory is also important in the study of creoles, and especially their origins and structures. Ansaldo (2009) touches upon this in his volume, which will be discussed further in the next chapter, where we look at literature related to Asian and Luso-Asian creoles. Aboh & Ansaldo (2007) propose that rather than seeing the formation of creoles as emergence out of a simplified language, or one that resulted in a break of transmission, we should rather consider them as languages that collect features from the languages that are surrounding them, what they term the Feature Pool (FP). An emerging

creole language will not take items fully from one language and incorporate them into its syntactic system, but will absorb features from both lexifier and substrate. Sometimes this happens in full, but it often involves a mixture of the features from each language. For instance, in the English-lexified and Gbe-substrate creoles of the Caribbean, language structures include lexical and functional features from both English and Gbe.

In many creoles, including those selected for this study, many of the function elements have their origin in words with heavier lexical content. This includes the Kristang future aspect marker *logu*, which derives from the Portuguese adverb *logo* "later". Aboh & Ansaldo (2007: 40-1) point out that the shift to a creole language is not simply a process of simplification of features from the lexifier to the creole, but rather a competition between features that exist in the FP. Their framework states that "only features with semantic content enter this competition" (Aboh & Ansaldo 2007: 41) We will see in Chapter 6 more about how this happens in the selected creoles under study, but it is apparent that this type of competition is relevant for the four Luso-Asian creoles in this thesis. As mentioned earlier, the TAM markers are originally from Portuguese adverbs and verbs. We also can find many serial verb constructions where the minor verb in these constructions functions in a way that in Portuguese would involve verbal morphology (see Chapter 6, Section 2).

With this in mind I take a viewpoint somewhere between the substratists and superstratists. There are many structures in the creoles that so resemble the substrate languages that their influence must not be discounted, but influence from Portuguese can be found as well.

3.5 Conclusion

This section of the literature review shows how opinions have long been divided within the field of creolistics. Proponents of creole exceptionalism tend to be data-driven and there have been many articles, chapters, and books looking into the structure of individual creole languages, identifying similarities and seeking for the origins of such similarities in the origin of the languages themselves, via the processes of pidginization and subsequent creolization. Although my thesis argues against creole exceptionalism, I must acknowledge this strand of research without which I would not be able to conduct my research. Despite disagreements I share their enthusiasm for the advancement of creole studies.

Creole languages arise out of socio-cultural contexts that have an exceptional amount of inter-ethnic mixture. Although the lexifier may be spoken by the more dominant group in terms of politics and power, there is also evidence that there is heavy cultural input from both

the culture of the lexifier as well as that of the substrate. As we have seen in Chapter 2, the Portuguese had been in Malacca for about a century by the time they left, sufficient time for multiple generations to acquire and pass on their knowledge of Portuguese, Malay, and what would eventually become Kristang. The emergence of a creole language stems primarily from social mixing, and language acquisition is only one small part of that. As significant social mixing occurred in Malacca and the descendants of the Eurasians continued to live in Macau and Jakarta, the societal impacts endured. It is obvious that creolization played a part in the structuring of these languages—there would be no such languages as these four without it—but subsequent language contact that occurred in the four different communities continued to shape them and led to their divergence.

It is the ideas of creolization as a social phenomenon reflective of the particular socio-cultural, historical and language ecology that informs the structure of this study. As I am comparing the similarities and differences between Kristang, Makista, BC, and TC, I will note the societal and linguistic contact impacts that have led to certain morphosyntactic features of these languages that set them apart from each other.

As we will see in following chapters, I take a mixed approach that acknowledges the importance of substrate as well as lexifier influence. One major factor in the historical difference between the languages is the presence of the different substrates (Malay, Cantonese, and others). As for lexifier, Macau and Jakarta in particular contrast with each other, given the centuries-long rule of Macau by Portugal, and the fact that Jakarta was never under direct Portuguese control.

Chapter 4: Literature Review: Luso-Asian Creoles

4.1 Introduction

Creoles can be found all over the world and can have emerged out of any lexifier/substrate combination, and as seen in the previous chapter much has been written about them, especially for creoles spoken in areas where there are large populations of speakers of such languages. Luso-Asian creoles form a subset due to their lexifier and the location they are spoken in, but they also have a particular history distinct from the formation of other creoles. Unlike many of the Atlantic creoles, Luso-Asian ones came to be spoken in largely urban areas without a populous Portuguese presence. This can be attributed to the history of the Portuguese Empire and the ways they set up colonies, as explained in Chapter 2. The substrates of these languages therefore remained in close contact with them, and their influence can be seen today. I will review some of the previous work done on the Luso-Asian creoles, and I will also set out my contribution to this sub-field. There is some overlap between corpora and descriptive and analytical work on these languages, especially in the case of Santos-Ferreira's Makista literature and Maurer's texts and grammar of BC and TC. This chapter will only address the contributions to linguistics, and a discussion of corpora and data sources will be in the following chapter. As a comparison between Kristang, Makista, BC, and TC this thesis adds a comparative viewpoint to the literature on Luso-Asian creoles.

4.2 Literature on Portuguese Asian Creoles

An early scholar in the study of creoles, including Southeast Asian Portuguese-lexified creoles in particular, was Hugo Schuchardt (1883, 1889, 1890). Working in correspondence with local priests in modern-day Indonesia, he was one of the first to describe creoles in the area. Contrasting with the belief of Coelho (1881) that languages which would later be termed creoles were an imperfect and incomplete learners' version of the European vernacular, Schuchardt argued that these languages had systems of their own and were an "intentional creation by European colonists out of necessity" (Hancock 1975: 212-213).

Ansaldo (2009) is an important volume on the linguistic history and landscape of Southeast Asia. Coming from a functional perspective, Ansaldo argues that the social aspects of the situations in Portuguese Asia are what led to the emergence of the creole languages. The mixing of cultures that was already present before the Portuguese arrived, plus the addition and influence of the Portuguese gave the environment ample opportunity to provide lexicon and structure that aided in the creation of these creoles. Ansaldo (2009: 5) argues that

contact language formation "is not the consequence of a negative process [ie, a broken transmission of language (see Bickerton 1981, McWhorter 2001a, Bakker et al 2011)] but rather of a constructive process involving speakers' agency." He also proposes that "it is the communal, societal output of language that we need to concentrate on" (Ansaldo 2009: 12) in the study of contact linguistics, because society played a very important role in the development of grammatical features.

The studies in Cardoso et al (2012) provide a good framework and methodological model for the comparisons of languages and Portuguese creoles specifically. Several of its eleven chapters focus on a feature or features and look at its function across various Portuguese creoles in India and Southeast Asia. Smith's chapter in Cardoso et al (2012) "Measuring substrate influence: Word order features in Ibero-Asian Creoles" in particular takes a quantitative approach that helps to address the question of how and to what extent the substrates in different creoles influenced word order. He selected nine features in the word orders of Portuguese-lexified creoles in India and Southeast Asia, from a questionnaire in the *Atlas of Pidgin and Creole Structures*. From there he formulated a metric to give each creole a score based on the data that reflects the extent of substrate influence on word order. Baxter & Bastos (2012) wrote about post-possessor genitives in a range of Portuguese creoles spread over a large geographic area. They contend that this feature has its origins in Dravidian languages in the south of India, and was brought east to Malaysia, Indonesia, and Macau where the structure was enforced by similar features found in Bazaar Malay and Cantonese. In the same chapter, they look at the phenomenon of post-possessor and pre-possessor genitives in Kristang specifically, which has interesting patterns in regards to the semantic type of the noun that is being modified. While there were no absolutes, the post-nominal genitive was preferred in some categories, and preferred more strongly in some than in others. In Laub (2014), I examined the same categories in Makista, and found comparatively very few examples of the post-possessor genitive. However, the places where they did occur did follow the preferences that Baxter & Bastos (2012) identified in Kristang, suggesting that Makista has remnants of a similar genitive system. These features will be further discussed in Section 6.7.

4.3 Literature on Kristang

Hancock (1975) sets out the roots of Kristang in European, African and Asian languages. In addition to the obvious connection between Asia and Europe in the creole of Malacca, Hancock argues that some of its features are ultimately derived from African languages. One

example of this is the linking particle *ku* that is ultimately derived from the Portuguese *com* ("with"), as in *yo mira ku eli* "I looked at him". This is found in other Portuguese-lexified creoles, including Papiamento, spoken in the Caribbean, which has no Asian connection but does have African influences on its structure. The early Portuguese trading empire was a string of colonies located around the African and Asian coasts. As Malacca was one of easternmost, there were not only Portuguese people there but also peoples from Africa and India accompanying them. This multinational group of people were from various linguistic backgrounds, meaning that they not only brought Portuguese, but brought their own pidginized versions of it as well. If we consider the Portuguese only being the superordinates, this complicates the four possible social positions in language contact situations from Thomason (2001) (see Section 2.4). The immigrants in this case may well be speakers of the subordinate or superordinate languages, especially in their relations to the indigenous Malaccans, which explains the existence of the various features that Hancock illustrated.

Baxter (1988) provides a comprehensive grammar of Kristang. In addition to basic information about the structure of the language, he also posits sub- and superstrate origins for some of the lexicon and structure of the language. The book also contains a brief history of the area and how Kristang began to be spoken there, as well as predictions of its future. Baxter (1988: 218) also makes the interesting point that due to their location in Malaysia, where Malay is the official language, decreolization is taking place in a direction different from most cases, ie away from the lexifier. He also makes suggestions for how the community could work on language maintenance, including uses in church or in school.

4.4 Literature on Makista

Previous studies of Makista have taken a similar approach to those of Kristang, looking at corpus data and analyzing structures and features of the language in regards to their function and their substrate or superstrate origins and influence. Ansaldo & Matthews (2004) explored reduplication, a feature present in both Malay and Cantonese, and its use in modern Makista. Laub (2014) examines this as well, and I argue that there are reduplication patterns resembling those of Malay in the Makista of the corpus of works by Santos Ferreira (1973a, 1973b, 1990). Although there are features resembling Cantonese as well (Arcodia 2017, Lebel 2018), it appears that the initial Malay influence in the reduplication had a more profound impact, as the patterns of reduplication in Makista are more similar in form compared to Malay than to Cantonese. Further discussion of reduplication in the four creoles can be found in Section 6.8.

Santos Ferreira's poetry and prose is an invaluable source of data for research into Makista, especially since fieldwork is not a viable option due to the moribund state of the language in Macau, but it does not provide complete coverage language itself. Arana-Ward (1977) looked deeply into many features of Makista, based on fieldwork with speakers of Makista living in Hong Kong. There were some differences in Arana-Ward's materials and that which is found in Santos Ferreira's literary works; one major feature where they differ is reduplication. Santos Ferreira contains a very large amount of it, while Arana-Ward's informants mentioned that it was an obsolete feature (Arana-Ward 1977, 91). This will be discussed further in Section 6.8.3.

Pinharanda Nunes (2008) studied demonstratives in Makista, using data from various corpora. His aim was to compare the sources to each other, to the adstrate languages, as well as to other Portuguese-lexified creoles. An interesting feature he points out is that demonstratives in Makista are derived from the Portuguese demonstrative form plus *unga*, originating in Portuguese *um/uma* "a(n), one". He compares this with data from Kristang, Sri Lankan, and North Indian Portuguese creoles, which show a different form, not employing *unga*, which is found only in Makista. In addition, he shows that both Malay and the Chinese languages of Hokkien and Cantonese have a similar form to the demonstratives shown by Makista. Both Malay and Chinese use numerals and classifiers, a feature that Standard European Portuguese lacks. Further evidence of Chinese sources of influence like this is shown in Hong Kong English Pidgin, which has a demonstrative form with a numeral present. Pinharanda Nunes (2008: 9) terms the demonstrative without a numeral as a "semi creole" feature and argues Makista uses both forms.

Pinharanda Nunes & Baxter (2004) examined Makista pre-verbal markers with a lexical origin in Portuguese. Similar to the other studies, they investigate Portuguese-based words or morphemes and analyze their use based on a corpus, in this case various sources of data from different times dating back to the 19th Century, as well as contemporary oral and written data sources. The morphemes they analyzed were pre-verbal tense-aspect-mood (TAM) markers, whose structural behavior shows similarities with Malay and Cantonese. They also noted that there were instances where a pre-verbal marker's presence could have been grammatical, but the speaker or writer chose to omit it. For instance, the particle *já*, lexically derived from the Portuguese word for "already", may occur in Makista to mark the past tense or perfective aspect, but there are many instances of past tenses with just a bare verb. They propose two hypotheses: (1) the optionality of the particles is a sign of

decreolization; and (2) the particles are influenced by Cantonese, which also has optional aspect particles in the verb phrase.

4.5 Literature on Batavia and Tugu Creoles

Because BC and TC both became extinct before the wide availability of recording equipment, there are no audio records of them, to my knowledge. Like most of the Makista data presented in this thesis, the information on BC and TC come from written sources only. Maurer (2011) provides a thorough grammar based on the small amount of texts as well as a compilation of glossed example sentences collected in BC and TC. In addition to the glossing in English there are the original translations into Dutch, German, or Malay. My examples in BC and TC in this thesis are glossed according to Maurer (2011), with some exceptions. For example, for consistency, I have glossed the BC and TC perfective particle *dja* as "PFV", the same abbreviation that I have chosen for the Kristang and Makista examples. In addition, translations of examples may be adjusted to put them in context; this was not necessary for the majority of examples.

Due to the written nature of the sources, it is impossible to test grammaticality judgments in the way I have been able to with Kristang. Although some of the texts include a dialogue, most are narratives or stories, which also affects what kind of constructions occur.

4.6 Contributions to the Literature

This thesis is meant to be a continuation and extension of previous work in Luso-Asian creoles through a comparison between four related varieties. Much of the previous work on these creoles takes a descriptive approach, and I wish to draw on the various descriptions for a cross-linguistic study. I also contribute to the existing data in Kristang through my own fieldwork on issues relevant to this thesis, with the goal of providing a deeper understanding of these languages which are either extinct, moribund or endangered and at risk of being lost. This research also contributes to creolistics in general, as an account of how the sociopolitical history of a place can have an effect on the structural aspects of a language. DeGraff (1999: 36) notes the need for studies such as this:

> Given this volume's bias (toward internal factors in morphosyntactic
> development), the findings described in the following chapters urgently need
> to be complemented with facts outside of syntax and outside of grammar—for
> example, facts of a sociohistorical, demographic, and cultural nature, as
> advocated by Lightfoot (and others).

This thesis thus aims to address the decades-old need and desire within the creole studies community to tie demographic and sociohistorical facts to morphosyntactic and other structural ones. We are primarily concerned with the interactions of demographics and language ecology on the structure of languages. We will see this in the methodology (see Chapter 5) and in the data analysis (see Chapter 6). In addition to comparing the data from creoles against each other, I also compare them with the languages with which they came into contact.

Language endangerment is another important factor in the choice of languages studied in this thesis. Kristang is the most populous by far of the four in this study, however it is still endangered, with younger generations shifting to English at home. Makista is a moribund language, although performances are conducted in it. BC and TC are no longer spoken. Cardoso (2012: 117) points to the importance of studies like this one, which focuses on small languages:

> Given that this study is based on a comparative analysis of a single construction… my generalisations must be taken as hints rather than holistic, established facts. … Hopefully [this study] can provide a framework for future comparative studies and … produce a more solid picture of the Luso-Asian Creoles, their sources and their inter-relationships.

Smith (2012: 128) also comments:

> [One of the concerns] is the availability of data. Some languages are well described, others less so. So, even if it were possible to arrive at a set of unbiased features, there is no guarantee that data on these features would be available for all the languages under consideration.

And finally Sippola (2012: 150) notes:

> The limitations of this study stem from the scarcity of the data and the generalizing bias of typological studies. Although there are a growing number of detailed descriptive grammars of the Ibero-Asian creoles, these seldom pay much attention, if any, to the range of functional properties of indefinites.

Much of the previous literature that has informed this study involves comparisons between creole languages. The choice of comparands is not arbitrary, and usually is because they have common origins as creoles (such as in Holm & Patrick 2007) or a narrower focus (as in Cardoso et al, 2012). The significance of the present study arises from the fact that the four creoles are very closely related to each other, having been descended from the variety that emerged in Malacca, and their connections are consequently stronger than their simply being

Luso-Asian creoles. Comparison of a Portuguese-lexified creole in Malacca and its daughter languages can help us show how creoles can develop over time. Their origins and evolution in such a linguistically diverse region also has implications for how language contact is involved in the formation of creole morphosyntax.

The description of features here is a morphosyntactic one, but this is not meant to be a comprehensive grammar. Much work has already been made on the grammars of the individual languages (such as Baxter 1988, Maurer 2011), and my work builds on previous grammatical descriptions. This is ultimately a comparative project, that aims to synthesize the various morphosyntactic features found in Kristang, Makista, TC, and BC, and look at how their morphosyntactic structures have developed within different linguistic ecologies.

Michaelis (2008: xiii) notes that "Even today… we are still woefully lacking systematic cross-creole data, and we are still in need of more fine-grained data-based studies of the various structural aspects of different creole languages." It is with this in mind that I have decided to investigate the structures of these four varieties and how they relate to each other. This can provide a look into specific creole structures, and how they fit within the characteristics of creoles as a whole.

Chapter 5: Data, Methodology, and Methods

5.1 Introduction

In this chapter I discuss the sources for data from each of the four languages, the methodology I used in collecting data, and the methods I used for conducting my research. The type of data used varies between written and spoken, and for Kristang data I collected myself, and data that others collected.

Although I was able to conduct my own fieldwork on Kristang, my time and funding were limited so I have relied heavily on the recordings of other researchers and the writings of speakers of Kristang and Makista. As BC and TC are no longer spoken, I have relied upon Maurer (2011) for data. His volume is a compilation of data from BC collected by Schuchardt (1891a, 1891b), whose original source was a 1780 book, published in Jakarta, titled *Nieuwe Woordenschat, uyt het Nederlanduitsch in hed Gemeene Maleidsch en Portugeesch, zeer gemakkelyk voor die eerst of Batavia komen[6]*, together with TC materials derived from an anonymous manuscript (printed in França 1978), and Wallace (1978). Part of my fieldwork in Malaysia involved elicitation exercises with multiple speakers of Kristang. Below I include a sample of the pictures that I used in these exercises.

5.2 Data
The following section outlines the data I have used for each of the creole varieties.

5.2.1 Makista Data Sources

Makista has very few native speakers left in the 21st Century. According to the UNESCO *Atlas of the World's Languages in Danger*, there were 50 speakers in 2000, but the number is certainly smaller now. As ethnically Makista people are a small minority in Macau, and Makista-speakers a small minority among them, it is difficult to find speakers to record for research without an intimate relationship with the community. Also, as there is no Makista-speaking neighborhood of Macau, it is also hard to find speakers without knowing exactly where to go. This thesis uses two main sets of corpora, one from the late 19th and early 20th centuries, and one from the mid-to-late 20th century.

The more recent Makista corpus used as a source in this thesis is primarily written work by José dos Santos Ferreira, also known as Adé. A writer and poet who lived from 1919 to 1993, his works include plays, poems, and stories about life in Macau. A resident of

---

[6] "New vocabulary, from Dutch to Common Malay and Portuguese, very convenient for those coming to Batavia for the first time" (Maurer 2011, 1)

Macau his whole life, Santos Ferreira was a prolific writer whose volumes include poetry, prose, and theatre, in both Portuguese and Makista. While useful for the purposes of studying the Makista language itself, his work also provides a glimpse into the life of Macau as it was during his active period of writing. His volumes include *Macau sã Assi* (1968), *Qui Nova Chencho* (1974), *Poéma di Macau* (1983) and *Macau di tempo antigo* (1985), all of which have had an impact in the cultural life of Makista, including being used in the Patuá theatre by *Dóci Papiaçám di Macau*. Many of Adé's works are original, but he also made adaptations of existing songs and stories in his work, such as *The Godfather* (Santos Ferreira 1973) and *Romeo and Juliet* (Santos Ferreira 1967).

Adé's work is written purely in Makista, with no translated version in either Portuguese or English. This presented a challenge in ensuring an accurate and faithful morphemic glossing and translation of each example presented in this thesis. However, his volumes did include glossaries with definitions of some terms in Portuguese. Additionally, functional morphemes in Makista have very similar forms to those in Kristang, such as the TAM markers (*já, lôgo, tá*), the possessive marker *sua*, and prepositions such as the locative *na* and possessive *di*. It was through these similarities with Kristang and the glossaries with translations into Portuguese that I was able to present the data with interlinear glossing.

An obvious limitation is that by relying on Santos Ferreira's work, we have the language of a single person. Individual variation, personal choices, or peculiarities in language will inevitably come through in material of this type, compared to a corpus collected from and cross-checked with a larger number of people. As a writer of a minority language, Santos Ferreira wanted to promote it and declare it as a distinct language from Portuguese, referring it to the "Sweet Speech of Macau" (from which *Dóci Papiaçám* take their name). One such example is his usage of reduplication, where Arana-Ward (1977) contends that speakers of Makista reported that reduplication is an archaic feature of the language, known to the speakers but not often used by them. Santos Ferreira makes extensive use of reduplication, especially reduplication of nouns to make them plural. One reason for this could be that standard Portuguese does not use reduplication productively as Makista does, so the author wanted to exploit, and perhaps overuse, this feature to emphasize Makista's distinctiveness. With this in mind, it may be the case that Santos Ferreira exaggerated a number of the more creole-like or Makista-like (ie, non-Portuguese) features such as reduplication, in his work, either consciously or not. However, it is not the case that he invented this feature, but rather used it in a more extensive way than probably was used by speakers of Makista during the mid-20th Century. This type of usage has been noted in

endangered languages, where as they decline, the users adhere to an idea of linguistic purism (Lüpke 2005, Florey 2004).

The older corpora consist of a collection of letters, journal entries, poetry, and theatre from various Macanese authors. These were published in the periodical *Ta-ssi-yang-kuo* between the years 1889 and 1901 (Pereira 1901) as well as the academic journal *Renascimento* (Barreiros 1944). These sources, like the Santos Ferreira corpora, are also completely written, and potentially edited from their original forms (Lebel 2021). However, this has a benefit to the present study as the original authors were multiple people rather than a single one as was the case for the newer corpora. Chapter 6 will also demonstrate that many of the patterns identified in Santos Ferreira's work are also present in the writings from 50-70 years earlier.

5.2.2 Kristang Data Sources

There is much more data and analysis available for Kristang than Makista, owing to its larger population of speakers, and a larger number of researchers having worked on it. Baxter (1988) provides a good insight into the structure of the language, and concepts that can be tested in fieldwork.

Stefanie Pillai, University of Malaya, recorded Kristang speakers in the Portuguese Settlement of Malacca, and I have made use of the corpus found in her deposit in Endangered Languages Archive[7] as well as other recordings that she personally shared with me. Her corpus includes audio and video recordings and covers a range of genres, including monologues, interviews, and conversations between multiple people. Her deposit in ELAR consists of 20 recordings, 15 of which contain video. Of these, nine are classified as narratives, six as interviews, two as conversations, one as a cooking demonstration, and one as a song. Much of the recordings are involving activities and aspects that relate to the culture of the Malacca Eurasians, such as fishing, talking about festivals, local cuisine, and games played by the residents. In all, 22 speakers are identified as being participants in the recordings. The type of speech differs between the recordings, as well. For instance there is a cooking demonstration in which a woman is demonstrating how to make *achar pesi*, a fermented fish pickle used as a condiment. In this recording, she is the only speaker and is talking her way through the process of making the dish. Another narrative involved two Kristang-speaking participants, who speak about the San Pedro festival celebrated in the Portuguese Settlement. This is similar to the *achar pesi* recording in that the only speakers

---

[7] https://elar.soas.ac.uk/Collection/MPI130545

are the participants (and not the researcher), but involves two men instead of a single person. The recording with the most participants is one in which four women demonstrate the card game *chiki*. They explain how it is played and talk amongst themselves while playing the game.

There are also recordings that are in the form of interviews. The interviewer in each recording is the researcher, and asks questions in Kristang. Although the majority of the language used in the interviews is Kristang, there are some instances of English being used, both in the form of code-switching and sometimes entire sentences.

In addition to the recordings available on ELAR, Pillai supplied me with more recordings made in the Portuguese Settlement. This consists of 12 recordings made in a similar manner to the ones on ELAR. One of the recordings is an interview, where the interviewer is a child who asks questions in Kristang. In addition to the interviewer and the interviewee, the mother of the child sometimes add to the discussion, also in Kristang. In another interview, the questions are asked in English, but the answers are given in Kristang. This group of recordings also has narratives, and groups of people playing *chiki*. Like the recordings on ELAR, the content of the recordings is based on the participants' lives, including their personal histories, parts of Kristang culture, and daily activities.

None of the recordings were based on elicitation with an obvious linguistic goal in mind, and for the purposes of this thesis they were used for some of the quantitative methods (see Chapter 6). These were also used as a basis for the subsequent recordings I made in Malacca, as the counting of certain tokens and examples of features helped guide the elicitation tasks that I conducted. The material had English translations of the lines spoken in Kristang, but there was no line-by-line morpheme glossing. The examples from Pillai's corpus were glossed by myself, based on the grammar of Kristang from Baxter (1988).

One of the ways I used this corpus was to investigate short and long pronoun forms. In Pillai's transcriptions they are all written in a single form, so knowing which one was used required searching through the transcriptions for instances of the pronouns I was analyzing, and listening to which form was spoken in each instance.

Pillai collected metadata, such as the age and gender of speakers, and was able to record speakers in their own homes, and make video recordings. She asked the speakers to talk about their jobs, what it was like growing up in Malacca, and other questions about daily life.

The following table has a summary of the type of recordings in Pillai's corpus:

Table 7: Summary of recordings in Pillai (2011)

Title	Genre	Topic	Language(s) used
MPC01	Monologue, Song	Kristang version of "Silent Night" with spoken introduction	Kristang
MPC02	Monologue	Speech about the history of the Kristang language	English, Kristang, Malay*
MPC03	Monologue	Fisherman talking about fishing	Kristang, Malay*
MPC04	Monologue	Speech about the history of Malacca and its musical traditions	English*, Kristang
MPC05	Monologue	Man speaking about Kristang dance	English*, Kristang, Malay*
MPC06	Monologue	Man speaking about Kristang dance	English*, Kristang, Malay*
MPC07	Monologue, Song	Kristang song with spoken introduction	Kristang
MPC08	Monologue	Woman demonstrating how to make fish pickle	English*, Kristang, Malay*
MPC09	Monologue	Man speaking about his life	English*, Kristang, Malay*
MPC10	Monologue	Man speaking about the *Intrudu* festival	English*, Kristang, Malay*
MPC11	Interview	Personal profile of interviewee	English, Kristang, Malay*
MPC12	Interview	Personal profile of interviewee	English, Kristang, Malay*
MPC13	Interview	Personal profile of interviewee	English, Kristang, Malay*
MPC14	Interview	Personal profile of interviewee	English, Kristang, Malay*

MPC15	Interview	Personal profile of interviewee	English, Kristang
MPC16	Interview	Personal profile of interviewee	English, Kristang
MPC17	Conversation	Family speaking about the importance of speaking Kristang	Chinese*, English*, Kristang, Malay*
MPC18	Monologue	Man speaking about his work as a fisherman	Kristang, Malay*
MPC19	Conversation	Four women playing cards	English, Hokkien*, Kristang, Malay*
MPC20	Dialogue	Two men talking about the San Pedro Festival	English, Kristang, Malay*
MPC21	Monologue	Woman describing events during Lent	English, Kristang, Malay*
MPC22	Conversation	Four women playing cards	English, Kristang, Malay*
MPC22 (second batch)	Monologue	Speech about the history of the Kristang language	English*, Kristang, Malay*

The usage of a language marked with an asterisk (*) is minimal, and only consists of a single word or particle. English is the only language aside from Kristang in which full sentences are used, however all of the recordings are overwhelmingly Kristang.

Pillai (2011) transcribed the recordings and translated each line into English, as in the following example:

Beng nakih papiah kada dia
I come here everyday to talk

Lent nus start di Intrudu
Lent we start with Intrudu

Intrudu nus bringkah ku agu
We play with water on Intrudu

Kabah akeli bringkah ku agu
After playing with water

Dos oras di oitura pamiang nus bringkah ku agu sampai dos oras midia

From eight o'clock in the morning until mid day twelve o'clock we will play with
water

As can be seen above, the transcription has English translations of each line, but there is no morphemic glossing, nor are there word-by-word translations. Each recording is also paired with an ELAN file with one tier each for transcription into Kristang and translation into English.

Picture 1: Screenshot of ELAN file (Pillai 2011)
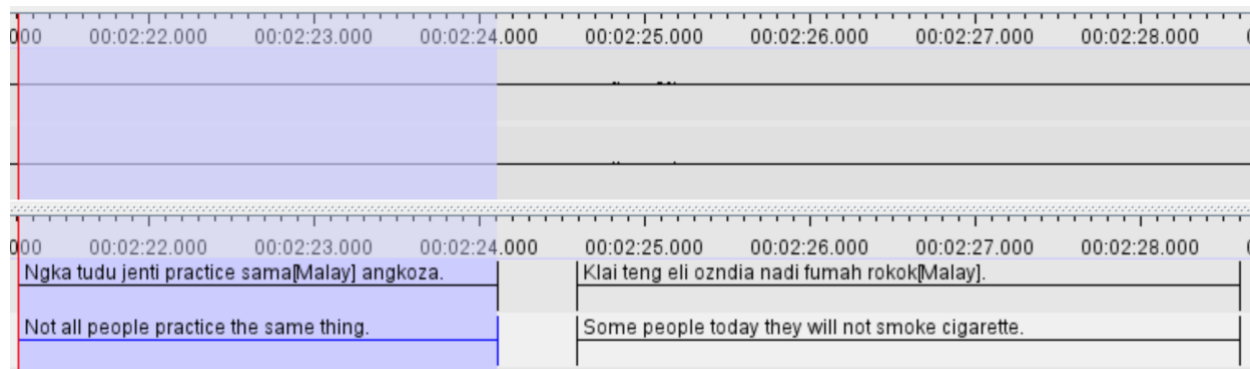
Another small corpus was collected by myself during two field trips to Malacca in 2017. These were part of initial contact with the Kristang-speaking community, and therefore much of the time spent there was working on establishing rapport with the community. The first trip resulted in a recording, although it was not elicited or structured enough to have been included in this study. The second trip produced a corpus that resulted from work through elicitation with members of the Kristang-speaking community whom I had recorded before. I did not make video recordings. I wanted my speakers to be at ease as much as they could, so I recorded them within the Portuguese Settlement, in places where they felt comfortable. The recordings I made were at the homes of Kristang speakers, where people often gather together to sit and chat about their lives. I also gave informed consent with a form modeled after Pillai's that explained the purpose of my study as well as the rights of the participants to recall their consent at any time. More of my fieldwork will be discussed in Section 5.4.

The corpus itself consists of recordings from May: 1:33:14 and an extra 10 minutes. This recording was almost completely unstructured and was a way for me to introduce myself as a researcher and to show my methods of recording. This was unstructured speech in a mix of Kristang and English, with 5 participants, male and female, aged between 55 and 65 years.

The October recordings consisted of a 40-minute recording between myself and a participant whom I recorded in May. He is a male aged 65 at the time of recording, who did not live within the actual boundaries of the Portuguese Settlement but close by, and would often socialize within the Settlement. The recording session took place at the home of a friend within the Settlement itself. The next session 35 minutes of guided activities. This was at the home of the participant from the first recording, and included himself and two male friends of a similar age, plus some recordings with a 36-year-old male who lived in the Portuguese Settlement. The majority of the examples from my data in this thesis are from the October recording sessions.

All participants, no matter the size of their role in the recording, received informed consent in the form of a written statement. I printed enough consent forms for each participant to read, sign, and date. They were written in English as the residents of the Portuguese Settlement are English speakers, and it was modeled after the consent form Pillai used in her recordings in the settlement. Consent was given by all participants, and was not relinquished after the recording sessions by any participant. See Appendix 1 for a copy of the consent forms I used.

5.2.3 Batavia Creole and Tugu Creole Data

All of the BC and TC data in this thesis comes from Maurer (2011), which he collected into a single volume. The BC data is from Schuchardt (1891), who in turn collected the data from a 1780 guidebook for visitors to Jakarta. The TC data is from Schuchardt (1891), as well as a manuscript written by a Dutch priest in 1937, printed in França (1978) amd Wallace (1978) (Maurer 2011: 1-2). This data is glossed by morpheme, and there is also a line of English translation. I used the English translations without changing them in this thesis.

5.3 Methodology

Baxter & Bastos (2012) on the genitive in Luso-Asian creoles, with a particular focus on Kristang, provided a methodological background for an aspect in the current study. They described the ways possession was expressed in Kristang, and investigated the origins of the structures and the distribution of their use. In Laub (2014), I used the same categories as Baxter & Bastos (2012) and counted tokens of Makista genitives. What I found was that the pre-possessor genitive was overwhelmingly preferred, but there were remnants of the post-possessor genitive as they appear in Kristang. Interestingly, the prominence of the post-nominal genitive was in the same places where it was prominent in Kristang. In the present

study I adopt similar quantitative methods to Baxter & Bastos (2012) and Laub (2014). One of the features I examine is pronoun variation in Kristang (see Chapter 6, Section 4). Three of Kristang's personal pronouns (2SG, 3SG, 1PL) have variants in their pronunciations in which the final phoneme is sometimes dropped. I aim to address what could be the causes of this by counting tokens of these various pronoun forms in the corpus in terms of their environments. Unfortunately, the corpus is too small to provide a sufficient number of tokens to be able to draw statistically significant conclusions, but it does provide some hints that give us a clue about what could be the cause of the variations.

While it is important to look at when a feature is present and when it is absent, we must also look at why it occurs. As this is a study looking into creole languages' morphosyntactic structures and how languages within their ecologies have influenced them, we need to focus qualitatively on how the features behave within the individual languages. As Kristang is spoken in Malacca, I make comparisons certain features that can be found in Malay, the majority language of the city. It is because of this that I have chosen to use a mixed quantitative-qualitative methodology that aims to describe how a certain feature is distributed, but also why.

The qualitative aspect of this methodology is useful for the descriptive and comparative parts of the study. By exploring current theories on features we can then see how they work in the creoles of this study, and then compare them to each other and the languages with which they came into contact. The qualitative method can best examine these features by zooming in and discussing how they work in detail.

I have chosen to incorporate some quantitative methods as well, because this can help reveal some of the ways the syntax of the languages works. Thus, Baxter & Bastos (2012) undertook a quantitative study on genitives in Kristang uncovered a pattern that a solely qualitative method would otherwise not. Applying the same method to the other languages can help us compare functional differences in a way that might not be seen with only a qualitative method.

One caveat is that the corpora used in this study are too small to make the sort of strong conclusion that one would be able to do with larger data sets. For this reason I use the quantitative methods to compare the creoles with each other, but not with their super- and substrate languages. Languages like Malay and Cantonese have millions of speakers and are very well-documented both in writing and in speech, and the conclusions we could make from corpora of them are different from those of the smaller languages under consideration.

Section 5.4 discusses the Kristang data, which somes from both outside sources and my own fieldwork, although I will include some words about how I applied different methodologies to the data from different sources. The quantitative elements of the methodologies will be applied to data collected by other researchers, including those collected by Pillai (2011), because much of this data consists of monologues, interviews, and discussions among speakers of the creole languages. These were collected without the purpose of investigating certain features that I will discuss in Chapter 6, and are therefore more natural speech. On the other hand, the data I collected was largely the result of elicitation work aimed at finding certain features. The participants' language was therefore guided by my exercises and would produce words or structures in a distribution that would not likely be found in unguided speech. It is this data that is used in the qualitative areas of my research, as we can see which environments would trigger which structures (such as a long or short pronoun in Kristang, as seen in section 6.4).

As explained above, mixed methods will be used in this thesis. To demonstrate the ways the lexifiers and substrates have influenced the resultant structures of the creoles, we will investigate different features of their structure. While by no means exhaustive, the features chosen will give the reader an idea of how their structures in Kristang, Makista, BC, and TC compare with those of the lexifiers, the substrates, as well as with other creole languages and with each other. Chapter 6 is divided into sections about different chosen features, and will feature a description and definition of the features to establish what each feature exactly is (for instance, what distinguishes a true serial verb construction from a two-verb sentence). Where appropriate, I will apply some quantitative methods. This includes the number of tokens of a genitive structure, as previously studied by Baxter and Bastos (2012).

5.4 Methods

In 2017 I made two trips to Malaysia to become acquainted with the Portuguese Settlement and to conduct fieldwork. The first was in March and April, where I first got to know the people in the Settlement and their language. Although people were happy to talk to me about their language, asking them to be recorded was a little more difficult. As Kristang is a creole language, the image of it to the speakers may be that it is broken Portuguese, or Portuguese mixed with Malay. I brought with me a Roland Edirol R-09 24Bit WAVE/MP3 digital stereo recorder, and a Sony ECM-MS957 electret condenser stereo microphone. I also brought a table stand as my fieldwork was conducted while sitting at a table. Other equipment that was required was available in Malacca or in Kuala Lumpur, where I started my time in Malaysia.

With Kristang having its origin in Malacca, today it is easiest to find speakers in the Portuguese Settlement located there. The language is still spoken in the city that it was born in, and the influence of other languages on Kristang can be seen in daily life in the city. The Portuguese Settlement is the best neighborhood in which one can encounter Kristang speakers, and one can also see the influence of Portuguese religion and culture in the decorations in their houses and names of streets and businesses.

An advantage of conducting fieldwork on Kristang is their concentration of speakers within the Portuguese Settlement, even though the language is endangered. If someone is not a regular speaker of Kristang, they will know someone who is. This was useful not just for my own recordings but for acquiring knowledge of speaking the language, which I intended to use for the data collection. Speakers of Kristang are in general very happy to share their language and their knowledge with those who are interested. Among the older generations, speakers are eager to talk amongst themselves and with interested outsiders about the usage of their language, what words they may or may not use, and ask for suggestions for new words that may not be in the everyday language (this is especially common for technological advancements of the 20th Century, such as the refrigerator or airplane). While I did make valuable connections with speakers of Kristang, I was not there long enough to establish rapport to schedule a recording session that would provide any meaningful data. I was able to record an informal conversation among five friends who gathered to talk about Kristang. This was a mixture of Kristang and English, and also included singing. This recording may prove useful as a corpus for Kristang, especially with the songs, although for the purposes of the present study it was unfortunately inadequate. Some issues included multiple speakers, but only one microphone, leaving some voices harder to hear than others. It did leave me little to work with, but I was able to make connections with people who I know would be open to further recording sessions.

My first trip to Malacca thus did not produce sufficient data for the present study. However, it was a worthwhile trip because of the personal connections I made while in the community. Bowern (2008: 7) points out that "a better personal relationship between the linguist and the consultants will result in data collection." As I was an outsider in the beginning, it was important that I make my intentions in the community clear well before I attempted to make any data collection. Kristang is a very small minority language, putting it in a position that is subordinate to many other languages in the multilingual city and nation. Many of the speakers I met hold their culture and language dear to them, so it was vital that I

treated it with respect. Because of this, I could not realistically start recording and collecting data immediately after my arrival in the field.

Another advantage of my first trip to Malacca was the opportunity to learn Kristang. Before I arrived, I became familiar with the language through studying its grammar in Baxter (1988). I also received a copy of *Beng Prende Portugues Malaká* (Singho et al 2016) which I used to gather a basic knowledge of spoken Kristang. As Kristang is a Portuguese-lexified creole, I also had the advantage of being familiar with many words already, due to my own knowledge of basic Portuguese. This background knowledge of Portuguese and Kristang helped me use the language with the residents of the Portuguese Settlement, and in the process I learned more about Kristang and became more confident in speaking it. During fieldwork, making the effort to learn the language one is researching can be very helpful for both personal and research reasons (Bowern 2008). This helped me in my activities both within the actual collection of data and elsewhere. My willingness to use the language and learn from the speakers gave a positive impression on the community members, which helped me with the personal relationships in the community. I was able to learn about and participate in the community, which helped both myself and Kristang speakers become comfortable by the time it came to record the language and collect data. This also helped me in my linguistic research, as some of the activities I conducted made use of prompts in Kristang. I was not, and still am not at the time of this writing, a fluent speaker of Kristang, but this basic knowledge greatly sped up the process of creating Kristang prompts.

My second trip was in October of 2017, funded by a bursary from the Center for Southeast Asian Studies at SOAS. This trip was only for one month, but with the personal connections I had made before I was able to conduct more successful recording sessions. My first session was with one of the participants in the recording I made earlier in the year. One advantage of conducting a one-on-one recording session was that I never had an issue in hearing his voice, as I did not need to move the microphone. Although the microphone was pointed at him, I was near enough that my voice was also audible. The recording session was made outside, as the weather in Malacca is always warm enough to be outside, in front of a house in the Portuguese Settlement. The streets in the Portuguese Settlement are not busy, but there is the occasional car or motorcycle that drives by. Fortunately, the microphone picked up his voice well enough that even traffic noises did not affect the recording too much.

Because the previous corpus by Pillai was more naturalistic speech, the methods I used in my fieldwork was elicitation and guided exercises, in order to investigate specific

phenomena. The following will lay out the methods I used in the collection of my corpus, with examples of each.

5.4.1 Method 1: Translation (English into Kristang)

One advantage of working with speakers of Kristang is their knowledge of spoken and written English. My consultant for the translation activity was a good candidate, as someone who is educated and literate in English, but also a frequent user of Kristang. Translation is not an activity that is unfamiliar to him; as a hobby he translates Christmas carols from English to Kristang. Code-switching and frequent use of both languages are an everyday occurrence in his life, so the act of translating a simple sentence from Kristang to English is not something that he would have found difficult. A disadvantage is that the sentences I chose were based on ones that I saw in Makista, and wanted to see what structures a sentence with a similar meaning would result in Kristang. The sentences I used for translation were presented by themselves in a list without context, so the use of one particular grammatical structure would not necessarily rule out other structures.

I wanted to compare the structure of *cavá* in Makista with that of *kabah* in Kristang, so I wrote sentences down in English based on the Makista structure, and then asked my consultant to read them and translate them into Kristang. I also used the opportunity to ask for other translations where one would use *kabah*.

a. After Christmas, we celebrate the New Year.
b. After the war, he started a family.
c. After the war ended, he started a family.
d. After the war began, he enlisted in the Army.
e. After the main dish, we will drink coffee.
f. After (we eat) the chicken, we will eat rice.
g. The boys will dance after the girls.

This translation exercise gave me examples of *kabah* in Kristang, which I can use to compare against the Makista examples using *cavá*, as well as add to the corpus of Kristang. My consultant seemed to prefer to use *kabah* before a verb phrase, as can be seen in his translations of sentence (e).

1. Kristang (Laub fieldwork, 2017)

   kabah   nus   kumih

   after   1PL   eat

   "After we eat"

2. Kristang (Laub fieldwork, 2017)

kabah	ja	kumih,	nus	lo	bebeh	coffee
after	PFV	eat	1PL	FUT	drink	coffee

"After we eat, we will drink coffee"

3. Kristang (Laub fieldwork, 2017)

kabah	kumih	bebeh	coffee
after	eat	drink	coffee

"After we eat, we will drink coffee"

These examples show that there are a variety of ways to translate my sentence e, but my consultant preferred to say it using the verb *kumih* rather than say "main dish".

I continued with the translation activity with more sentences inspired by constructions I saw in Santos Ferreira's Makista.

h. I escaped death!

i. I want to sleep.

j. I need some sleep.

k. I avoided sleep.

l. If you do that again, I will beat you!

m. If you do that again, I will give you a beating!

5.4.2 Method 2: Video stimulus

Another method I used in the one-on-one sessions was a video stimulus. One issue that had not been looked at before in detail in previous studies of Kristang is the long and short forms of pronouns in Kristang. There are, however, enough examples in the previously existing data to formulate a hypothesis to help me get further data in my own fieldwork. It was through Pillai's recordings that I concluded that short-form pronouns are not solely phonologically dependent (as claimed by Baxter), and that it is not an individual variation.

From this data I gathered that information structure might play a role in the choice of pronoun, so I used resources that would help elicit certain structures that might trigger one form or another. One of my sessions was one-on-one with a Kristang speaker. In one of Pillai's recordings I noticed that the speaker uses the generic "you" while describing how to make food. In this case, she exclusively uses the Kristang short form *bo* (rather than *bos*). I wanted to elicit a similar monologue from my consultant, and I first showed him a video from Dóci Papiaçám di Macau, in which the narrator describes things one should and should not

do as a visitor to Macau[8]. This video uses a general "you" throughout in the form of *vós*, an identical form to what is found in other environments where it refers to a specific person as the second person pronoun.

Picture 2: Screenshot of DÓCI VAI RUA – FICÁ VIVO NA MACAU

One issue with this is that Makista is a very similar language to Kristang, so the consultant listening to a narrator speaking the former language could have influenced the Kristang that resulted in the monologue. Additionally, there were subtitles in Chinese, Portuguese, and English.

I then asked my consultant to speak freely on the same topic, telling visitors to Malacca and the Portuguese Settlement things they should and should not do when they come. What resulted was a monologue containing usage of the general "you", and exclusively the short form. This occurred in the short form despite the Makista video including the pronoun *vós*, which has a pronunciation more similar to that of the long form in Kristang. Future research would benefit from elicitation that would result in a similar monologue using a general "you", but with the prompt not being an example in another language, especially one so similar to Kristang. However, the focus of this elicitation exercise was the **content** of the video, not the **structure**, and the consultant was presumably thinking about advice for tourists in Malacca while speaking during the recording.

---

[8] DÓCI VAI RUA - FICÁ VIVO NA MACAU https://www.youtube.com/watch?v=93YUQCwejNY&t=2s

5.4.3 Method 3: Grammaticality Judgments

In this recording session I had prepared some questions asking about grammaticality. I was curious about the usage of pre-verbal TAM markers, and which are grammatical and which are not. This is discussed in more detail in Section 6.5. My goal was to see which markers can appear without a main verb. There is a common exchange of greetings in Kristang:

4. Kristang (personal knowledge)

    ja   kumih?

    PFV   eat

    "Have you eaten?"

    ja

    PFV

    "Yes (I have)"

This perfective marker *ja*, which can appear in the response by itself, is derived from a Portuguese adverb, as is the future-irrealis marker *lo(gu)*. In contrast, the progressive marker *ta* is derived from a Portuguese verb. I was curious to see if the category status of the Portuguese original had had any effect on whether the marker can occur on its own grammatically, so I asked my consultant questions about that, such as:

5. Kristang (Laub fieldwork, 2017)

    "But if I ask a question like *Ja kumih?* [you'd] say *ja*?"

    "Ja, straightforward, straightforward, ja."

    "[If] you ask me *Bos ja bebeh?*, I'd say *ja*?"

    "Yes, *bos ja bebeh*, also can."

    "What about … *bos ta bebeh*?"

    "*Bos ta bebeh,* yes."

    "But then [can I] answer *ta*?"

    "No."

It was through questions and answers in this section that I determined that answering a question with a bare *ja* is grammatical, while *ta* is not.

Another feature in Kristang is the usage of multiple elements before the main verb. This includes the TAM markers mentioned above but also includes modals and auxiliary verbs. One such example is *tokah*. On its own it means "to touch" but it also occurs in an auxiliary function. As described by Baxter (1988) it can be used to express an "adversative

passive", meaning, ie a passive that carries a negative or unfortunate connotation. Marbecks (2004) has an example of *tokah* being used as an adversative passive, in a story in which a tiger gets caught in a trap:

6. Kristang (Marbeck 2004: 53)

istri	tigri	ja	tokah	pegah	na	lastru
this	tiger	PFV	touch	catch	in	trap

"This tiger here was caught in a trap"

In this sentence, *tokah* does not have the meaning of "touch", but rather makes the verb *pegah* "to catch" into a passive. As the tiger presumably does not want to be caught, this is an appropriate use of the adversative passive, as per Baxter.

Interestingly, in the same story, a character is talking about the tiger having been caught in the trap, but the main verb *pegah* does not appear, unlike the previous example:

7. Kristang (Marbeck 2004: 53)

bos	olah,	istri	tigri	ja	tokah	na	trongku
2SG	see	this	tiger	PFV	touch	in	trap

"You see, this tiger was caught in a trap"

In this case, we have the verb *tokah* followed directly by the goal *na trongku* "in a trap", with the verb *pegah* unstated. I wrote this example sentence down for my consultant, gave him a copy and read it to him. He agreed that it was grammatical, so I asked him about other verbs that might be grammatically removed after *tokah*.

Based on the sentence from Marbeck and my own knowledge of Kristang, I made the following sentence:

8. Kristang (Laub fieldwork, 2017)

istri	tigri	ja	tokah	pegah	na	trongku,	kun	istri	tigri	ja
this	tiger	PFV	touch	catch	in	trap	but	this	tiger	
										PFV

tokah	(pegah)	na	riu
touch	catch	in	river

"This tiger got caught in a trap, but this tiger got caught in the river."

I first asked my consultant if the version with the second *pegah* was grammatical, and he agreed that it was. I then asked if it was okay if I took out the second *pegah*, by speaking the

sentence to him in Kristang. He repeated it back to me including *pegah*, but after some more questioning he said that it must be included to be grammatical.

In Kristang, *tokah* has another meaning expressing obligation. Its construction is the same as the adversative passive, which led me to look into if a verb after *tokah* in the obligation sense can be dropped (as is the case in sentence 3 for the adversative passive).

The sentences I constructed were a mix of those from previous Kristang data and those that I made up myself. The original sentences did have the main verb after *tokah*, so I asked if it was okay to use it with the verb removed. Also, the obligation sense widens the types of verbs that can be used, including intransitive, which the adversative passive may not. Because of this, I wanted to look at different thematic relations for the element that comes after the verb to see which ones were acceptable and which ones were not.

One of the examples was taken from Pillai (2011). It has men describing the Intrudu festival that marks the beginning of Lent. In this festival, people throw water at each other, but not at people who are older than the thrower. The following sentence describes this rule:

9. Kristang (Pillai 2011):

nus	tokah	(pinchah)	agu	ah	nus	sa	keng	teng	igual	sa
1PL	touch	throw	water	to	1PL	GEN	who	have	same	
	GEN									

idadi

age

"We must throw water at those who are of the same age"

The original sentence had the verb *pinchah* "to throw", so I asked if it was possible to say it without the *pinchah*. My consultant said that this was not possible to say. The thematic relations this verb has are the theme *agu* "water" and the goal *nus sa keng tend igual sa idadi* "those who are of the same age."

I then asked about a sentence that I made from my own knowledge of basic Kristang:

10. Kristang (Laub fieldwork, 2017):

nus	tokah	(sperah)	na	kaza
1PL	touch	wait	at	house

"We have to wait at home"

In this example, the verb *sperah* "to wait" is intransitive, and *na kaza* is the location. As with example 9, my consultant said that it was not possible to say it without *sperah*.

The next sentence was also from my own knowledge. In this scenario I asked my consultant to imagine a parent asking their child to eat their vegetables. The examples from Marbeck (2004) in which *pegah*, the main verb, is dropped, comes after a sentence where it is stated in full. Because of this, I wanted my consultant to imagine a parent repeatedly asking their child to eat their vegetables, where a different form might show up given the context. I asked him about the following sentence:

11. Kristang (Laub fieldwork, 2017)

   bos   tokah (kumih)   bos   sa   bredu
   2SG   touch eat   2SG   GEN   vegetable
   "You have to eat your vegetables"

In this sentence, the verb *kumih* "to eat" is transitive, and *bos sa bredu* "your vegetables" are the theme. My consultant said that one cannot say the sentence with *kumih* missing, although he offered *Bos tokah kumih*, or "You must eat" as a grammatical way to remind one's children to eat their vegetables.

Because of time constraints and other activities planned for this session, I did not look further into the case of *tokah*. In all the examples I presented *tokah na lastru* or *tokah na trongku* are the only ones that were acceptable. Additionally, these were the only examples of *tokah* without the following verb in the whole of Marbeck's written work (*lastru* and *trongku* are different words, but she translated them both as "trap"). It became clear that for at least this particular consultant, it is not grammatical to remove the verb after *tokah*, although further research to confirm the restriction is required.

5.4.4 Method 4: Picture Elicitation

The final method used in the collection of the corpus was elicitation using pictures. Each of the picture activities is explained below in this section. This method differs from the previous ones in that the priming was a mix of language and pictures, rather than language only (in the translations and grammaticality judgments) or language through a video (in the video method). This also differed in that this method was applied during a multiple-person recording session. The activities were not solely conducted through the use of pictures, and I did prepare questions in Kristang to ask the participants. I also gave instructions in English when I was eliciting a dialogue between two consultants.

Advantages of using this are that it constructs a scenario in the minds of the participants through which they can express what is going on. Unlike in the translation method, which provided sentences without a wider context, this one provides more of a

context which can then be used in the consideration of the types of syntactic structures that were used.

One major disadvantage of this is that sometimes the pictures were not clear enough for the participants to understand without a brief explanation. Another issue was that I intended to investigate pronoun forms, however the participants would avoid using pronouns and use other peoples' names. However there were enough examples that provided instances of pronoun usage for my corpus.

For further investigation into the usage of short- and long-form pronouns, I adapted the following activities from the University of Potsdam's *Questionnaire for Information Structure* (QUIS).

Activity 1 is about an animal's birthday party. The goal is to get informants to say which gift belongs to which animal. In one part, Informant B, who knows who the animals are, asks Informant B which gifts belong to which animal. In this case, the gifts are all different objects.

Picture 3: Activity 1
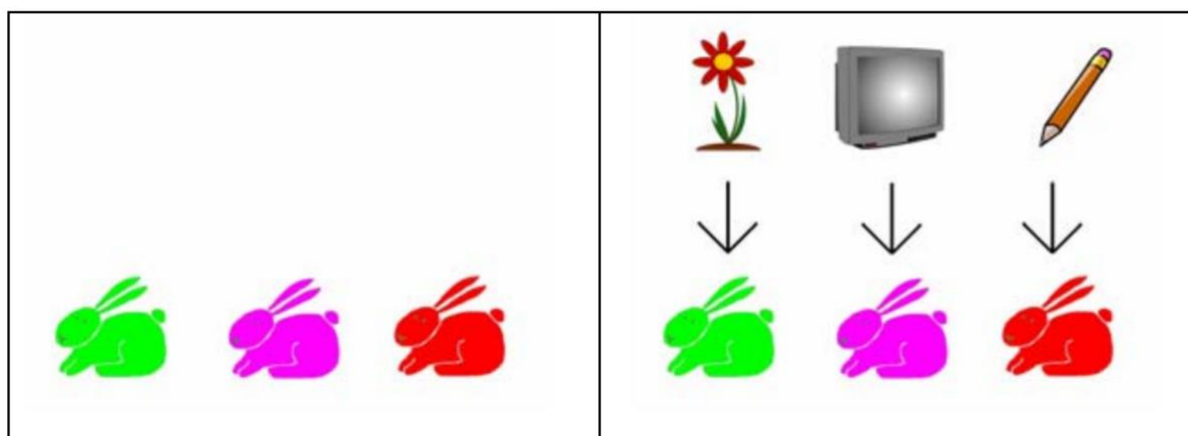
Another version is where Informant B knows the gifts, but not which animal each one belongs to.

Picture 4: Activity 1
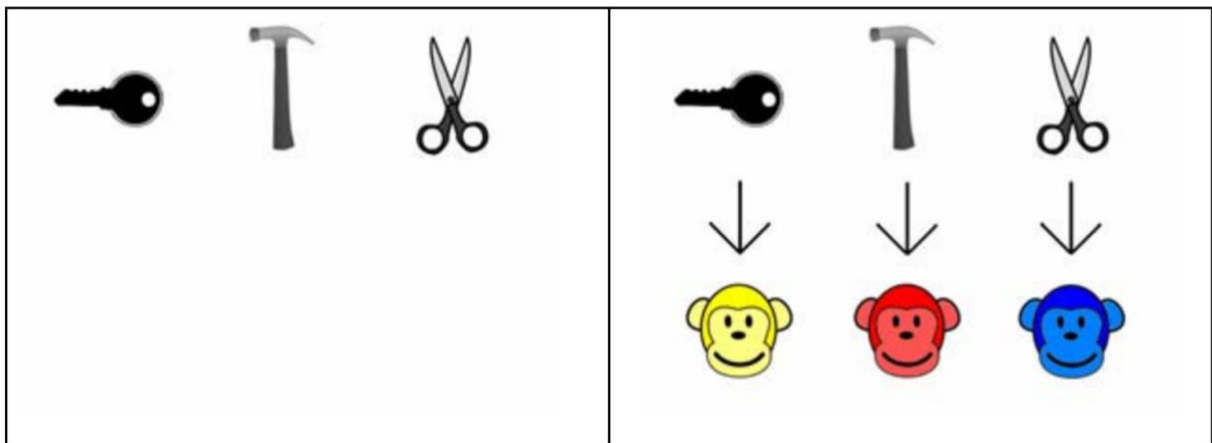
The next version is the same as the first one above, but the gifts are similar with small differences.

Picture 5: Activity 1
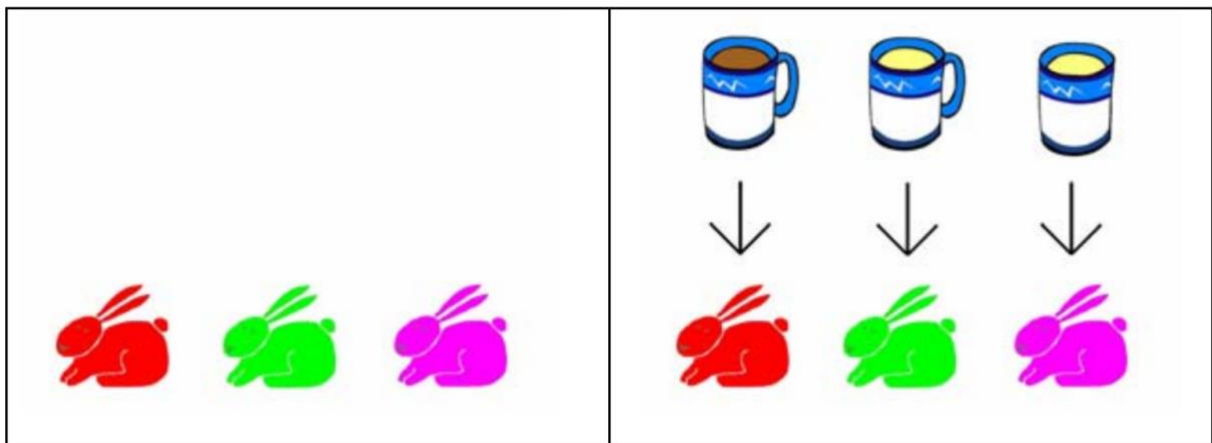
As well as the inverse.

Picture 6: Activity 1
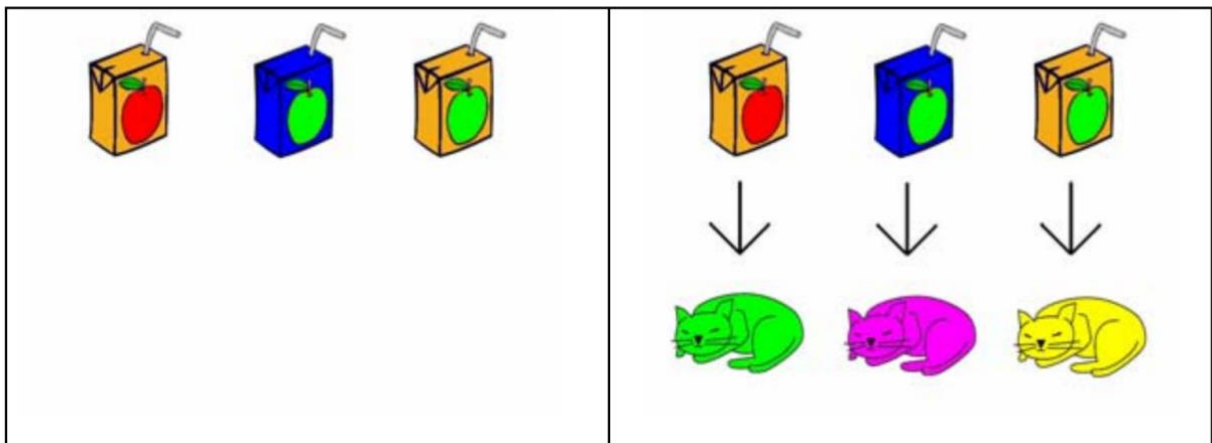
The instructions for Activity 1 were as follows:

> To both informants: "A birthday party! Our three animal friends will each get a present. But who gets what?"
>
> To informant A: "Please answer the question of the other person. You must tell them what to give to whom."
>
> To informant B: It is your task to give the presents to our three animal friends. But you don't know who gets what. You will see a card which shows either three animals or three objects. You must ask the director for instructions. You must ask about all three animals or all three objects at the same time."

The next activity had pictures of people doing things to other people or objects. I asked the consultant to describe what they see based on the questions I ask them.

Picture 7: Activity 2

Sheet 1: Woman hitting man, girl hitting boy, man kicking chair, man pushing car

Sheet 1, questions for session 1

1. There, where the blue sky is: Is <u>a woman</u> hitting the man?
   a. <u>Femi</u> atakah omi?
2. In the room: Is the man kicking <u>a table</u>?
   a. Omi chutah <u>meza</u>?
3. In the garden: Is the girl hitting <u>a girl or a boy</u>?
   a. Fila atakah <u>fila o filu</u>?

4. In front of the well: <u>Who</u> is pushing the car?

       a. <u>Keng</u> push kareta?

The following activity involved handing pictures of various items (see Picture 6) to the consultants and to myself. I then asked them questions regarding who was holding what.

Picture 8: Activity 3
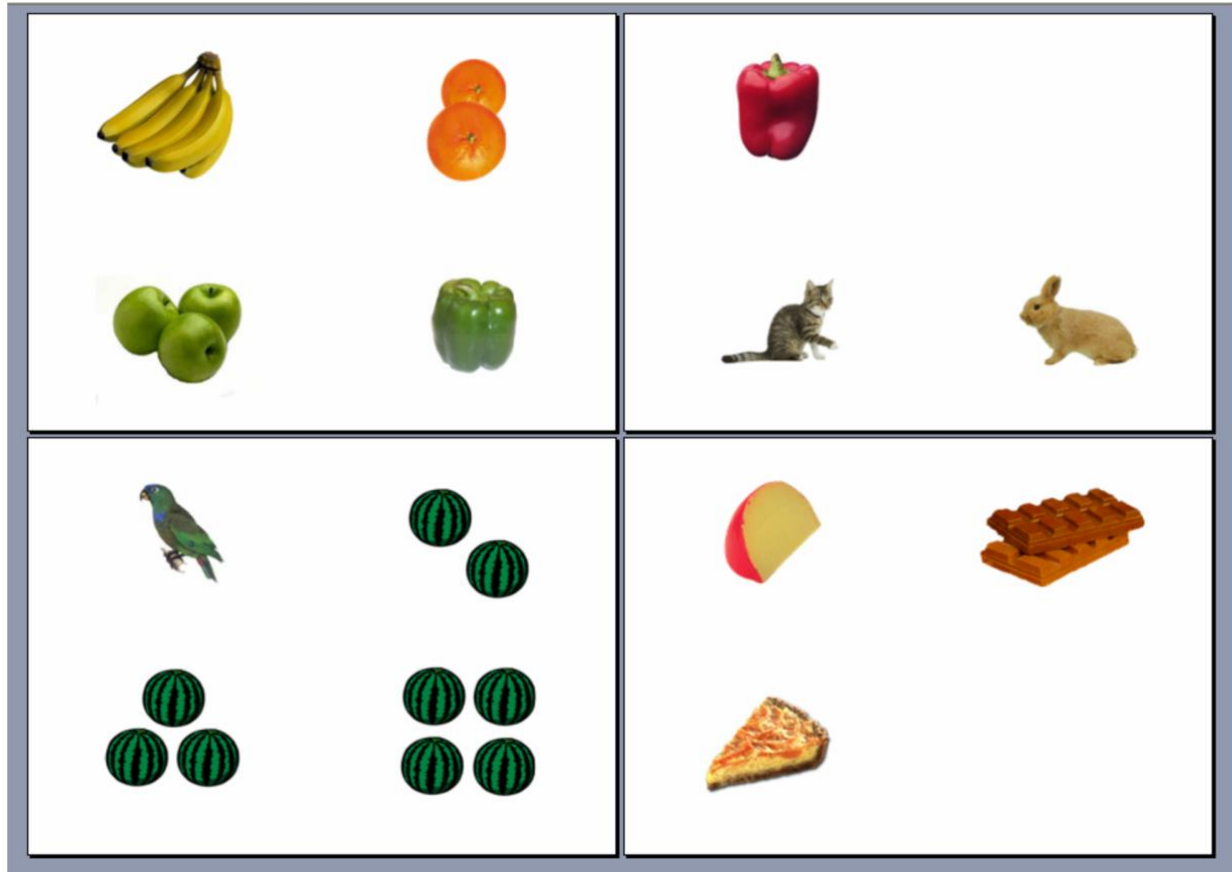
The instructions were as follows. In italics are the instructions for the fieldworker, followed by the questions to ask the consultants. The original English questions are listed, then the translations into Kristang provided by a member of the community.

1. *Give a parrot to X.*

   Who has a parrot, X or Y?

   Ki teng ungua parrot, X o Y?

2. *Give a red pepper to X and a green pepper to Y.*

   Describe what you see!

   Diskobrih ki bos olah!

3. *Give three melons to X.*

   Does X have four melons?

   X teng quartu melon?

4. *Give a rabbit to X.*

   Does X have a rabbit and a parrot?

   X teng ungua kuelyu ku ungua parrot?

5. *Give a rabbit and a cat to X.*

   Does X have a cat?

   X teng ungua gatu?

For activities 2 and 3, I did not want there to be any internal translating between English and Kristang, so with the help of a speaker I translated the questions I wanted to ask, and asked them to the best of my ability. Some issues included problems with my pronunciation and fluency, as well as the consultants' being used to me speaking English to them. Activity 3 was particularly problematic in that I encouraged my consultants to use pronouns as much as they could, but they often would speak in incomplete sentences and use people's names instead of pronouns. This could have been due to the scenarios standing on their own without a larger context; each new scenario was a new situation, and therefore the consultants could have felt the need to express participants' names rather than a pronoun without a referent. This method could have use in looking at the effect of information structure on the construction of individual sentences, such as word order, but it did not result in sufficient data for the purposes of describing pronoun forms, as was originally intended.

5.5 Discussion and Conclusion

As this is a comparative study, I endeavored to look into structures in one language that I could find in another. Kristang data was available to me through fieldwork, so I aimed at looking for structures that could also be found in Santos Ferreira's work. This included the usage of serial verbs and patterns using pre-verbal markers. One feature that was investigated, the long and short pronoun forms, is not as easily found in the Makista data. Since it is only a spoken form and not written in Kristang, it could also be the case that Makista had used different forms that were not expressed in writing. However, based on the performances in Makista, it appears that the short/long pronoun distinction is not as present as it is in Kristang. As for comparisons to the lexifiers and substrates, this was much less of an issue due to Portuguese, Malay, Cantonese, and English having a much larger corpus as well as easy access to native speakers of these languages.

There were various reasons why I chose to conduct the research myself, rather than rely on a native speaker for the elicitation tasks. One major reason was time. I was working on a limited schedule and arranging for someone to meet to introduce myself, introduce them to the material, and explain what exactly I was doing would not have been feasible. With myself as the person conducting the recording, scheduling with participants, consent, and elicitation tasks, I was in control of the sessions. As I was working on limited time, it was essential that my participants take part in the task and that any speech outside of the realm of what I was looking for would be limited. As the recording sessions took place near the end of my stay in Malacca, I had been able to learn a little bit about speaking the language and its pronunciation. I also sought the help from a Kristang speaker who helped me with translating the questions. Fortunately, communication and understanding one another was never an issue, as I am a native speaker of English and the participants speak English either natively or as their primary tongue.

While comparing these languages, we must also consider the time in which the data was collected. These languages have been spoken for centuries, so there will have been changes that occurred diachronically. Although an important factor in the study of creole languages' evolutions, this thesis looks at the language as they have been spoken in the late 20th and early 21st Centuries. This can show how the languages have diverged from each other since creolization in the 16th Century, and differences or similarities they have between their input languages can show us how they have affected their structures. As I am looking at these languages through a viewpoint that argues language ecology plays a role in the structure of a language, comparison with the substrate languages is an important factor in the description of these languages.

Part Two: Description of Selected Features

Chapter 6

6.1: Introduction

In Part Two I will show selected features in Kristang, Makista, BC, and TC, as well as possible sources for these features. They are compared with each other and in also with lexifier and substrate structure. The features examined in this part are serial verb constructions, relative clauses and cleft constructions, short and long pronouns, verbal ellipsis after a TAM particle, verbal morphology from Portuguese influence, genitive patterns, passive, reduplication, constituent order, and question formation. Language ecology will play a role in this investigation, and examples from Chinese, Malay, and Portuguese will be found throughout as a reference point for what could be a cause in the features in these four languages.

The base of this chapter is primarily descriptive. Much of the descriptive work, especially those of the substrate languages, have been laid out by previous scholars, such as Matthews & Yip (1994) and Mintz (1994). Many of the examples of the creoles are from previous sources, especially Baxter (1988) for Kristang, and Maurer (2011) for BC and TC, although the organization of this previous work has been made to fit with the structure of this thesis and these chapters specifically. I also provide description of the data from other sources (Pillai, Maurer, Santos Ferreira), as well as some of the data I collected myself.

There are various reasons why each feature was chosen. Some were chosen because they are a distinctive part of the creoles not found in Portuguese, but are in the substrate languages. This includes serial verb constructions, reduplications, and ellipsis after a TAM particle. Others were chosen because of the way the feature works within each language, and ones where we can clearly see a parallel from Portuguese, Malay, or Cantonese, or a combination of them. This includes relative clauses and cleft constructions, genitive patterns, and passive. One section does not look at a particular feature per se, but looks at the way the lexifier has had influence on the four creoles. A large part of the syntactic structure of all four of these creoles is likely originated in, or at least reinforced by, the presence of substrate languages and its contact. There are some features that are more Portuguese-influenced, so this section looks at the differences between the creoles insofar as they have Portuguese morphology.

This chapter is structured by feature, with subsections on each creole, as we will see in the outline of sections of this chapter. Within each section we will see a basic description and theory of each feature discussed. We will then see how the input languages make use of these features. Then, we will look at examples from the creole languages. Each section will have a summary of how the phenomena in each creole compares with those of the input languages, and what it means for the influence of language ecology.

The first feature investigated in section 6.2 is Serial Verb Constructions, or SVCs. This is the phenomenon of two or more verbs within a single clause, which share tense and aspect, often share arguments, and have no overt subordination or coordination between them. This is a broad category that can serve different functions and can have differing structures. One of the major ways to divide SVCs in categories is between symmetrical SVCs, in which all verbs carry the same functional weight, and asymmetrical SVCs, in which one of the verbs comes from a closed class and serves a more syntactic role. SVCs can be found in all of the creoles investigated in this paper, and is also present in Malay and Cantonese. These are notably not part of most European languages, including Portuguese, and this can give us insight into the influence of the substrate languages on Kristang, Makista, BC, and TC. SVCs are also found in other creole languages and are sometimes used in creole universalist arguments, but I will also look at SVC examples in other non-creole languages, with a focus on those from the substrates of these four creoles.

Section 6.3 looks at relative clauses and cleft constructions. Cleft constructions are a structure that does not change the underlying semantic meaning of a phrase, but puts one of the constituents into focus. Consider the English example *John is eating a sandwich*. Using a cleft construction, we can put the subject *John* into focus: *It is John who is eating a sandwich*. Conversely, we can put the object *a sandwich* into focus: *It is a sandwich that John is eating*. Cleft constructions in English are marked by a relativizer, such as *that* or *who*, as in the previous examples. An Indo-European language, Portuguese also uses a relativizer *que* to mark a cleft construction. Although it is in the Austronesian family, Malay also uses a relativizer *yang* in a very similar construction to those in Portuguese or English. We will see that the creoles in the Malayo-Indonesian-speaking areas, Kristang, BC, and TC, use cleft constructions such as these. Of the four languages, Makista sticks out in cleft constructions because of its usage of copula clefts. This is the insertion of a copula to bring a constituent into focus, something that does not happen in Malay or Indonesian, but does happen in Sinitic languages including Cantonese. We will see examples where the Makista copula *sã* has scope

over an entire verb phrase, including the aspect markers, and how it behaves similarly to Sinitic examples.

Section 6.4 looks at what I call short and long pronouns. Like the previous section, this identifies a feature to highlight the distinctiveness of just one of the four languages, in this case Kristang. All four of the creoles have no distinction in their pronominal inventory of subject or object pronouns, something that Portuguese has but Malay and Cantonese lack. However, some Kristang personal pronouns have been found to have variants of pronunciation, one longer and one shorter. While not related to the syntactic relation or to the thematic role they play, they do seem to have influence from information structure and definiteness. An issue is that this phenomenon only shows up in spoken data, and written data is often not consistent in its recordings of the type of pronoun used. Written data for BC and TC do not show this type of variation between pronouns, nor do the writings of Santos Ferreira. However, performances including spoken Makista also do not have these types of distinction, whereas in Kristang this variation shows up often.

Section 6.5 investigates verbal ellipsis after a TAM particle. In these four languages there are three main pre-verbal particles. In Kristang these are *ja* for perfective, *ta* for progressive, and *lo(gu)* for future/irrealis. In the other languages these serve a similar purpose and have similar pronunciations. Etymologically, *ja* and *logu* come from Portuguese adverbs, while *ta* is from a verb. I investigated if this difference in etymology has an influence on verbal ellipsis, and it does seem to be the case. Again, I was only able to test this with certainty in Kristang, where I could ask consultants whether certain constructions were grammatical or not. However, in the other languages the original Portuguese functions of these words seem to have continued in their influence of the creoles.

Section 6.6 describes the genitive constructions. There are three forms of genitive that can be found in these four languages. The first is the usage of the preposition *di* "of", a direct descendant of Portuguese *de*. As in Portuguese, the word order is X di Y, where X is the possessee and Y is the possessor. All four of the creoles exhibit this construction. The second is the usage of a post-possessor particle, which in Kristang is *sa*. The word order in these constructions is Y sa X, (again where X is is the possessee and Y is the possessor). This is also found in all four of the creoles, but in Makista has limited usage. The third type is expressed with word order only, where the possessee is expressed before the possessor, with no prepositions or particles. This is only found in TC. I will show the distributions of these based on semantic categories laid out in Baxter & Bastos (2012). We will see the patterns

Baxter and Bastos showed with their research in Kristang, and how the other three creoles show similar patterns.

Section 6.7 explores passive constructions. There is more variation between the four languages on their usage of passive than in other features. Kristang tends to use word order or a serial verb to express passive, while Makista has examples which exhibit Portuguese influence. TC, having a strong Malay influence compared with its neighbor BC, makes use of morphology directly borrowed from Malay to express the passive.

Section 6.8 is on reduplication, which occurs in all four of the creoles. Reduplication is the phenomenon of repeating a word or morpheme twice, with a semantic or syntactic effect. Reduplication can occur on words of different syntactic categories and has different effects on each one. For instance, a reduplicated noun can indicate a plural, whereas a reduplicated adjective can indicate intensity. Like Serial Verb Constructions, reduplication is often cited as a typically creole feature, but it can also be found in Malay and Cantonese, and we will see how these have had an effect on the creoles' structures.

Section 6.9 is about the order of constituents in a phrase, and what they have in common with their substrate languages. It also features a discussion of two-object constructions, and how the dative is expressed in the four creoles.

Reference will be made to the Atlas of Pidgin and Creole Language Structures, or APICS, for relevant features, and where parallels exist in other creole languages (Michaelis et al 2013).

6.2 Serial Verb Constructions

Serial verb constructions are a common feature across the world's languages. They are the usage of two or more verbs as a single unit, used within a single clause without overt coordinating or subordinating morphology. There has been debate over what exactly constitutes a serial verb construction, or SVC, as there can be variation cross-linguistically or even within a single language as to how these multi-verb clauses appear. One definition of SVC is a phenomenon in which a single clause contains two finite verbs, each capable of being independent verbs on their own. They are generally used to denote the same action or two consequent actions. They also have the same tense and aspect as each other (Sebba 1987).

SVCs are considered to be the usage of two or more verbs in a single clause, but they do not make up only one grammatical category (Aikhenvald 2006). There are different

categories for SVCs that will be discussed below. However, Aikhenvald (2018) laid out properties that must exist for a construction to be considered a SVC:

a. A SVC consists of two or more verbs, each of which could also function as the sole verbs in a clause

b. There is no mark of dependency – such as coordination, subordination, or dependency of any sort – between the verbs within a SVC

c. A serial verb is monoclausal – it functions as a single predicate

d. The serial verb construction itself will have its own transitivity value

e. There is usually at least one core argument shared by all the verbs in a SVC

f. The SVC is conceived as describing a single event

Aikhenvald also pointed out that point f is a semantic rather than a syntactic property. What can be conceived of as a "single event" depends on context, both within a speech event and within a culture. In this sub-section I will be marking SVCs in **bold**.

Serial verb constructions can be classified into two broad classes which Aikhenvald terms symmetrical and asymmetrical serial verb constructions. Symmetrical SVCs consist of two or more verbs which generally imply a sequence, and thus the order in which they appear are pertinent to the message being conveyed. Baxter (1988) describes these as linking serializations. Such constructions describe an event that may consist of different segments or steps, but is conceived as a single even taken all together.

An example of a symmetrical SVC can be found in Kristang, describing a sequence of actions:

12. Kristang (Baxter 1988)

    eli   **pega   matah** la
    3SG   catch   kill   PRT
    "He catches and kills (children)"

This shows the sequence of catching and killing children, in the order in which they occur. Although the object of the clause is unstated, the arguments for both verbs are the same. He catches the children, and he kills the children.

Symmetric SVCs can also express simultaneous action.

13. Cantonese (Matthews 2006)

    keoi^3 **paak^3**-zyu^6 sau^2 **coeng^3** go^1
    3SG   clap-PROG   hand   sing   song
    "She claps her hands while singing"

As this is a simultaneous action, Cantonese allows for this to be stated in either order:

14. Cantonese (Matthews 2006)

keoi^3 **coeng^3**-zyu^6 go^1 **paak^3** sau^2

3SG sing-PROG song clap hand

"She claps her hands while singing"

In these examples the continuous verbal marker *zyu^6* is stated after the first verb in the sequence, although it takes scope over the whole sentence.

Another related function of symmetrical SVCs are cotemperal actions, which show two parts of a single event (Aikhenvald 2018: 74)

15. White Hmong (Aikhenvald 2018)

cov poj-niam **npaj** zaub **npaj** mov

CLF.PL woman:pl prepare vegetables prepare rice

"The women prepare the food" (lit. prepare.vegetables prepare.rice)

This example shows two actions that are culturally considered to be integral in the preparation of food. In addition to the wider context of the conversation or story, the cultural context also needs to be considered to investigate whether sub-events are considered by the speaker to be inseparable parts of a whole. The simultaneous preparation of vegetables and rice by the women is considered to be integral to the preparation of food, and therefore can be expressed with a symmetrical SVC (Jarkey 2015, Aikehnvald 2018).

Asymmetrical SVCs consist of a verb from a restricted class, often a verb of motion, adding a sense of direction or a tense-aspect meaning to the construction, called the major verb. The other verb is from a non-restricted class, called the minor verb. The major verb in an asymmetrical serialization can often be grammaticalized (Aikhenvald 2006). The structure of an asymmetrical serialization can resemble the usage of a modal or auxiliary verb. (This is especially the case with Kristang, Makista, BC, and TC which make use of pre-verbal elements to convey tense and aspect and have no inflection on the verb.) However, modal verbs have no other meaning aside from those used in multiple-verb clauses, but the major verb that can be found in an asymmetrical serial verb construction can be used in a single-verb construction. Baxter describes asymmetrical serial verb constructions as modifying serializations.

Although SVCs share tense and aspect, they do not necessarily share arguments. For instance, an SVC consisting of two transitive verbs can share a subject but have different

objects (as in sentence 4). Whether an SVC is symmetric or asymmetric does not usually affect the transitivity of the verbs in the construction.

For example, the Kristang verb *tomah* "take" is in this case used with the argument *faka* "knife" as an instrumental.

16. Kristang (Baxter 1988: 212)

eli	ja	**tomah**	faka	**kotrah**	kandri
3SG	PFV	take	knife	cut	meat

"He cut the meat with a knife"

Tomah has the same valency as the marker of an instrumental within an SVC as when it is used alone:

17. Kristang (Pillai 2011)

yo	tomah	stall	lah
1SG	take	stall	PRT

"I used to take a stall" (ie, I used to sell food at a stall)

Li and Thompson (1973) describe four types of serial verb that have a temporal or purposeful relationship, with examples in Mandarin:

Consecutive:

18. Mandarin (Li and Thompson 1973, in Sebba 1987: 2)

Zhang1-san1	**chuan1-shang4**	yi1fu	**tiao4**	zai4	di4-shang
Zhang-san	put-on	clothes	jump	on	floor

"Zhang-san put on his clothes and then jumped on the floor"

Purpose:

19. Mandarin (Li and Thompson 1973, in Sebba 1987: 2)

wo3men	**kai1**-hui4	**tao3lun4**	nei4ge	wen4ti3
we	hold-meeting	discuss	that	problem

"We are holding a meeting to discuss that problem"

Simultaneous:

20. Mandarin (Li and Thompson 1973, in Sebba 1987: 2-3)

Zhang1-san1	**qi2**	jiao3ta4che1	**zou3**	le
Zhang-san	ride	bike	leave	PFV

"Zhang-san left riding his bike"

Alternating:

21. Mandarin: (Li & Thompson 1973, in Sebba 1987: 3)

Zhang1-san1   tian1tian1   **hui4**-ke4   **xie3**   xin4

Zhang-san   every day   receive guest   write   letter

"Every day Zhang-san receives guests and writes letters"

All of these four categories are more symmetrical SVCs, but this demonstrates that within this subcategory there are further classifications that can be made. This illustrates that SVCs happen crosslinguistically, but the way they function can vary from language to language. The verbs within an SVC are not always contiguous. Intervening elements depend on the language. For instance, Sri Lanka Malay SVCs are contiguous:

22. Sri Lanka Malay (Nordhoff 2012: 320)

aajuth=yang   buurung   mà-**angkath**   **baapi**   su-**diyath**

dwarf=ACC   bird   PURP-lift   take.away   PAST-try

"The bird tried to carry the dwarf away"

This is an example of a contiguous SVC, which does not allow constituents to intervene between the verbal elements (Aikehnvald 2018: 92).

Contiguity of the verbs within an SVC also raises the issue of wordhood, whether the two elements are considered a single word or separate words. In languages that have SVCs they are not always limited to one-word or multiple-word constructions. Cantonese is an example of such a language (see sentences 2 and 3 for multiple-word SVCs with intervening constituents). Aikhenvald (2018: 94) points out that there are single-word SVCs in Cantonese:

23. Cantonese (Matthews 2006)

kéoi5   **sai3-laan6**-zo2   di1   saam1

3SG   wash-torn-PFV   PL   clothes

"Her washing the clothes left them torn" (lit. she washed (clothes) (they were) (torn))

However, some Cantonese single-word SVCs are not contiguous, as in the following example:

24. Cantonese (Matthews 2006: 82)

keoi5   **daa2**   m4   **laan6**   faai3   bo1lei1

3SG   hit   NEG   break   CLF   glass

"He cannot break the glass"

Cantonese SVCs are not strictly contiguous, although they are not strictly discontiguous either. They may be intervened by the modal marker *dak¹* or the negative *m⁴* as in sentence 7, or "prosodically light" elements (Matthews 2006: 82). Full constituents, such as noun phrases, cannot intervene in this type of SVC.

In the Luso-Asian creoles, the components of an SVC do not need to be contiguous, especially if each of the verbs has a different object. This is likely due to the fact that they have an SVO word order, so the object is usually stated directly after the verb. The TAM particle, however, appears before the whole verb phrase and not in between the verbs.

All four of the Luso-Asian creole languages discussed contain SVCs, both symmetrical and asymmetrical, despite the fact that such constructions do not exist in Portuguese. However, Malay and Cantonese exhibit examples of SVCs, as do their Austronesian and Sinitic relatives, respectively.

Influence from Portuguese can be seen as a cause of a reduction in the usage of SVCs. Aikhenvald (2018: 217) points out that in Tetun Dili from Timor Leste, there is a massive influx of Portuguese loanwords that consist of a single verb. Such an example is *informa* "to inform" which has replaced the SVC *fa hatene* (give know). As Portuguese increases as the language of official communication in Timor Leste, these types of verbs have incorporated themselves into Tetun Dili.

Three types of SVCs are described in APICS: directional SVCs with "come" and "go" (Chapter 84, Maurer and Michaelis 2013), SVCs with "take" (Chapter 85, Maurer 2013), and SVCs with "give" (Chapter 86, Maurer 2013). While none of the South Asian Portuguese lexified creoles of Diu (Cardoso 2013), Korlai (Clements 2013) and Sri Lanka (Smith 2013), are attested to have directional SVCs with "come" and "go", all of the creoles in this thesis exhibit their use. Other creoles that have different lexifiers but are also in the Malay-speaking region, Singlish (Lim & Ansaldo 2013) and Singapore Bazaar Malay (Aye 2013) do have this feature.

6.2.1 Serial Verb Constructions in the Substrates

In this section we will see examples of SVCs in the substrate languages for Kristang, Makista, BC, and TC. Serial verb constructions are an important part of Cantonese grammar. Cantonese SVCs are often used to express what in English and other European languages is expressed with prepositions (Matthews & Yip 1994, 143).
Asymmetric SVCs in Cantonese usually have the major verb first.

25. Cantonese (Matthews & Yip 1994: 143)

> ngo5 **bong1** lei5 **da2**-din6wa2
>
> 1SG help 2SG call-phone
>
> "I'll phone for you"

While the major verb *bong1* in isolation means "help", in this construction it is used in a benefactive sense, and does not convey the meaning of literally helping someone.

Another asymmetric SVC in Cantonese uses *jung6* "use" as an instrumental:

26. Cantonese (Matthews & Yip 1994: 144)

> keui5dei6 **jung6** di1 cin3 **heui3 maai5** lau2
>
> 3PL use CLF money go buy flat
>
> "They're buying a flat with the money"

As in Sentence 14, the major verb comes first.

These examples of asymmetrical SVCs are discontiguous. As the object of both verbs in Sentence 15 is different, and the basic word order of Cantonese is SVO, the objects come directly after the verb for which they are arguments. Despite this discontiguity they do constitute a single clause without overt coordination.

Coordination can be overtly expressed in Cantonese.

27. Cantonese (Matthews & Yip 1994: 290)

> ngo1dei6 sin1 **sik6**-faan6, gan1zyu6 **heui3 tai2** hei3
>
> 1PL first eat-food following go see film
>
> "We'll have dinner first, then go and see the film"

This example uses the conjunction *gan1zyu6* "following", indicating a sequence of events.

28. Cantonese (Matthews & Yip 1994: 290)

> ngo1dei6 **haau2 jun4** si5 zau6 **heui3 waan2**
>
> 1PL take finish exam then go play
>
> "When we've finished exams we're going to have some fun"

The conjunction *zau6* "then" is also used to coordinate events in a sequence. Since taking and exam and having some fun are not generally considered to be connected, there is no serialization but rather coordination.

As is common on other languages, Cantonese also uses *go* and *come* in directional SVCs.

29. Cantonese (Matthews 2006: 76)

> lei5 **lo2** di1 **saam1** lai4
>
> 2SG take PL clothing come
>
> "Bring some clothes"

In this type of asymmetric SVC, the major verb *lai* "come" comes after the minor verb.

In the languages of Malaysia, the presence of serial verb constructions is not unique to Kristang, as Bazaar Malay has similar patterns:

Motion verbs:

30. Bazaar Malay (Aye 2006: 119)

hari-hari	**jalan-jalan**	sini	**datang**	Geylang Balu	Kampong
every day	walk-walk	here	come	Geylang Bahru	village

"Every day, I walk here to Geylang Bahru"

"Give" constructions:

31. Bazaar Malay (Aye 2006: 119)

dia	**kasi**	**tahu**	lembu, Gila!	*Hey*	lu	mao	tengok saya	*ah*.
3SG	give	know	ox mad	hey	2SG	want	look.at 1SG	PRT

"He told the ox, the mad guy, you look at me"

Linking serialization/purpose:

32. Bazaar Malay (Aye 2006: 119)

diaorang	selalu	**datang**	sini	**cari**	murah	punya
barang						
3PL	always	come	here	look.for	cheap	GEN
thing						

makan-makan	pakai-pakai	semainlah
eat-eat	use-use	etc.

"They always come here to look for cheap food and things, etc"

This will be discussed further in the next section 6.2.2.

6.2.2 Serial Verb Constructions in Kristang

Within Kristang, Baxter (1988: 211) describes both symmetrical ("linking") and asymmetrical ("modifying") SVCs.

Symmetrical serialization represents a sequence of actions as subparts of a whole:

33. Kristang (Baxter 1988: 211)

eli	**pega**	**matah**	la
3SG	catch	kill	PRT

"He catches and kills (children)"

Asymmetrical serialization is the description of a single, indivisible event. In asymmetrical serialization, one of the verbs is the main verb, and the other "does not retain its full semantic

value" (Baxter 1988: 211). Baxter also notes that when there is no intervening constituent between verbs, the first verb's stress pattern changes, from stress on the ultimate syllable to the penultimate. Baxter then describes modifying serialization further, using the following categories:

1. Relator Role
2. Aspectual modification
3. Modal modification
4. Passive indicator
5. Serialization involving *da* (give)
6. *Ke* (want) serialization

Relator role serializations involve a modifying verb functioning "as a relator indicating the function of an NP within the clause." One such example is the use of *beng* (come) and *bai* (go) to indicate a goal NP

34. Kristang (Baxter 1988: 212)

eli	ja	**kure**	**bai/beng**	kaza
3SG	PFV	run	go/come	house

"He ran home"

Another example of a relator serialization is the use of the verb *tomá* (take), which marks an instrumental NP:

35. Kristang (Baxter 1988: 212)

eli	ja	**tomah**	faka	**kotrah**	kandri
3SG	PFV	take	knife	cut	meat

"He cut the meat with a knife"

Aspectual modification is when the major verb modifies the aspect of the main verb:

36. Kristang (Baxter 1988: 213)

kora	yo	ja	**chegah**	nalih	eli	ja	kaba	**bai**
when	1S	PFV	arrive	there	3SG	PFV	finish	go

"When I arrived there he had gone"

Modal modification is when the major verb modifies the mood of the main verb:

37. Kristang (Baxter 1988: 213)

eli	ja	**acha**	**bai**	Singapura
3SG	PFV	receive	go	Singapore

"He got to go to Singapore"

The passive indicator is the verb *toká* (touch) used as passivization on the main verb, but only when "the patient is adversely affected":

38. Kristang (Baxter 1988: 213)

    akeh   kaza   ja   **toka**   **kemah**
    that   house   PFV   touch   burn
    "The house got burnt"

Baxter describes two types of serialization with *da* (give). One is to give the meaning "allow", when followed by an object:

39. Kristang (Baxter 1988: 214)

    pa   nadi   **da**   ku   yo   **sai**   fora
    father   NEG   give   to   1SG   go.out   outside
    "Father won't let me go out"

The other type is when the addition of *da* to a transitive verb, the resulting verb is ditransitive. For instance, the following example uses *da* (give) with *kumí* (eat) to create "feed":

40. Kristang (Baxter 1988: 214)

    papa   galinya   olotu   **da**   **kumih**   ku   olotu   pa
    bizah
    porridge   chicken   3PL   give   eat   to   3PL   for
    watch
    "They feed them chicken porridge for watching (the dead)"

There is a third type of serialization with *da*. Kristang uses the particle *ku* "with, and" as an object marker for human objects. This is uses for both direct and indirect objects.

41. Kristang (Laub fieldwork, 2017)

    yo   da   lus   ku   bos
    1SG   give   light   to   2SG
    "I give a light to you"

It is not grammatical for non-human objects to take this marker, and there is usually no marker before the object. However, the verb *dah* can also be used as an object marker for inanimate objects.

42. Kristang (Laub fieldwork, 2017)

    e   ta   **dali**   kosi   **dah**   kadera
    3SG   PROG   hit   kick   give   chair
    "S/he is kicking the chair"

Ke serialization (Baxter 1988: 215) is described as the major verb of a reduced subordinate clause. This serial verb construction cannot take TAM markers or be negated, and may not have an overt subject:

43. Kristang (Baxter 1988: 215)

eli	midu	**ke**	**bai**	bos	sa	kaza
3SG	fear	want	go	2SG	GEN	house

"He is afraid to go to your house"

Ke serialization is generally preceded by a predicate that is emotive, such as the example above.

Baxter's categories of verb serialization provide a good insight into Kristang-specific patterns. However, verb serialization is a phenomenon found across language families, as well as in other creoles. It is important to look at these patterns of serialization within the context of other descriptions of serial verbs typologically and cross-linguistically.

All of these examples would fall under Baxter's category of linking (symmetrical) serialization, but there are examples in Kristang of linking serialization in various categories from Li & Thompson.

44. Kristang (Marbeck 2004: 63)

mulher	di	kaza	lo	**tomah**	akeh	pesi	**fazeh**	kari,
frizih								
woman	of	house	FUT	take	this	fish	make	curry	fry
fazeh	**sambal**		ka	**fazeh**	sup			
make	sambal		or	make	soup			

"The women of the house would make curry with this fish, fry it for sambal or for soup"

Example 41 contains two serial verb constructions, one symmetrical and one asymmetrical. The first one, *tomah akeh pesi fazeh kari*, uses the verb *tomah* (take) marking its object *pesi* (fish) as instrumental. The second serial verb construction, *frizih fazeh sambal* "fry (the fish) to make sambal" is an example of symmetrical serialization.

Although both examples 30 and 41 have instances of symmetrical serialization, they do have some notable differences. Both verbs in example 30 are understood to have the same object (although not explicitly stated), while example 41 has intervening objects between the first and second verb. Using Li & Thompson's categories, example 30 is interpreted as a consecutive action, while 41 is used for purpose, as the fish is fried to make soup.

Patterns of verb serialization in Kristang are not dissimilar to those of other languages that feature this phenomenon. Sebba (1987) described serial verb constructions in Sranan and other languages and noted seven categories of them:

1. Coordinating and subordinating serialization
2. "Take" constructions
3. "Give" constructions
4. Other three-place verbs
5. Motion verbs
6. Lexical combinations
7. Comparative constructions

Much like in Kristang, the "take" constructions are used as an instrumental relator (sentences 21, 30). Structures identical to that found in Kristang can be found in other languages:

45. Yoruba (Stahlke 1970, in Sebba 1987: 168)

mo	**fi**	ada	**ge**	igi
1SG	take	machete	cut	tree

"I cut the tree with a machete"

46. Saramaccan (Sebba 1987: 170)

hen	**tei**	di	faka	**koti**	di	bee
3SG	take	the	knife	cut	the	bread

"He cut the bread with the knife"

"Give" constructions (see sentence 26) also have parallels in other languages:

47. Ijo (Williamson 1963, in Sebba 1987: 178)

egberi	**gba**-ni	u-**piri**
story	say-to	him-give

"Tell him a story"

The presence of *bai* (go) and *beng* (come) in serial verb constructions in Kristang (see example 31) should come as no surprise. They are near universal in languages with such constructions, and they occur as a complement to the first verb, indicating motion and the direction thereof (Sebba 1987: 184). Such constructions can be found in other languages:

48. Akan (Christaller 1875, in Aye 2006: 119)

miguaree	**mebaa**	mpono
I-swim-PAST	I-come-PAST	shore

"I swam to the shore"

49. Jamaican Creole (Winford 1993, in Aye 2006: 119)

dem	a	**waak**	**go**	a	maakit
3PL	PROG	walk	go	LOC	market

"They're walking to the market"

Examples from other languages and other categorizations of serial verb constructions show that Kristang has phenomena that typologically resemble those of other languages. Similar constructions in Bazaar Malay (see examples 30-32) suggest a possible source for serial verb constructions in Kristang. Although found in other creole languages, the presence of such phenomena in both Bazaar Malay and Kristang show the influence of their input languages in this particular feature.

It is clear in the previous examples that SVCs in Kristang do not require contiguity. Indeed, sentence 35 shows that when there are two transitive verbs with different objects, the word order rules of Kristang require that there be intervening elements in the SVC.

However, SVCs without intervening constituents do exhibit a shift in the stress of the first verb. The verb *tokah* "touch" when used on its own has stress on the ultimate syllable, much like the vast majority of Kristang verbs. However, when it is used in the sense of an adversive passive, the stress shifts off of the ultimate syllable (as in *toka kemah* in sentence 38). Sentence 35, which has the intervening constituent between the verbs, has the same stress patterns as if these verbs were used in isolation.

6.2.3 Serial Verb Constructions in Makista

Serial verb constructions are also found in Makista. They occur both symmetrically and asymmetrically. Examples of symmetrical serialization can be found in José dos Santos Ferreira's writings in Makista:

50. Makista (Santos Ferreira 1967: 85)

tudo	dia	pramecido	**pegá**	lata	**corê**	casa-casa	**vendê**
every	day	early in the morning	take	tin	run	house-house	sell

merenda
snack

"Every day early in the morning I took cans, ran to the houses, and sold snacks"

51. Makista (Santos Ferreira 1967: 97)

Mána-Chai	já	**vai**	botica	**comprá**	"bacalhau"
Mána-Chai	PFV	go	shop	buy	salt cod

"Mána-Chai went to the store to buy salt cod"

52. Makista (Santos Ferreira 1973, 148)

vôs **levá vai tambá**

2SG take go repair

"You take it and go fix it"

53. Makista (Santos Ferreira 1973, 155)

querê **vai** Inferno **buscá** iou... **jugá** dado co iou?

want go hell seek 1SG play dice with 1SG

"Will you go to hell and search for me, and play dice with me?"

Example 50 is an example of consecutive symmetrical serialization (compare Kristang example 33 and Mandarin example 18), while 51 is purposeful symmetrical serialization (compare Kristang example 41 and Mandarin example 19). Examples 50 and 52 also have triple serialization, where there are three verbs in the clause, representing subsequent actions. The older corpora also have instances of symmetrical serialization:

54. Makista (Pereira 1899: 240)

corê rua **pidí comê**

run street ask eat

"Walking the streets asking for food"

55. Makista (Barreiros 1944: 138)

pegá mão de sua pai **tomá bença** dentro de greza

hold hand of his father take bless inside of church

"Hld his father's hand to bless him in the church"

56. Makista (Barreiros 1944: 484)

iou **sai vem** casa **olá** mamã

1SG exit come house see mother

"I came out to come to the house to see mother"

These examples have different objects, but the same subjects. Other symmetric SVCs have the same object:

57. Makista (Barreiros 1944: 358)

commo nôsso fatiáng que **vai- vem** Hong-Kong assim

like our boat REL go come Hong Kong such

"Like our boats which go to and come from Hong Kong"

These likely were influenced by Cantonese structure, as triple serialization can occur in Cantonese and other Sinitic languages, and is not attested in the Bazaar Malay literature:

58. Cantonese (Yip & Matthews 2000: 67)

bong1 ngo1 **wan2** zi1-liu2 **se2** bou3-gou3

help 1SG find material write report

"Get some material for me to write my report."

Examples 50 and 51 have objects intervening between the verbs, which may lead one to classify this as an instance of coordination. However, Cantonese exhibits serial verb constructions with the same structure and word order:

59. Cantonese (Matthews 2005: 72)

keoi5 dei6 seng4jat6 **haang4**-gaai1 **tai2**-hei3

3PL always walk-street see-movie

"They're always going out shopping and going to movies"

This is in contrast to overt coordination:

60. Cantonese (Matthews 2005: 72)

keoi5 dei6 seng4jat6 **haang4**-gaai1 tung4maai4 **tai2**-hei3

3PL always walk-street and-also see-movie

"They're always going shopping and going to movies to"

It also contrasts with intervening particles:

61. Cantonese (Matthews 2005: 72)

keoi5 dei6 **haang4**-gaai1 laa1, **tai2**-hei3 laa1

3PL walk-street PRT see-movie PRT

"They go shopping, go to movies…"

Makista also uses coordination to express actions in a sequence.

62. Makista (Santos Ferreira 1973: 157)

dôs americano jóvi já **intruzá**, cavá já **dále**

Two American young PF deceived then PFV attack

"Two American boys deceived her, then hit her"

Example 62 uses *cavá* in a similar way as Cantonese uses *tungmaai* in sentence 60, suggesting that Cantonese has influenced the distinction between serialization and coordination.

There are examples in the Pereira corpus with non-overt coordination, but with clear examples of SVCs within the sentence. The non-serialized verbs are in italics, while the serialized verbs are also in bold.

63. Makista (Pereira 1899: 260)

gora	elle	tá	mutu	contente	naquelle	cadera,	já nunca
now	3SG	PROG	very	happy	in.that	chair	PFV never

lembrá	ôtro	cusa	más	si	nam	*assisti*	festa,	*companhá*
remember	other	thing	more	if	NEG	attend	party	follow

pruciçám,	*tentá*		tudo	janella janella,	***vai***	***vem***	palacio,
procession	pay.attention	all	window-window		go	come	palace

sumiá	cóvi coco	na	lugar	di	guardá	vaca	di	puçá	careta
put	cabbage	in	place	of	keep	cow	of	pull	car

di	suzo
of	garbage

"Now he is very happy in that chair, he doesn't remember anything other than if he does not attend parties, following processions, paying attention to all the windows, going to and coming from the palace, and putting cabbage in the place where they keep the cows that pull garbage carts"

64. Makista (Pereira 1899: 324)

cêdo,	cêdo,	já	*tem*	na	rua,	*tirá*	telhêro	di	botica,	*rancá*
early	early	PFV	exist	in	street	take	shed	of	store	rip.ou

pagôde	di	porta	di	china china,	***cortá***	rua	***fazê***	drêto,	qui
pagoda	of	gate	of	Chinese-Chinese	cut	street	make	right	REL

fazê	gusto	ólá
make	like	see

"Very early, they are in the streets, they take the sheds from the stores, rip out the pagodas of the Chinese gates, cut the streets to make them right, so whoever makes it likes what they see"

These examples show that several verb phrases are coordinated in a long list, but the ones that are serialized have a link to them. Example 63 has the SVC *vai vem palacio* "going to and coming from the palace", and Example 64 has *cortá rua fazê drêto* "cut the streets to

make them right". What these have in common are that each one shares their object, being *palacio* in 63 and *rua* in 64.

Serialization with *virá* "turn" is another structure found in Makista. *Virá* on its own is a transitive verb, meaning "to turn" or "to transform":

65. Makista (Santos Ferreira 1967: 20)

têm	hóme-hóme,	olá	unga	vassóraco	saia	tamêm	pôde	virá	ôlo
have	man-man	see	one	broom	with	skirt	also	can	turn	eye

"There are men who would even turn their eyes if they saw a broom with a skirt"

66. Makista (Pereira 1995: 323)

governo	nôvo	sam	capaz	e	já	virá	tudo
Governor	new	COP	capable	and	PFV	transform	everything

"The new governor is capable and has changed everything"

In an SVC, *virá* appears first and marks the action as subsequent to previous actions.

67. Makista (Santos Ferreira 1967: 85)

tánto	qui	já	prendê	fazê	merenda,	qui	cavá	já	**virá**
									vendê
much	that	PFV	learn	make	snack	that	finish	PFV	turn	sell
					merenda				
					snack				

"I learned much about making snacks, that I then began to sell them"

68. Makista (Santos Ferreira 1967: 136)

iou	**virá**	**respondê**
1SG	turn	respond

"I then responded to him"

69. Makista (Santos Ferreira 1967: 95)

justo	cavá	guéra,	panhá	na	tudo	sapéca,	já	**virá**	**vai**	Portugal
just	after	war	take	in	all	money	PFV	turn	go	Portugal
ficá										
stay										

"Just after the war, he took all his money and then went to live in Portugal"

70. Makista (Pereira 1899: 59)

eu	**virá**	**fállá**	cô	elle
1SG	turn	speak	with	3SG

"I then spoke to him"

Virá appears to be a verb and not an adverb due to its position in the sentence. Just as in Kristang, verbs can be marked with a pre-verbal TAM marker, with no intervening elements (Pinharanda Nunes and Baxter 2004). Compare with the adverb *justo* "just" below:

71. Makista (Santos Ferreira 1967: 97)

Janjám	justo	ta	contá	qui	êle	na	mar	Vemêlo,	já
Janjám	just	PROG	tell	that	3SG	on	sea	red	PFV

olá	pesse	"voador"
see	fish	flying

"Janjám is just telling that on the Red Sea, he saw flying fish"

In this case, *justo* falls outside of the scope of the TAM marker, while in example 69, *virá* immediately follows the marker, putting it within its scope.

Another marker of subsequent action is *cavá*, when used as a verb meaning "finish". It can appear sentence-initally and act as an adverb:

72. Makista (Santos Ferreira 1967: 91)

cavá	vôs	falá	Maria	sã	feióna
then	2SG	say	Maria	COP	very.ugly

"And then you say that Maria is very ugly"

Serialization with *ramatá* (finish) occurs in Makista. As a main verb, *ramatá* is a transitive verb meaning "finish":

73. Makista (Santos Ferreira 1967: 128)

ramatá	vôs	sua	estória
finish	you	GEN	story

"Finish your story"

Ramatá in a serial verb construction marks the clause as perfect:

74. Makista (Santos Ferreira 1973)

puliça	têm-qui	vai	bebê	café.	Si	temá,	**limpá**	**ramatá**!
police	have-to	go	drink	coffee	if	insist	clean	finish

"The police have to go drink coffee. If you insist, finish cleaning"

75. Makista (Santos Ferreira 1967: 128)

> **fazê**   **ramatá**
>
> do   finish
>
> "Do it to the end"

76. Makista (Santos Ferreira 1967: 128)

> já   **cunvidá**   **ramatá**
>
> PFV   invite   finish
>
> "I ended up inviting you too"

77. Makista (Santos Ferreira 1967: 95)

> já   **rabixá**   **ramatá**   Maria, ilôtro   sua   fila
>
> PFV   bring   finish   Maria 3PL   GEN   daughter
>
> "They ended up bringing Maria, their daughter"

It has a sense of a perfective aspect, but comes second in a serial verb construction. It may also co-occur with the perfective marker *já*.

Similar constructions are found in Papuan Malay, a Malay-lexified creole:

78. Papuan Malay (Donohue 2011: 420)

> so=lapar   jadi,   **makan**   itu   **habis**
>
> PFV=hungry this   eat   that finish
>
> "You're hungry, so eat it all up"

Another example of an SVC using finish or complete as the major verb, with the major verb coming last is from the Papuan language Kalam.

79. Kalam (Pawley and Lane 1998: 206)

> mnm **ag**   **d**-p-al
>
> word   say   get/complete-PRF-3PL
>
> "They have finished talking"

While Kalam is not a creole, we can see a similar word order using *complete* as the major verb, as the second verb in the SVC.

There thus appears to be a typological tendency for the major verb *finish* to be the second part in an asymmetric serial verb construction, especially when conveying the meaning of perfective.

Alleyne (2000: 127) notes that:

> Saramaccan has the serial verb 'finish', in the form *kaba*, in post-verbal
> position. Jamaican still has it (in the form *don*) in that position, but the ecology
> of its use indicates a old form which is archaic and becoming obsolete. Syntactic
> reanalysis is underway in Jamaican and *don* is in the process of being
> reinterpreted as an aspect marker.

The Saramaccan verb *kaba* "finish" is derived from Portuguese *acabar* "finish", just as the equivalents in Makista and Kristang are. However, its usage seems to be more similar to the Makista use of *ramatá* than to *cavá*, coming at the end rather than in the beginning. Additionally, *cavá* has a meaning more akin to "afterwards" than to "finish" in the same way *ramatá* does.

Siegel (2007: 180-1) also notes the influence of Cantonese onto the aspect of a creole, that of Hawai'i Creole. An optional perfect marker in Hawai'i Creole is *already*, from the English, and it occurs post-verbally:

80. Hawai'i Creole (Masuda 1998: 232, as cited in Siegel 2007: 180)

da	tako	no	come	in	already	Olowalu-	side
the	octopus	NEG	come	in	PFV	Olowalu	area

"The octopus doesn't come to the Olowalu area"

He also notes that such a construction also exists in Singapore English:

81. Singapore English (Platt and Weber 1980: 66, as cited in Siegel 2007: 181)

I only went there once or twice already

In these instances, he argues that *already* has come to represent the perfect because it has a parallel in the Cantonese verbal marker, which comes post-verbal. He also points out that the reason that this particular English adverb has shifted to be an aspectual particle, as opposed to others, is that it often comes after a verb in English, whereas other adverbs tend to come before the verb:

> He already went there.
> He went there already.
> He usually goes there.
> ?He goes there usually.

It is this convergence of English structure and lexicon, with Cantonese structure that allowed for this structure to emerge in Hawai'i Creole and Singapore English.

In the case of Makista *ramatá*, this could also be a reflection of Cantonese influence. Makista already has a perfective particle, *já*, which in Portuguese occurs pre-verbally. Since

there is also an identical structure in Kristang, BC, and TC, it is almost certain that this structure predates the splitting off of Makista. With *já* already extant in the language, it would not be likely for it to shift to a post-verbal position in a short amount of time. However, Makista had in its grammar the possibility of serial verb constructions, in which a full verb can be used in a multi-verb construction to have some aspectual or modal implication. Following Siegel's model, it could be the case that *ramatá* was able to enter into such a construction by Makista allowing for these constructions. Its location at the end of an asymmetric serial verb construction is also a reflection of the Cantonese placement of aspect particles after the verb. It is thus the convergence of the structure of Makista allowing for such an innovation to take place, the lexical input of Portuguese verbs (or the verbs that already existed in the proto-Malaccan creole), and the word order input from Cantonese, that led to the instance of V2 *ramatá* structures. Similar structures in other languages indicate that there could be a typological influence as to why this particular SVC is structured this way, but it is the circumstances of the Makista language and its linguistic ecology that led to its emergence.

Serialization with *querê* (want) has been found in the corpora:

82. Makista (Santos Ferreira 1973, 155)

querê	**vai**	Inferno	buscá	iou…	jugá	dado	co	iou?
want	go	hell	seek	1SG	play	dice	with	1SG

"will you go to hell and search for me, and play dice with me?"

Querê means "to want", but here appears to have the additional meaning of marking the future.

Detges (2000) writes about the grammaticalization of languages, and argues that many processes that are marked as creolization are actually grammaticalization. One feature cited is the formation of a future tense, one example of which is "I want to do it → I'll do it" (Detges 2000: 137). Although Makista already has a future marker in the form of *lôgo*, this example with *querê* appears to be the beginning of the formation of another structure that indicates future.

We find another parallel with Chinese languages in contact, in the case of Baba Malay, spoken by the Peranakan community of the Straits of Malacca. With Malay as one of the input languages, it uses pre-verbal aspectual markers, much like the Luso-Asian creoles in this thesis. As Malay is considered the lexifier, they take the same form in Baba Malay (Ansaldo, Lim, and Mufwene 2007: 216). This also includes a general future marker, from the Malay *nanti* "later".

83. Baba Malay (Ansaldo, Lim, & Mufwene 2007: 217)

nanti satu haru li pun nanti charek dia

later one day 2SG also FUT seek 3SG

"You too will need her one day"

This is a very similar structure to the usage of *lôgo* in Makista, not to mention the future-irrealis markers in the other creoles.

Ansaldo and Matthews (1999: 59) mention that there is an alternative future marker, *mo*, which marks an immediate or an intentional future. They point out that *mo* is derived from Malay *mau* "to want".

84. Baba Malay (Ansaldo, Lim, & Mufwene 2007: 217)

lu pun mo charek dia

2SG also FUT seek her

"You are also going to need her"

This could be from Hokkien influence, where there is a distinction between a general future and one of immediacy (Pakir 1986, Bodman 1987, Ansaldo, Lim, & Mufwene 2007)

85. Hokkien (Ansaldo, Lim, and Mufwene 2007: 217)

î běq laí la

3SG FUT come PRT

"He is coming"

86. Hokkien (Ansaldo, Lim, & Mufwene 2007: 217)

î těq běq laí la

3SG FUT come PRT

"He is on his way now"

When this language has a distinction between a general and an immediate future, the language which it influences can also have such a distinction. As we can see in the Makista examples, there is usage of *querê* as an intentional or immediate future, while *lôgo* still exists as a future marker alongside it. What is interesting to note is that the parallels also extend to the lexical origins of the words, with Portuguese *logo* and *querer* and Malay *nanti* and *mau* meaning "later" and "to want", respectively. Donohue also points out that Singaporean Malay has serial verbs, as do some other varities of Western Malay, possibly due to Chinese influence, and they are not present in standard Malay.

The verb *vai* "go" is also used as a marker for future, and both this and the FI marker *lôgo* can be used by the same speaker, as seen in the following example:

87. Makista (Barreiros 1944: 357)

Macau **vai** **fazê** tanto festa; aqui tá fallá qui Macau lôgo tem

Macau go do many festival here PROG speak REL Macau FI

have

quatru dia di festa

four day of festival

"Macau is going to have many festivals; here they are saying that Macau will have four days of festival"

Serialization with verbs of motion is also found in the earlier corpora:

88. Makista (Pereira 1899: 517)

vai **cazá** co moça

go marry with girl

"Go marry the girl"

89. Makista (Barreiros 1944: 574)

Nhun Chico, **vem** **ovi** iou falá

boy come see 1SG talk

"Boy, come see me and talk to me"

Mandá (send, order) is used in an asymmetric SVC, found in several examples in the earlier corpora:

90. Makista (Pereira 1899: 57)

eu já **mandá** **sentá**

1SG PFV order sit

"I ordered [him] to sit"

91. Makista (Pereira 1899: 192)

mandá **abri** dicionario

oder open dictionary

"[he] ordered the dictionary to be opened"; "[he] ordered [us] to open the dictionary"

92. Makista (Pereira 1899: 323)

mêstre nômquêro que eu fazê mutu força, e **mandá**

teacher neg.want REL 1SG do much force and order

tomá ninho di pastro

take child of bird

"The doctor did not want me to push too hard, and prescribed bird's nests"

93. Makista (Barreiros 1944: 348)

já	**mandá**	**chómá**	tudo	daya	Egitana
PFV	order	call	all	power	Egyptian

"[the Pharaoh] summoned all of the power of Egypt"

94. Makista (Barreiros 1944: 481)

este	tem	que	ser	Josefina,	io	já	**mandá**	**chomá**
this	have	REL	be	Josefina	1SG	PFV	order	call

"This must be Josefina, whom I've called over"

6.2.4 Serial Verb Constructions in Batavia Creole

BC has instances of SVCs, both symmetrical and asymmetrical. Asymmetrical serializations include usage of *da* "to give" are found in BC:

95. Batavia Creole (Maurer 2011: 131)

isti	belu	**da**	**sabe**	kung	ile	ki
this	old.man	give	know	to	3SG	REL

"The old man told him that…"

96. Batavia Creole (Maurer 2011: 132)

akel	belu	**da**	**intindi**	ki
this	old.man	give	understand	REL

"The old man explained [to him] that…"

This give-construction occurs in the other creoles, and is also common crosslinguistically in instances of asymmetrical serialization. Although this may resemble a verb-complement construction, we can see that it functions as one SVC, especially in example 68. In this example, both arguments of the verbal construction appear after *da sabe*: *kung ile* "him", and *ki*, the relativizer introducing the subordinate clause.

Symmetrical serialization also occurs in BC. The two components of symmetrical SVCs often may have different objects, but they share a subject and generally convey a single or connected action.

97. Batavia Creole (Maurer 2011: 129)

noybu	**irgi**	di	sonu	**rakuli**	fatu
groom	get.up	of	sleep	set.order	belongings

"The bridegroom awoke and set his belongings in order"

98. Batavia Creole (Maurer 2011: 139)

lanta komer **tridji** na medja

raise food bring LOC table

"Bring the food to the table"

99. Batavia Creole (Maurer 2011: 141)

supri **fika** kaladu nomas

suffer become quiet only

"Just suffer it and keep quiet"

These examples include SVCs that contain both transitive and intransitive verbs, but share the same aspect and subject. Unlike in Makista, there are no instances of symmetrical serialization with more than two verbs.

These examples show that symmetrical serialization could occur in BC, although coordination was much more common.

Another multiple-verb clause found in BC is the usage of *bi desa nos* "come let us" as a first person plural imperative.

100. Batavia Creole (Maurer 2011: 139)

bi desa nos apusta

come let 1PL bet

"Come on, let us bet"

101. Batavia Creole (Maurer 2011: 141)

bi desa nos fadji asi

come let 1PL do so

"Come on, let us do it like this"

This is not an example of an SVC, but it is notable in its similarity with Malay and Indonesian use of *mari kita* as a first person plural imperative.

6.2.5 Serial Verb Constructions in Tugu Creole

TC has examples of asymmetrical and symmetrical SVCs, with more Malay- or Indonesian-influenced structures than BC does. Asymmetrical serialization is found in the TC corpus with the following verbs: *get, make, touch, give, go,* and *turn*. *Get*-serialization indicates a passive-like construction, or a shift of the subject to a patient role:

102. Tugu Creole (Maurer 2011: 147)

filu	filu	**acha**	**prende**	dikausa	les,	iskrebe	mas

konta

hild	child	get	teach	of.cause	read	write	and

count

"The children are taught reading, writing and counting"

There is an example of *get*-serialization that seems to act as an emphasis for ability:

103. Tugu Creole (Maurer 2011: 163)

podi	nos	**acha**	**ponta**	pertu	pertu	kal	grandi
can	1PL	get	shoot	near	near	REL	big

"We can shoot big ones from very close"

Make-serialization indicates a causative construction:

104. Tugu Creole (Maurer 2011: 148)

ele	**fay**	**say**	mas	dianti	sua	kote kote	neli
3SG	make	come.out	more	early	GEN	grains	rice

"[The rice] drops first its rice grains"

105. Tugu Creole (Maurer 2011: 150)

poy	**difila**	**fay**	askura
can	PASS.pound	make	hulled.rice

"It can be pounded in order to make hulled rice"

Touch-serialization indicates an obligation or an adversative passive.

106. Tugu Creole (Maurer 2011: 150)

i	**toka**	**mara**
REL	touch	bind

"[The ears of rice] which have to be bound"

Give-serialization involves changing intransitive verbs to transitive, or can convey a meaning of permission:

107. Tugu Creole (Maurer 2011: 152)

iste	pesua	mas	**da**	**sabe**
DEM	secret	more	give	know

"We will reveal this secret"

108. Tugu Creole (Maurer 2011: 159)

sua	tawdu	podi	**da**	**kumi**	kabalu
GEN	tawdu	can	give	eat	horse

"Its *tawdu*[9] can be fed to the horses"

109. Tugu Creole (Maurer 2011: 177)

kusir,	**da**	**pasa**,	pasta
coachman	give	go	enough

"Coachman, go ahead, right now"

Go-serialization indicates the direction of the action:

110. Tugu Creole (Maurer 2011: 162)

bi	nos	**anda**	**ponta**	porku	na	matu
come	1PL	go	shoot	pig	in	forest

"Let us go hunting wild boars in the forest"

111. Tugu Creole (Maurer 2011: 164)

teng	djenti	ki	kontenti	**anda**	**peska**
COP	people	REL	HAB	go	fish

"There are people who go fishing"

Turn-serialization expresses doing an action again. This type is different in that the major verb occurs second rather than first.

112. Tugu Creole (Maurer 2011: 158)

djenti	lanta	ayak	**supra**	**torna**
people	pick.up	sieve	blow	turn

"We take the sieve and winnow it again"

Turn-serialization is also present in Makista, but it has a different function than that in TC. For one, the SVCs in Makista place the major verb *virá* "turn" before the verb being modified, while in the TC examples it comes at the end of the sentence. While turn-serialization in Makista is used to indicate the progression of a narrative, in TC it indicates the repetition of an action.

Symmetrical serialization is another common feature in TC. As is the case cross-linguistically, they express distinct parts of a single event.

Transitive-Transitive with same subject and object:

---

[9] "*Tawdu* is a mixture of rice and bran, a by-product of milling used as animal feed" (Maurer 2011: 159)

113.   Tugu Creole (Maurer 2011: 150)

asmis   **lanta**   **garde** neli

immediately pick.up   store rice

"He immediately picks up and stores the rice"

114.   Tugu Creole (Maurer 2011: 162)

djenti kontenti   **kudji fridji** kechap

eople   HAB   cook   fry   *kecapi*

"People cook and fry *kecapi*"

These examples have the object stated after the two verbs. As TC has pro-drop features, this type of SVC can also take the following structure:

115.   Tugu Creole (Maurer 2011: 159)

nobu   nos   **toma   tara**

new   1PL   take   plant

"Then we plant them"

Transitive-Transitive, where object of 1 becomes subject of 2, and have a shared patient:

116.   Tugu Creole (Maurer 2011: 150-1)

djenti kudji **djadi**   **aros**

people cook   become   cooked.rice

"They cook it, so that it becomes cooked rice"

Transitive-Transitive with same subject but different objects:

117.   Tugu Creole (Maurer 2011: 152)

mas   dianti nos   **machika**   aka   neli   **fay**   **say**

more   early   1PL   tread   DEM   rice   make   come.out

sua   gaba

GEN   unhulled.rice

"First, we tread the rice and make its kernels come out"

Transitive-Symmetric SVC:

118.   Tugu Creole (Maurer 2011: 154)

djenti	kustuma	**fay**	fuma	**fay**	empe	pondok	na	sua
People	used.to	make	smoke	make	stand	hut	in	GEN

sawa

wet.rice.field

"One usually makes smoke and builds a hut in one's wet rice fields"

Intransitive-Intransitive with same subject:

119.   Tugu Creole (Maurer 2011: 157)

subi	**santa**	su	riba	ola	neli
go.up	sit	GEN	top	watch	rice

"[We] sit on [the hut] and watch the rice"

120.   Tugu Creole (Maurer 2011: 165)

kadju	lugar	djenti	**fika**	**drumi**	**irgi**
house	place	people	stay	sleep	be.awake

"A house is a place where people sleep and are awake"

Intransitive-Intransitive with different subject:

121.   Tugu Creole (Maurer 2011: 159)

kalu	dja	**say**	rais	**say**	fola	kumpridu
when	PFV	come.out	root	come.out	leaf	long

"When the root and a long leaf have sprouted"

Contrast this example with the following, which use coordination instead of serialization.

122.   Tugu Creole (Maurer 2011: 159-60)

chega	seti	anu	say	fula	dan	fruta
arrive	seven	year	come.out	flower	and	fruit

"After seven years flowers and fruit will sprout"

123.   Tugu Creole (Maurer 2011: 160-1)

djenti	kontenti	fay	orta:	tara	pipinyu,	djagong,	garang
people	HAB	make	garden	plant	cucumber	corn	bean

kumpridu

long

"People cultivate a garden: to plant cucumbers, corn, and long beans"

Intransitive-Transitive with same subject

124.   Tugu Creole (Maurer 2011: 157)

lo	**bi**	tukang	**panya**	dretu,	singku	panya	unga
FUT	come	laborer	take	right	five	take	one

"The collector of dues will come, take one part of five"

One such example uses *anda* "go" as the first element of the SVC, but it is not the same as the go-serialization indicating direction of an action.

125.   Tugu Creole (Maurer 2011: 176)

agora	io	kere	**anda**	na	badjar	**gasta**	bredu	bredu
now	1SG	want	go	to	market	spend	vegetable	vegetable

"Now I want to go to the market to spend [money buying] vegetables"

Although *anda* is the first verb of two in this phrase, there are ways we can distinguish it from the go-form of asymmetrical serialization. In the previous example about shooting a pig, the location was stated at the very end, while in here, it is stated right after the verb. Therefore, the intended meaning in this case, as seen in Maurer's English gloss, is that this person will go to the market and buy vegetables at the market, and it is presented as being one singular event with constituent parts.

The TC corpus has examples of SVCs in which one of the two verbs is directly borrowed from Malay, rather than a calque into a Portuguese-derived word. The following examples show these constructions, with the Malay-derived words in italics.

126.   Tugu Creole (Maurer 2011: 148)

ate	sua	kaske	kote	neli	ake	dja	*mulain*	**kebra**
until	GEN	husk	grain	rice	DEM	PFV	begin	break

"Until its husk has begun to break off"

127.   Tugu Creole (Maurer 2011: 150-1)

djenti	**kudji**	*djadi*	aros
people	cook	become	cooked.rice

"They cook it, so that is becomes cooked rice"

128.   Tugu Creole (Maurer 2011: 154)

kaba	djenti	*barisin*	singku	singku	**arma**
finish	people	line.up	five	five	arrange

"Then one lines it up in rows of five, and arranges it"

129.   Tugu Creole (Maurer 2011: 163)

iste   mes   porku   *malain*   **toma**   muler

DEM   month   pig   start   take   wife

"This month the boars start to mate"

130.   Tugu Creole (Maurer 2011: 163)

agora   bi   nos   karta   ataw   *gotong*   **laba**   na   nos   sua   kadju

now   come   1PL   carry   or   carry   bring   LOC   1PL   GEN   home

"Now let us carry them home"

131.   Tugu Creole (Maurer 2011: 166)

fay   askura   *buat*   **komi**   askura

make   hulled.rice   make   eat   hulled.rice

"They make hulled rice in order to eat hulled rice"

132.   Tugu Creole (Maurer 2011: 166)

djenti   *randam*   **kaba**   **lanta**   **fay**   farinya

people   submerge   finish   raise   make   flour

"They soak it and then take it out of the water in order to make flour"

Most of these are symmetrical serializations, in which the Malay verb that had been chosen likely has no other significance than those of Portuguese origin. The ones with *mulain* and *buat*, however, are examples of asymmetrical serialization and may be used in a similar fashion in Malay.

Symmetrical serialization in TC can have two different functions, that of describing two or more events that happen within the context of a larger event, and that of doing one thing in order to do something else.

Linking serialization describing subsequent events or parts of a whole.

133.   Tugu Creole (Maurer 2011: 150)

asmis   **lanta**   **garde**   neli

immediately   pick.up   store   rice

"He immediately picks up and stores the rice"

134. Tugu Creole (Maurer 2011: 153)

kaba	miste	**djunta**	**konta**	kantu	maradu	teng
finish	must	gather	count	how.much	bundle	COP

"Then one must gather them and count how many bundles there are"

135. Tugu Creole (Maurer 2011: 155)

nos	miste	**abri**	**ola**	**seka**
1PL	must	open	look	dry

"We must open, watch, and dry it"

This example (Sentence 108) has three verbs in the SVC.

136. Tugu Creole (Maurer 2011: 157)

ele	**pare**	**susdi**	tantu
3PL	give.birth	become	much

"The seeds sprout and multiply"

137. Tugu Creole (Maurer 2011: 158)

kaba	**pila**	**mara**	**pedja**
finish	grind	bind	weigh

"After grinding it, we bind and weigh it"

This example (Sentence 110) is a sentence that only contains verbs.

Symmetrical serialization describing doing one action in order to do another.

138. Tugu Creole (Maurer 2011: 150)

poy	**difila**	**fay**	askura
can	PASS.pound	make	hulled.rice

"It can be pounded in order to make hulled rice"

139. Tugu Creole (Maurer 2011: 155)

agora	miste	**kompra**	ikël	puchuk	ataw	bambu	tali
ow	must	buy	palm.leaf.rib	top	or	bamboo	string

fay	korda	**mara**	aka	neli
make	string	bind	DEM	rice

"Now we must buy tips of palm leaf ribs or bamboo for making strings to make strings and bind the rice together"

140.   Tugu Creole (Maurer 2011: 165)

kadera	lugar	djenti	**santa**	**iskrebe**	ka	unga	medja
chair	place	people	sit	write	with	a	table

"A chair is a place where people sit in order to write at a table"

141.   Tugu Creole (Maurer 2011: 182)

iskur,	**para**	fogu	**fika**	lumi
dark	install	fire	become	light

"It's dark, prepare the fire so that it will become light"

There is, however, a purpose conjunction *per* that can also be used instead of the symmetrical SVC.

142.   Tugu Creole (Maurer 2011: 157)

kaba	fay	fuma	fay	empe	pondok	per	subi	santa	su	riba
finish	make	smoke	make	stand	hut	PURP	go.up	sit	GEN	top

ola	neli
watch	rice

"Then we make smoke, erect a hut in order to go up, sit on it and watch the rice"

Another conjunction *ampe* is glossed as "so that".

143.   Tugu Creole (Maurer 2011: 162)

fifay	sua	aros,	nobu	da	kumi	ka	aka	filu	nang

ampe

PASS.make	GEN	rice	new	give	eat	OBJ	DEM	child	NEG

so.that

eli	fomi
3PL	hungry

"When the rice is made fine, they feed the children, so that they are not hungry"

In this example where *ampe* "so that" is used, there is an adjectival predicate *eli fomi* "they are hungry" rather than a verb, so the conjunction is used in order to convey this meaning. This shows parallels with Malay and Indonesian. There are conjunctions that convey the meaning of "in order to" or "for", but it is also possible to have this meaning without such conjunctions. Mintz says that when a verb contains the meaning of "in order to" as a central part of the meaning of the verb, it is likely that such a construction can occur.

144.   Malay (Mintz 1994: 383)

saya   **pergi menolong**   Hamid

1SG   go   help   Hamid

"I went [in order] to help Hamid"

145.   Malay (Mintz 1994: 382)

Hamid **keluar mengambil**   barang-barang yang   tersebar   di
perkarangan

Hamid go.out take   item-item   REL   spread.out   at   yard

"Hamid went out [in order] to get the things strewn about the yard"

While the verb *keluar* "go out" does not have "in order to" contained within its central meaning, it is still grammatical for such a construction to occur.

It is possible to insert a conjunction *untuk* "for; in order to" and the sentence's main meaning does not change.

146.   Malay (Mintz 1994: 382)

Hamid keluar untuk mengambil   barang-barang yang   tersebar   di
Hamid go.out for   take   item-item   REL   spread.out   at
perkarangan
yard

"Hamid went out to get the things strewn about the yard"

Mintz mentions that the example without the conjunction *untuk* is more colloquial.

As in BC, there is the construction with *bi* "come" where rather than indicating the direction of an action, it refers to the first person plural imperative. The word order is similar too, with *bi* [1PL] V, in which *bi* comes first, followed by an optional stating of the first person plural pronoun, then the verb.

147.   Tugu Creole (Maurer 2011, 176)

bi   nos   subi   kareta

come   1PL   go.up carriage

"Let us get on the carriage"

148.   Tugu Creole (Maurer 2011, 179)

bi   kumi   aros   ku   tasab

come   eat   rice   with   jerked.meat

"Let's eat rice with jerked meat"

Bi "come" can also co-occur with *anda* "go", showing that *bi* is not an instance of serialization, as *anda* is the major verb of an asymmetrical SVC.

149.   Tugu Creole (Maurer 2011, 176)

Bapa Siun,   bi   nos   anda   korta   neli

Mister Siun   come   1PL   go   cut   rice

"Mister Siun, let's go and cut rice"

6.2.6 Summary and Discussion

The following is a table that summarizes the asymmetric SVCs that occur in all four creoles:

Table 8: Serial Verb Construction Summary

	Kristang	Makista	BC	TC
come	directional			
finish (kaba)	aspectual, perfective			
finish (ramata)		aspectual, perfective		
get	present			present
give	present		present	present
go	directional			directional
make				present
take	instrumental			
touch	adversative passive			adversative passive
turn		narrative		repetition

Serial Verb Constructions are common not only in creoles but in languages across the world. As they occur cross-linguistically for a determination to be made that the substrates affected the creoles' usage of SVCs it must be demonstrated that there are parallels in their structures. It has been demonstrated that Cantonese makes use of symmetrical serialization (which may just be non-overt coordination), and Makista uses it in a similar way. There is also evidence that symmetric purpose SVCs in TC have a parallel with Malay. While they can be used with a conjunction (*per* in TC, *untuk* in Malay), they may also occur without one.

6.3 Relative Clauses and Cleft Constructions

6.3.1 Introduction

A cleft construction is a sentence with multiple clauses that can be expressed in a single clause, the purpose of which is to put a constituent within the sentence into focus (Hartmann & Veenstra 2013). This shifted focus is different from what a sentence with one clause would have. Both Kristang and Makista make use of cleft constructions.

6.3.2 Relative Clauses and Cleft Constructions in Kristang

Cleft constructions in Kristang are formed using the structure *teng* (copula) + definite NP + relative clause (Baxter 1988).

150. Kristang (Baxter 1988: 210)

teng	bos	sa	familia	ki	yo	ta	lantah,	retu?
COP	2SG	GEN	child	REL	1SG	PROG	carry	right

"It's your child I am carrying, right?"

Similar structures can be found in both Portuguese and Malay

151. Portuguese (native speaker)

é	a sua	criança	que	estou	levando,	não	é?
COP	your	child	REL	am	carrying	NEG	COP

"It's your child I am carrying, right?"

Just like in Kristang, Portuguese has the same order of copula (conjugated form of verb *ser*) +NP + relative clause. The Kristang relativizer *ki* is a direct descendant of the Portuguese *que*, and is used in the same way here.

152. Malay (native speaker)

kopi	yang	anda	sedang	minum,	betul	kah?
coffee	REL	2SG	PROG	drink	right	Q

"It's coffee you're drinking, right?"

Malay has the same word order as the Portuguese and Kristang cleft constructions, with the notable absence of a copula. However, we do see the order of NP + relative clause, with the relativizer *yang* functioning in the same way as *ki* and *que*. Where the lexifier and the substrate have identical or near-identical structures, the resultant creole structure is not likely to be very different, as is seen here.

6.3.3 Relative Clauses and Cleft Constructions in Makista

Makista makes more use of the copula more than Kristang does. One way it can be used is in the cleft construction. The copula *sã*, when used in a cleft construction, is found before the VP of the second clause.

153.   Makista (Santos Ferreira 1967: 24)

mas	êle	nunca	sã	fazê	co	maldádi
but	3SG	never	COP	do	with	wickedness

"But he was never one to deal with wickedness"

Although sentence 4 shows two verbs adjacent to each other, it is not an example of a serial verb construction, where all verbs within the construction have the same tense or aspect. We can see the copula *sã* falling outside the scope of the TAM marker in the following examples:

154.   Makista (Santos Ferreira 1967: 85)

iou	sã	já	nacê	na	Macau
1SG	COP	PFV	born	in	Macau

"It's that I was born in Macau"

155.   Makista (Santos Ferreira 1967: 157)

iou	nunca- sã	querê	matá	vôs
1SG	NEG COP	want	kill	2SG

"I never wanted to kill you"

As in Kristang, no word may come between the TAM marker and its associated verb. In this case, *já* forms a constituent with *nacê*, while *sã* is found outside of this constituent. This shows that it is two clauses, rather than a single clause with a serial verb construction.

This is a cleft construction resembling a form in Sinitic languages, such as Mandarin:

156.   Mandarin (Sun 2006: 190)

ta3	shi4	zuo3tian1	mai3	cai4	(de)
3SG	COP	yesterday	buy	vegetables	PRT

"It was yesterday that he bought vegetables"

Cantonese also takes a similar approach with cleft constructions formed using the copula.

157.   Cantonese (Matthews and Yip 1994: 307)

go1	hai6	duk6	Jing1man4	go3
1SG	COP	study	English	PRT

"It's English I study"

In these Chinese examples, we can also see that there are elements that come between the copula and the verb in the second clause.

The cleft in Cantonese is also used to express contrast:

158.   Cantonese (Matthews and Yip 1994: 307)

hai6	lei5	lam2	cheut3	lai4	g3,	m4hai6	ngo1
COP	2SG	think	out	come	PRT	NEG.COP	1SG

"It was you that thought it, not me"

We can see a contrast in the types of cleft in Cantonese between examples 157 and 158. In 157, the copula follows the subject pronoun and precedes the verb phrase. This is similar to the examples in 153-156. Example 158, however, shows the copula preceding a pronoun, which is not dissimilar to other languages, such as English "It is you that thought it", although without an overt dummy pronoun like the English *it* and without a relativizing pronoun.

The construction of such a clause is similar in both Makista and Chinese. For instance, example 154 could be expressed in a non-cleft form (and have the same meaning, but different focus) as *Iou já nacê na Macau*, "I was born in Macau". The Mandarin example 156 can also be expressed without the copula, and have the same meaning but a different focus. However, such a construction is not present in Kristang.

This type of cleft is found in the older corpora as well:

159.   Makista (Barreiros 1944: 468)

vosso	acunga	diabinha	divéra	sã	querê	pa	iou
2PL.POSS	that	little.devil	really	COP	like	to	1SG

"That little devil of yours really does like me"

160.   Makista (Pereira 1899: 781)

oze em dia	tudo	nhonhonha	sã	falá	portuguezado
nowadays	all	girls	COP	speak	Portuguese

"Nowadays all the girls do speak Portuguese"

The below examples show that the copula is not part of the VP, as TAM and negative markers are:

161.	Makista (Pereira 1899: 124)

Rê	Salmão	mesmo	nun	sã	já	vivo	na	sua
palácio								
King	Solomon	even	NEG	COP	PFV	live	in	his
palace								

com	um	mil	môlêr?
with	one	thousand	woman

"Didn't King Solomon even live in his palace with a thousand women?"

162.	Makista (Barreiros 1944: 129)

cazamento	são	lôgo	fazê	na	Sé
wedding	COP	FI	do	in	cathedral

"the wedding will be done in the cathedral"

Other relative clauses are found in Makista. The older corpora contain relative clauses introduced by *de* "of" as the relativizer.

163.	Makista (Pereira 1899: 125)

já	cavá	tempo	de	Eva	enganá	Adão
PDV	finish	time	of	Eve	deceive	Adam

"after the time that Eve deceived Adam"

164.	Makista (Pereira 1899: 193)

pôde	vem	antes	de	eu	rezá	terço
can	come	before	of	1SG	pray	rosary

"[you] can come before I pray the rosary"

165.	Makista (Barreiros 1944: 34)

como	já	são	hora	de	vem	tudo	parecero	de	jogo,	eu
how	PFV	COP	hour	of	come	all	partner	of	game	1SG

já	nompôde	escrevê	más	novidade
PFV	NEG.can	write	more	news

"As it is now the time for the game partners to come, I can't write any more news"

166.   Makista (Barreiros 1944: 89)

antes	de	eu	vae	pra	Hongkong
before	of	1SG	go	to	Hong Kong

"Before I go to Hong Kong"

A large amount of relative clauses in Makista use the relativizer *que*, which has a direct link to Portuguese.

167.   Makista (Pereira 1899: 194)

sã	ung-a	comparação	que	eu	já	fazê
COP	one	comparison	REL	1SG	PFV	do

"It is a comparison I have made"

168.   Makista (Pereira 1899: 260)

são	tolicia	di	tudo	este	gente	qui	já	votá	pra
elle!									
COP	foolishness	of	all	this	people	REL	PFV	vote	for
3SG									

"It's the foolishness of all these people who voted for him!"

169.   Makista (Barreiros 1944: 36)

ung-a	dôr	de	cabeça	que	nunca	largá	atê	agora
one	pain	of	head	REL	never	drop	until	now

"a headache that by now hasn't stopped"

170.   Makista (Barreiros 1944: 129)

Tio	recebê	carta	de	Éropa	que	dá	novidade
Uncle	receive	letter	of	Europe	REL	give	news

"Uncle received a letter from Europe bringing news"

Another one uses the preposition *pa* "for" and relative pronoun *quim*.

171.   Makista (Santos Ferreira 1973: 152)

fazê	favô	pa	quim	pedí
do	favor	for	who	ask

"[He] does favors for those who ask"

This is a parallel with the Portuguese structure of the same kind of relative clause.

172.	Portuguese

ele	faz	favores	para	quem	pergunta
3SG.M	do.3SG	favors	for	who	ask.3SG

"He does favors for those who ask"

6.3.4 Relative Clauses and Cleft Constructions in Tugu Creole

TC uses a variety of structures for relative clauses, some of which resemble those of Kristang, as well as Portuguese and Malay.

173.	Tugu Creole (Maurer 2011: 171)

siu siu	ki	teng	na	predjentu
sir-sir	REL	COP	in	jail

"Dear sirs, who are in jail"

In this sentence, the relative pronoun refers to the subject of the relative clause, the sirs who are in jail. This has a direct parallel in standard Portuguese:

174.	Portuguese (translation of Sentence 12)

senhores	que	estão	na	cadeia
sirs	REL	COP.3PL	in.the	jail

"Sirs who are in jail"

There are examples in TC where there is no overt relative pronoun, such as the following:

175.	Tugu Creole (Maurer 2011: 174)

teng	djenti	fala	chuma	sabe	iskrebe
COP	people	say	only	know	write

"There are people who say that they can only write"

This example is similar to the previous, in that the gapped pronoun in the relative clause is the subject. However, there is no overt marking of either a relative pronoun or a subject of the relative clause.

TC also uses similar strategies to Portuguese in which the relative pronoun refers to the object in a relative clause.

176.	Tugu Creole (Maurer 2011: 178)

aka	redjang	ki	io	dja	premetesang
DEM	story	REL	1SG	PFV	promise

"The story which I have promised"

177. Portuguese (translation of Sentence 15)

a história que (eu) prometi

the story REL (1SG) promised

"The story which I have promised"

This is another example of the structure having a parallel with Portuguese, in which there is an overt relative pronoun referring to an object in the relative clause which is gapped. There are also instances of non-overt relative pronouns in TC which refer to the object in the relative clause.

178. Tugu Creole (Maurer 2011: 181)

korpu sua nomi nos lumia aka

body GEN name 1PL name DEM

"These are all the body parts we have named"

179. Tugu Creole (Maurer 2011: 173)

teng mes chempra fil filu malay prende

COP self mix children Malay learn

"There are also Malay children who learn [there]"

The following example contains a relative clause in which the subject is the patient of a passive verb.

180. Tugu Creole (Maurer 2011: 176)

iste sawdu teng djenti dipindra

DEM Saturday COP person PASS.hang

"Next Saturday, somebody will be hung"

This also does not have a relative pronoun.

6.3.5 Discussion

Of the relative clause constructions in the four creoles, the ones that differ the most from the rest are the cleft constructions of Makista. As shown in previous examples, Portuguese and Malay cleft constructions are of similar form, with a relative clause fronted by the relative pronoun *que* (Portuguese) and *yang* (Malay). Makista exhibits a cleft construction which makes use of the copula *sã*, a feature not found in the other creoles. We can see that it closely resembles the structure of Sinitic languages, including Cantonese, where the copula introduces a cleft clause. Of the four languages Makista has had the most contact with a

Sinitic language, and the presence of copula clefts in Makista alone out of the four is likely due to this contact.

6.4 Short and long pronouns

6.4.1 Introduction

In much of the literature regarding Kristang (Baxter 1988, Pillai 2011, Singho et al 2017), pronouns are written according to their canonical forms. However, a casual observer of spoken Kristang will notice that there is variation in the pronunciations of them. The pronouns in question are the second personal singular with the variants *bos~bo*, the third person singular with the variants *eli~el~e*, and the first person plural with the variants *nus~nu*. This section investigates potential parallels in Malay, the main substrate of Kristang, and also what triggers the use of each pronoun variant. While I was not able to find a definitive reason, the short/long distinction appears to be a result of focus, as well as definiteness. I use examples from Pillai's corpus and supplement the data with examples from my elicitation of Kristang speakers.

This section is less comparative than the others in this chapter, in that the only creole discussed in detail is Kristang. The primary reason is that short and long pronouns are not attested in the other creoles - as the other corpora used in this thesis are written only it is impossible to know if there was variation in pronunciation that was not reflected in writing. There is also no chapter dedicated to this phenomenon in APICS, and therefore it is not considered for this section.

The feature of short and long pronouns is considered as it appears to be unique among the four creoles in the present study. As Kristang is spoken in the city where these creoles originated, this therefore appears to have developed after the other languages emerged. I aim to demonstrate that this feature emerged as a result of the linguistic ecology of Malacca, by comparing it to Malay pronouns.

6.4.1 Personal Pronouns in Malay

The Malay language also has long and short forms for some of its pronouns, expressed as pro- or enclitics. Also like Kristang, there are preverbal elements: aspect and negation markers and auxiliary verbs. However, the short pronouns in Malay must occur next to the verb stem, in between the verb and any markers. Mintz (1994: 92-94) refers to the short pronouns in Malay as "alternative pronouns". The only pronouns which may be expressed as the alternative form are the first person singular *aku* to *-ku*, the second person singular *kamu*

to *-mu* and the third person *dia* (singular)/*mereka* (plural) to *-nya*. These may be attached to a verb phrase to indicate the object, or to a noun phrase to indicate possession:

181.　　　Malay (Mintz 1994: 93)

Zainal cuba lari, tetapi penjahat memukul=nya

Zainal try run but criminal hit=3SG

"Zainal tried to run, but the criminal hit him"

182.　　　Malay (Mintz 1994: 93)

saya pinjam buku=nya

1SG borrow book=3SG

"I borrowed her book"

Mintz suggests that these alternative pronouns are remnants from a set of non-subject pronouns, as the third person pronoun does not match its alternative form as well as the others do. Modern Malay does not express case in the pronoun forms, but the alternative forms are only used as an object.

　　　Indonesian also has long forms and bound short forms of certain pronouns, all of which are singular. The following table, adapted from Sneddon (1996: 170), illustrates this:

Table 9: Indonesian Long and Short Form Pronouns

	long form	**prefixed short form**	**suffixed short form**
first person	aku	ku-	-ku
second person	engkau, kamu	kau-	-mu
third person	dia, ia	-	-nya

The short forms can indicate the subject or the object of a verb.

183.　　　Indonesian (Sneddon 1996: 170)

Dia mengambil kue itu lalu memakan=nya

3SG take take that then eat=3SG

"He took the cake and ate it."

184.　　　Indonesian (Sneddon 1996: 171)

Narti di=tunggun=nya

Narti PASS=wait=3SG

"He is waiting for Narti"

They can also follow nouns to convey possession:

185.　　　Indonesian (Sneddon 1996: 171)

    a. rumah=ku

       house=1SG

       "my house"

    b. mobil=mu

       car=2SG

       "your car"

    c. buku=nya

       book=3SG

       "her book"

Bound pronouns can also follow some prepositions, but not all. Notably, this is possible with *kepada* "to". Sneddon also notes that "[i]n all situations in which the bound forms can occur, corresponding free forms are also possible."

6.4.2 Short and Long Pronouns in Kristang

It is not immediately obvious what marks the difference between the two types of Kristang pronouns, and non-native speakers have no problem being understood using either variety. Little seems to be known about this phenomenon, both in the literature and amongst the speakers themselves (one speaker said about the short varieties that it was "bad Kristang").

Baxter (1988: 53) mentions the variation between /bos/~/bo/ and /nus/~/nu/ in Kristang, although there is no discussion that explains the reason for this variation. He also writes about the third person singular pronoun *eli* with variation between /e/~/el/~/eli/, and attributes the shorter forms of the pronoun to phonology, arguing that the shortened form may occur before a stressed syllable. However, it appears that information structure plays a role in this pronoun variation. The Kristang examples in this section are from recorded data, and I have written the pronouns according to how they have been spoken, rather than the canonical form. I have put the pronouns in slashes to show this is how it was pronounced in the recordings. The analysis of this phenomenon is quantitative, and I have limited this to the 1PL and 2SG pronouns. This is because there are only two varieties, rather than the three found in 3SG.

Upon a closer look at the data, patterns begin to emerge between the varieties of the pronouns. Pillai's recordings show people using both varieties, sometimes in the same utterance.

186.   Kristang (Pillai 2011)

/nu/   nang ubih   keng falah /nus/ papiah Kristang

1PL   NEG   listen who   say   1PL   speak Kristang

"We don't listen to those who say we speak Kristang"

Although both forms are frequently found in subject position, the long form is preferred when stated as an object. The following table shows totals of the pronouns with regards to their grammatical role.

Table 10: Tokens of pronouns in Pillai (2011); Grammatical Roles

	subject	subject with negative	object
/bos/	4	0	6
/bo/	30	3	0
/nus/	53	2	6
/nu/	51	5	3

The short form is strongly preferred before a negative particle. The only two examples of the long form before one is when there was a pause in the speech:

187.   Kristang (Pillai 2011)

mas   kifoi   /nus/…nang   kabalu…   ku   onsong   sa   disimentu

but   why   1PL   not   ashamed   with   own   GEN   ancestry

"But why should we not be ashamed of our own ancestry?"

188.   Kristang (Pillai 2011)

mas   /nus/…nungkafalah bai   na   Portugal   komprah   isi   libru

but   1PL   NEG   talk   go   to   Portugal   buy   this   book

"But we don't say we go to Portugal to buy this book"

This suggests that prosody may also play a role in the choice between a long or a short form. I tallied the totals of the pronouns based on their positions in an utterance:

Table 11: Tokens of pronouns in Pillai (2011); Position in sentence

	at the end of utterance	before a pause	elsewhere
/bos/	4	1	5
/bo/	1	0	29
/nus/	2	11	46
/nu/	0	3	51

While there is still insufficient data, this suggests a preference of the long form for *bos* at the end of an utterance, although it is less clear for *nus*.

Although there are some exceptions, short pronouns tend to be found before verb phrases. This can also explain why the long form is preferred in object position, since the word order is SVO. Of the 84 examples of short pronouns found in Pillai's recordings, only four are not before a verb, pre-verbal marker, or predicate:

189.   Kristang (Pillai 2011)

/bo/   pulak   kereh   interview   ku   jenti

2SG   again (Mal.)   want   interview   with   people

"You'll want to interview people again"

190.   Kristang (Pillai 2011)

ki   relijang   /bo/?

what   religion   2SG

"What religion are you?"

191.   Kristang (Pillai 2011)

/nu/   krensa ngka   celebrate

1PL   child   NEG   celebrate

"We children don't celebrate"

192.   Kristang (Pillai 2011)

komu /nu/   jenti   na   Malaka

like   1PL   people in   Malacca

"Just like us people in Malacca"

Among these exceptions, one is preceding an adverb, one is at the end of an utterance, and two are attributive pronouns.

It is also important to note that the verb phrase in Kristang may have aspect particles, negative particles, and auxiliary verbs, all that precede the main verb. The pronoun, long or short, always comes before particles if they are there:

193.   Kristang (Pillai 2011)

yo   podi   olah,   /bo/   mpodi olah?

1SG   can   see   2SG   cannot see

"I can see, you can't see?"

194.   Kristang (Pillai 2011)

agora	ja	singku sentu	anu	ki	/nu/	ja	fikah	na	isti	tera
now	PFV	five hundred	year	that	1PL	PFV	stay	in	this	

land

"Now it's already five hundred years that we have stayed in this country"

Another factor in the choice between long and short forms is definiteness. One recording by Pillai shows a woman explaining how to make a fish pickle. In the recording she uses the second person pronoun, and exclusively uses the short form.

195.   Kristang (Pillai 2011)

achar pesi	/bo/	kereh	jinjibri, kunyit,	alu.	/bo/	blend	fazeh

ungua

fish pickle	2SG	want	ginger turmeric	garlic	2SG	blend	make	one

"For fish pickle, you want ginger, turmeric, and garlic. You blend and make it on"

This is likely due to using the indefinite pronoun, as she is giving instructions. Her choice of the short form is a reflection of the pronoun not referring to a specific person, but rather anyone who is listening to her explanation, as she is the one making the fish pickle.

As seen before, the long form is preferred at the end of an utterance, as well as when the pronoun is an object (which also tends to occur at the end of an utterance). However, the indefinite *you*, regardless of its position in the sentence and its grammatical role, will take the weak form.

There was another example, however, where an indefinite *you* was used, but was pronounced /bos/.

196.   Kristang (Pillai 2011)

yo	mpodi	parah	si	ramidi.	/bos/	parah	duenti	logu	beng	ku

/bos/

1SG	cannot	stop	this	work	2SG	stop	illness	FUT	come	ACC

2SG

"I cannot stop this work. If you stop, illness will come to you"

We can see that he was talking about himself in this recording, by looking at the previous sentence in which he is referring to himself and his own work.

This particular recording has no other examples of the speaker saying *bos*, so it is not possible to determine if the distinction between /bos/ and /bo/ exists in his speech. However,

he does use the pronoun *nus* "we" in both the long and short forms within the same recording.

197.　　　Kristang (Pillai 2011)

/nus/	tokah	bai	buskah	kumih	is	sorti	lah
1PL	touch	go	find	eat	this	type	PRT

"We have to go find a living like this"

This example uses *nus* in the subject, and is in the long form.

198.　　　Kristang (Pillai 2011)

eli	logu	da	ku	/nu/	mpoku	doi	dos	mil,	kantu
3SG	PFV	give	to	1PL	a.little	money	two	thousand	how.much

anu	/nu/	kereh	bibeh
year	1PL	want	live

"They [the government] will give us a little bit of money, about two thousand [ringgit], how many years do we want to survive with it?"

199.　　　Kristang (Pillai 2011)

oltu	chumah	/nu/	falah,	Kristang	falah,	oltu	durian	/nu/
3PL	like	1PL	say	Kristang	say	3PL	durian	1PL

pipinyu
cucumber

"They are like what we say, what the Kristang say, they are durians and we are cucumbers"

In this case, the pronoun is stated after the object marker *ku*, but the short form is used. This could be because it is the indirect object, while the direct object *mpoku doi* "a little money" is stated directly after it.

　　　The second example shows a sentence of comparison. In this instance the focus is on the fruits that they are comparing each other to, the durian and the cucumber. The first part of the sentence also does not have *nus* as the main focus, but rather *falah*, or what they say. This could be what triggers the short form in this case.

The following examples are from the same speaker, and the same recording session:

200.　　　Kristang (Laub fieldwork, 2017)

kantu	/bo/	fazeh	di tona	yo	lo	dali	ku	/bos/
If	2SG	do	again	1SG	FUT	hit	to	2SG

"If you do that again I will beat you"

This sentence was elicited from a translation from the English. This sentence was chosen because of the usage of 2SG in both subject and object form. I also chose the verb *beat* because of its transitivity, and expecting that the accusative marker *ku* would be used with this pronoun, as it is a human object. Although this was elicited in isolation from any wider context, it is clear from the meaning that the *you* refers to a definite person. Before the elicitation I had expected the first 2SG, the subject would be in short form, while the second one, the object and the final word in the utterance, to be the long form.

As seen above definiteness appears to have an effect on the choice of pronoun form. In the same recording session I invited the consultant to speak freely about what one should do when they are a tourist in Malacca. I chose this as a topic because I anticipated usage of the general *you*, as was found in one of Pillai's recordings. I also anticipated the short form to be present, especially in sentences where a specific *you* would be stated in the long form. The following sentence is an example of what the consultant provided:

201.   Kristang (Laub fieldwork, 2017)

logu	nganah	ku	/bo/	rekah	ku	/bo/
FUT	deceive	with	2SG	deceive	to	2SG

"They will deceive you"

Even at the end of an utterance and as an object, the short form prevails, due to its non-specific use.

202.   Kristang (Laub fieldwork, 2017)

tudu	/nus/	nakih	isti	bairu	fikah	a	Portugis	di	Melaka
all	1PL	here	this	village	stay	to	Portuguese	of	Malacca

"All of us here in this village became a Portuguese of Malacca"

203.   Kristang (Laub fieldwork, 2017)

/nus/	par	akih	isti	bairu	di	Melaka
1PL	to	here	this	village	of	Malacca

"We here in this village in Malacca"

In this example the long form of the first person plural is present before a preposition.

When there is an explicit contrast, the focused pronoun tends to take the long form, while the other takes the short form.

204.   Kristang (Laub fieldwork, 2017)

agora	ngkoza	eli	teng	/bo/	nteng	eli	lo	pidi
now	something	3SG	have	2SG	NEG.have	3SG	FUT	ask

"Now he has something, you don't, he will ask [you]"

205.   Kristang (Laub fieldwork, 2017)

/bos/   pidi?

2SG   ask

"*You* ask?" (ie, not him)

206.   Kristang (Laub fieldwork, 2017)

/bo/	ke	dah	ku	yo	prezanti,	yo	lo	ku	/bos/
2SG	want	give	to	1SG	present	1SG	FUT	to	2SG

prezanti	ku	kuelu	bedri
present	to	rabbit	green

"You want to give me a present, I will give you a present for the green rabbit"

In an elicitation exercise, I held a picture of a red pepper, while another participant held one of a green pepper. I asked the participant to simply describe what he saw:

207.   Kristang (Laub fieldwork, 2017)

/bos/	ta	pegah	paparika	brumilu,	Juang	ta	pegah	pintura
2SG	PROG	hold	pepper	red	John	PROG	hold	picture

paparika	pretu
pepper	black

"You are holding a red pepper, John is holding a picture of a black [*sic*] pepper"

When there is no contrast or contradiction, but just a description, the long form was preferred. When there is contradiction about the object, rather than the subject pronoun, the long form was used:

208.   Kristang (Laub fieldwork, 2017)

yo	teng	pintura di	pasturinyu?
1SG	have	picture of	bird

"Do I have a picture of a bird?"

/bos/	ta	pegah	pintura di	gatu
2SG	PROG	hold	picture of	cat

"You are holding a picture of a cat"

209.   Kristang (Laub fieldwork, 2017)

/bos/   ta   pegah pintura di   kezu

2SG   PROG hold   picture of   cheese

"You are holding a picture of cheese" (After being asked if I held a picture of a

banana)

Although the focus in these answers are subject of the pictures being held, the long form of the pronoun was used. This could be a result of the contrast being with what the picture is depicting, not a contrast between two pronouns, as in the previous examples.

The long and short contrast between pronouns in Kristang do not appear to be strict rules, but there are some strong tendencies. When a pronoun is in focus, the long form is preferred. This is especially true when followed by the object marker *ku*. One exception to this is when a pronoun is non-specific. There is only a small amount of data, and more research is required, but every single instance of a non-specific pronoun uses the short form, even when followed by *ku*.

We can see some parallels between the long/short forms in Kristang and the free/bound forms found in Malay and Indonesian. In both languages only a subset of personal pronouns can take alternative forms. In Malay and Indonesian only singular pronouns can take the bound form, and in Kristang it is limited to the second- and third-person singular and first-person plural. In addition, they can occur before or after the verb, and after some prepositions. We can also find the long/free and short/bound forms in the same position in a sentence.

The differences between them is that Kristang does not use the short form post-nominally as an indicator of possession (a closer look at the genitive can be found in Section 6.6). Within a verb phrase, Malay and Indonesian bound pronouns must occur adjacent to the verb, while in Kristang there are often intervening elements, such as auxiliaries and TAM markers.

6.4.3 Personal Pronouns in the Other Creoles

There is no evidence of such differences between pronouns in Makista, BC, or TC, suggesting that this is an innovation that had occurred after the split. It is plausible that the rule present in all four creoles, that there can be no intervening element between the TAM particle and the head verb, had been long established before the pronoun variation came about in Kristang.

Usage of indefinite pronouns can be found in Makista, in the YouTube video entitled *Ficá Vivo na Macau* by Dóci Papiaçám di Macau[10]. It is a humorous explanation of what to do and what not to do for tourists visiting Macau. Makista's second person pronoun, "vôs", is used in a similar way, although the final /s/ sound is never dropped in the way that it is in Kristang.

210.　　　Makista (*Dóci Papiaçám di Macau*)

entre	na	Macau di	Tera China,	vôs	tem	que	usá	ropa
enter	in	Macau of	China	2SG	have	that	wear	clothing
forte
strong

"When you enter Macau from China, you have to wear strong clothing"

Cantonese is a pro-drop language that can express the subject clearly or not at all, (Luke et al 2001), but does not use pronoun variation to express prominence in the way that Kristang appears to. Cantonese does have variation in the pronunciation of its pronouns: for instance, the second-person singular can be pronounced as *néih* or *léih* (Yip and Matthews 1999), but this is due to sound change and is not a long-short distinction like can be found in Malay. Portuguese does have pronoun variation but it is due to the grammatical role they play within the sentence. Therefore it could be for this reason that Makista did not develop a short/long pronoun distinction in the way that Kristang did.

6.4.4 Discussion

It is not possible to come to a conclusion on short and long pronouns based on the current data available on BC and TC. However, none of the data contains pronouns spelled in ways that would suggest they were pronounced in a shorter form like they are in Kristang. The only alternation in TC pronouns are where the first-person singular is pronounced as *mi* (derived from Portuguese object form) or as *yo* (derived from Portuguese subject form). It is unclear why this is the case, as the pronouns in the four creoles are derived from subject forms.

The two likely causes of Kristang short/long pronoun variation are internal development, or influence from a contacting language. The use of alternative pronouns in Malay resembles that of Kristang, but not entirely. Malay bound pronouns adhere to a strict word order that free ones do not, while Kristang short pronouns have been found in environments where the long ones have as well. This is a feature that requires more attention in the future, especially as the Malacca Eurasian community shifts to English. If Kristang

---

[10] https://www.youtube.com/watch?v=93YUQCwejNY

inherited this feature from Malay, there was incomplete transmission where the rules for word order did not pass on to Kristang.

6.5 Verbal ellipsis after a TAM particle

6.5.1 Introduction

A distinguishing feature of Kristang, Makista, BC and TC is the usage of tense, aspect and mood (TAM) particles in pre-verbal position. In contrast to standard Portuguese, which expresses TAM through auxiliary verbs and verbal inflection, the four creoles express it through pre-verbal elements while the structure of the verb itself does not change. The three particles that will be discussed in this section are the perfective aspect particle (*ja* in Kristang, *já* in Makista), the future-irrealis particle (*lo* or *logu* in Kristang, *lôgo* in Makista), and the non-punctual aspect particle (*ta* in both languages).

This section will look at potential lexifier influence. Pre-verbally, all three of the main TAM particles act very similarly. They come right before the verb, and no intervening morpheme may come between particle and verb. The perfective particle may also occur before a predicate adjective (as in, *It is already late*). Two of the three may also occur in isolation, usually as the answer to a question.

6.5.2 TAM Markers' Origins

The lexical origins of these markers come from Portuguese. The perfective aspect particle is derived from the Portuguese adverb *já*, meaning "already". Although an adverb in Portuguese and not an aspect particle as it is in Kristang and Makista, the pronoun *já* comes before the verb in a word order similar to the creoles:

211.   Portuguese (Ganho & McGovern 2003: 178)

já   comi

already   eat.1SG.PST

"I already ate/I've already eaten"

212.   Kristang (Baxter 1988: 164)

eli   ja   sai   di   kaza

3SG   PFV   leave   of   house

"He left the house"

213.   Makista (Santos Ferreira 1967: 95)

sã   assi   qui   ilôtro   três   já   vai

COP   then   that   3PL   three PFV   go

"It was in this way that they went"

The future-irrealis particle is derived from the Portuguese adverb *logo*, meaning "soon", or "later". In contrast with *já*, it can come after the verb:

214. Portuguese (Ackerlind & Jones-Kellogg 2011, 180)

dissemos	que	sairíamos	logo
say.1PL.PAST	that	leave.1PL.COND	soon

"We said that we would leave soon"

215. Kristang (Baxter 1988:126)

amiang	otu	dia,	eli	logu	bai	mar
tomorrow	other	day,	3SG	FUT	go	sea

"The day after tomorrow he will go fishing"

216. Makista (Santos Ferreira 1967, 19)

si	sã	iou	tamêm	lôgo	pegá	dále!
tf	COP	1SG	also	FUT	give	attack

"If it were me, I would also attack him!"

The non-punctual aspect particle is derived from a shortened form of Portuguese *está*, the third-person singular present form of the verb *estar* "to be". In both European and Brazilian Portuguese, the progressive is marked with the form of *estar* in the appropriate conjugation and tense, followed by the main verb. European Portuguese marks the progressive with the preposition *a* + infinitive, while Brazilian Portuguese uses a gerund form of the verb:

217. European Portuguese (Ackerlind and Jones-Kellogg 2011: 212)

estamos	a	estudar
be.1PL.PRS	to	study.INF

"We are studying"

218. Brazilian Portuguese (Ackerlind and Jones-Kellogg 2011: 212)

estamos	estudando
Be.1PL.PRS	study.GER

"we are studying"

219. Kristang (Baxter 1988: 129)

eli	ta	fikah	godru
3SG	PROG	become	fat

"He is becoming fat"

220.   Makista (Santos Ferreira 1967: 96)

nôs   ta   vai   pa   acunga   rua

1PL   PROG go   to   this   street

"We're going to this street"

What is important to note is that the perfective and future-irrealis particles are derived from Portuguese adverbs, while the non-punctual particle is derived from a Portuguese auxiliary verb.

In Bazaar Malay, a yes/no question using a modal or a TAM marker may be answered in the affirmative using the marker without its head noun (Aye 2006). The following is a common greeting and response:

221.   Bazaar Malay:

sudah   makan?

PFV   eat

"Have you eaten?"

sudah

PFV

"Yes, I have"

6.5.3 Verbal Ellipsis in Kristang

Spoken Kristang employs the exact same greeting and response as in Sentence 11, including the word order and TAM marker response without the main verb:

222.   Kristang (personal knowledge)

ja   kumih?

PFV   eat

"Have you eaten?"

ja

PFV

"Yes, I have"

Similar to the perfective particle, the future-irrealis particle may also be used on its own as an answer:

223.   Kristang (Laub fieldwork, 2017)

bos   lo/logu   bebeh?

2SG   FUT   drink

"Will you drink?

Logu/*Lo

FUT

"Yes, I will"

In this instance, while the particle may be stated on its own without the main verb, it must be
stated in its full form *logu* rather than the short form *lo*.

Modals may also be used in the same way:

224.   Kristang (Laub fieldwork, 2017)

bos   misti   kumih?

2SG   must   eat

"Do you have to eat?

misti

must

"Yes, I do."

The progressive aspect particle, however, may not occur on its own in an answer:

225.   Kristang:

bos   ta   bebeh?

2SG   PROG drink

"Are you drinking?"

ta   bebeh/*ta

PROG drink   PROG

"I am (drinking)"

As mentioned earlier, the lexical origins of *ja* and *logu* are from Portuguese adverbs,
while *ta* is from a verb, and could be a contribution to their difference in use. Baxter (1988)
also points out some of the usages of these particles, and how there is still some Portuguese
influence on them.

226.   Kristang (Baxter 1988: 121)

kora   yo   chegah eli   ja   ta   kumih
when   1SG   arrive 3SG   PFV   PROG eat

"When I arrive he is/will be already eating"

227.   Malay (Baxter 1988: 121)

bila   gua   datang dia   sudah *sedang   makan
When   1SG   arrive 3SG   PFV   PROG   eat

When used on their own, the Kristang *ja* and *ta* generally are used in the same way as Malay *sudah* and *sedang*, respectively. This example shows a point of departure from the Malay parallel, where the coinciding of these two particles is grammatical in Kristang, but not in Malay. In Portuguese, as *já* is an adverb and *estar* is a verb, they may coincide:

228.   Portuguese (native speaker)

ele   já   está   a   comer.
3SG   already   be.3SG.PRS   to   eat

"He is already eating"

Baxter also points out that *ja* and *logu* can be used outside of a predicate in an adverbial manner:

229.   Kristang: (Baxter 1988: 123)

kora   yo   ja   chegah,   eli   ja   bai   ja
when   1SG   PFV   arrive   3SG   PFV   go   PFV

"When I arrived he had already gone"

230.   Kristang: (Baxter 1988: 137)

eli   lo   duenti, lo   fika   logu
3SG   FUT   ill   FUT   become   FUT

"He'll be ill, he'll become ill"

Additionally, the adverbial usage of the future-irrealis particle parallels the usage in a response without the head verb, in that the contracted form *lo* is ungrammatical; *logo* must be used in this case. This is not dissimilar to the case of short and long pronouns, where a long pronoun is preferred at the end of an utterance.

When it is not followed by the verb, but it is not the end of a sentence, the short form can be used:

149

231.   Kristang (Laub fieldwork, 2017)

o	ke	dah	ku	yo	prezanti	epal,	yo	lo	ku	bos
2SG	want	give	to	1SG	present	apple	1SG	FUT	to	2SG

sa	kuelu	brumilu
GEN	rabbit	red

"You want to give me a present of an apple, I will [give it] to your red rabbit"

This instance has ellipsis of the verb *dah* in the second clause, with its antecedent in the first clause.

The TAM particle on its own can answer a question, or it can be part of an answer, as in the following example:

232.   Kristang (Pillai 2011)

bos	ja	bai	skola?
2SG	PFV	go	school

"You went to school?"

seng,	ja.
Yes	PFV

"Yes, I did"

Verbal ellipsis can also happen in some cases after an auxiliary verb or the first part of a serial verb construction. Such an example can be found in Kristang:

233.   Kristang (Marbeck 2004: 53)

bos	olah,	istri	tigri	ja	tokah	na	trongku
2SG	see	this	tiger	PFV	touch	in	trap

"You see, this tiger was caught in a trap"

Although implied that the tiger had been caught in a trap the actual verb *catch* is not present in this example. In the same story about the tiger by Marbeck, we do have both verbs stated:

234.   Kristang (Marbeck 2004: 53)

istri	tigri	ja	tokah	pegah	na	lastru
this	tiger	PFV	touch	catch	in	trap

"This tiger here was caught in a trap"

As mentioned earlier, the verb *tokah* can be interpreted as an adversative passive, so within the context it can be implied that the tiger had been caught.

6.5.4 Verbal Ellipsis in Makista

Makista also exhibits responses to questions in the form of a TAM particle without the head verb:

235.　　　Makista (Santos Ferreira 1973: 155)

"jugá　dado　co　　iou?"

blay　dice　with　1SG

"'Will you play dice with me?'"

"lôgo,　lôgo",　Dom Vico　　consolá　　êle

FUT　FUT　Dom Vico　　console　　3SG

"'I will, I will,' Dom Vico consoled him"

236.　　　Makista (Santos Ferreira 1967: 92)

azinha　　　vai, Lita. Lembrá　　　　　virá-virá　　vêm…

quicklygo　Lita　Remember　　　sometime　　come

"Go quickly, Lita. Remember to come back sometime"

lôgo,　Marta

FUT　Marta

"I will, Marta"

In these instances, the characters in the first lines do not use the future-irrealis marker *lôgo*, although from the context it is implied that they are speaking about the future.

237.　　　Makista (Santos Ferreira 1967: 23)

Chácha　　　chomá iou?

Grandma　　call　me

"Are you calling me, Grandma?"

sã,　　filo

COP　son

"Yes, son"

238.   Makista (Santos Ferreira 1967: 90)

quelê-môdo	acunga	garidóna	já	fugí	de	casa?
how	that	flirt	PFV	run.away	of	house

"How did that flirt run away from home?"

sã	divéra.
COP	true

"It's true"

239.   Makista (*Dóci Papiaçám di Macau*)

"Cuza Dotôr"	já	falá	co	dotôr	falso	qui	já	fugí	pa
"What's Up, Doc"	PFV	talk	with	doctor	fake	who	PFV	escape	to

Philipino
Philippines

"(The comedy) What's Up, Doc?' told about a fake doctor that ran away to the Philippines"

Sã
COP
"Yes, it did"

Cantonese also uses the same structure, using the copula *hai6* as a response to show agreement (Yip & Matthews 1999: 33)

240.   Cantonese (Yip & Matthews 1999: 33)

hou2	gwaie3	wo3
very	expensive	PRT

"It's very expensive"

hai6	a3
COP	PRT

"Yes, it is"

The Pereira corpus has examples of the progressive marker *ta* used before non-verbal predicates, which was not found in the Kristang corpora. The below examples have adjectival predicates:

241.   Makista (Pereira 1899: 58)

Nhum Lorenço   jâ   tâ   vêlo

Mister Lorenço   PFV   PROG old

"Mr Lorenço is already old"

242.   Makista (Pereira 1899: 260)

quando   cofre tá   bem   inchido   lôgo   olá   então ôlo

when   coffer PROG well   stuffed   FI   see   then   eye

grande-grande

big-big

"When the coffers are full, we will see the big eyes"

243.   Makista (Pereira 1899: 260)

agora   elle   tá   mutu   contente   naquelle   cadera

now   3SG   PROG very   happy   in.that   chair

"Now he is very happy in that chair"

Additionally, there was one with a prepositional predicate:

244.   Makista (Pereira 1899: 780)

ven   falá   qui   ta   na   porta

come   speak who   PROG LOC   door

"Come speak to who is at the door"

This is similar to the Portuguese usage of *estar* "to be", and it appears that the distribution of the progressive marker in Makista is less restricted than that of Kristang.

6.5.5 Verbal Ellipsis in Tugu and Batavia Creoles

TC has an example of a TAM particle without the verb following.

245.	Tugu Creole (Maurer 2011: 193)

bos	nungku	limpa	korpu

2SG	NEG	wash	body

"You didn't bathe"

dja,	na	rio

PFV	in	river

"Yes I did, in the rive"

This is similar to other examples above, in which an answer to a question uses the TAM particle, in this case the perfective *dja*, without the verb explicitly stated.

Another example of the same verb can be found in BC.

246.	Batavia Creole (Maurer 2011: 129)

ile	bringka	kung	katana	seng	toka

3SG	play	with	short.sabre	without	touch

"He skirmished without getting wounded"

As in the case above, this example using *toka* as an adversative passive does not have the main verb; it is rather inferred from the wider context of the sentence. See 233-234 for similar examples in Kristang.

The case of *ta* sticks out because it comes from a Portuguese verb, and it cannot be stated on its own as a response to a question. Portuguese allows for answers to questions to simply be the verb.

247.	Portuguese (Ackerlind and Jones-Kellogg 2011)

você	leu	o	livro?

2SG	read.2SG.PAST	the	book

"Did you read the book?

ei,	sim

read.1SG.PAST	yes

"Yes, I did"

6.5.6 Discussion

Sadock and Zwicky (1985) point out three ways that languages can answer yes/no questions. The first is yes-no systems, which have particles *yes* and *no* for positive and negative

answers, respectively. Agree-disagree systems have a positive particle and a negative particle for responses that agree or disagree with the polarity of the question, respectively. Echo systems will repeat the verb or some part of the VP in the answer. As we have seen in the Kristang examples in this section, it appears to be an echo system. In the case of Kristang, questions are answered by repeating the TAM particle (with the exception of *ta*). Portuguese tends to use echo answers (despite it also having yes/no particles, *sim/não*). Santos (2002: 61) identifies four main simple answers to questions in Portuguese:

248.   Portuguese (Santos 2002: 61-2)

Q:	Ele	já	encontrou	a	chave?
	3SG.MASC	already	found.3SG	the	key

"Has he already found the key?"

A:
 a. Sim
"yes"
b. Encontrou
"found"
c. É/Foi
"is"/"was"
d. Já
"already'

The four main types shown here are (a) yes answer, (b) verbal answer, (c) *ser* (be) answer and (d) adverbial answer. The adverbial answer on the surface resembles the ability to answer *ja* on its own in Kristang, but as it is an aspect particle it is not produced with the same process.

        Lumsden (1999: 149-50) notes that functional words in creoles often come from lexical sources. This is certainly true in the case of TAM markers in the four creoles (as well as with the post-nominal genitive particle, see Section 6.6). Lumsden argues that such morphemes do not begin their life within a creole as a grammatical function, but are transferred into the creole as a lexical item. This would make the initial transfer not unlike the vast majority of words in the creoles, which took on very similar phonological shapes as their lexical origins, and most often kept the same lexical meaning as in the lexifier. Where they

get their grammatical functions, then, is within the creole language itself, through the process of reanalysis.

What appears to have happened with the TAM particles in these creoles is a typical case of grammaticalization. We can see some words that have not gone through complete grammaticalization, but might be on their way. The verb [Kr. *kereh*/Mak. *querê*] is such an example. As a verb in isolation, it lexically means *to want*, as is also the case in Portuguese *querer*. Makista has examples of *querê* being used to signify the future, while Kristang has other uses for it. Kristang also has the tendency to have optional shortening of pre-verbal syntactic elements, such as the pronouns (see Section 6.4), as well as certain particles such as *logu* → *lo*. *Kereh* also may shorten to *ke* when in a pre-verbal position.

The implications of this feature point to a substratist/superstratist explanation. Typologically, Kristang and Makista answer questions in a similar way as their substrates, as well as Portuguese. While certainly not a marked phenomenon, this parallel of features is just one of many that indicate the substrate's effect on the resultant creole structure. The TAM particles in Kristang also can be divided between those that are derived from a Portuguese adverb, and those that are derived from a verb. This has implications on how the lexifier's input was transferred to the creole, and how "broken" it may have been. As we can see, Kristang's TAM particles are similar to those in Malay, however the difference between the function of *ta* with the other particles does not have a parallel in Malay. The distinction is both the function of them, and whether they can occur as fragments, as well as that they come from different syntactic categories in Portuguese (Pinharanda Nunes 2012, 320).

6.6 Genitive Patterns

6.6.1 Introduction

In this section I investigate the genitive patterns that are found in the Luso-Asian creoles. In these four, three different genitive forms can be found, each of which has a parallel with a language within their ecologies. This section will outline the three types of genitives, as well as where we can find similar genitives in other languages. We will also see how these relate with each other, and where we are more likely to find one type of genitive over the other.

In addition to describing the possible origins of the types of genitives in the creoles, I also look at their patterns of usage according to a methodology in Baxter and Bastos (2012) whereby it looks at patterns according to semantics. We will see the relationship of genitive patterns with semantics and how alienable and inalienable possessives can be expressed in

the creoles. We will also look at the effects of decreolization and language ecology on the four languages, and how the genitive patterns have been affected.

Baxter & Bastos (2012) identified two different types of genitive found in Luso-Asian creoles: pre-nominal and post-nominal. What they term the pre-nominal genitive takes a form similar to that of Portuguese, in which a preposition *di* is used:

249. Kristang (Baxter & Bastos 2012)

palasu di rei

palace of king

"Palace of the king"

Because of this preposition, I refer to it as the **DI-type** in this thesis.

What they term the post-nominal is influenced by Asian languages where there was Portuguese contact, in which a particle following the possessor indicates genitive:

250. Kristang (Baxter & Bastos 2012)

eli sa kaza

3SG GEN house

"His house"

I refer to this in this thesis as the **SA-type**.

Baxter & Bastos (2012) indicate that the origin of the SA-type is likely to have been in Portuguese colonies in South Asia. They arrived in Malacca after having already conducted trade and set up colonies in Africa and the Indian subcontinent, so the pidginized Portuguese that arrived in Malacca in 1511 could have already contained the SA-type genitive. There are examples in the Indian languages where Portuguese colonies and creoles were present. One of these was Cochin, where the local language is Malayalam.

251. Malayalam (Asher & Kumari 1997: 173, as cited in Baxter & Bastos 2012: 49)

sittayuṭe viiṭə

Sita-GEN house

"Sita's house"

Another area of Portuguese colonization in India was in Korlai, present-day Maharashtra state, where Marathi is a language indigenous to the area. It also exhibits a genitive resembling the SA-form found in Luso-Asian creoles:

252.   Marathi (Clements 1996: 140, as cited in Baxter & Bastos 2012: 51)

Kapil-tsa   ghǝr

Kapil-GEN   house

"Kapil's house"

The contact varieties of Portuguese that arrived in Malacca from 1511 would have had SA-type genitives that could have had origins in these Indian languages. Once in Malacca, languages that were spoken there would have reinforced the SA-type, as their structure is also similar:

253.   Baba Malay (Baxter 1988: 92)

gua   punya rumah

1SG   POSS house

"My house"

254.   Hokkien (Baxter 1988: 92)

huà   é   cchǔ

1SG   GEN   house

"My house"

The language ecologies of the various Luso-Asian creoles repeatedly provided the structure that would eventually lead to the SA-type in Kristang and its related languages. The likely lexical origin of this genitive particle is the Portuguese 3SG possessive pronoun *sua*. When used in isolation it can still express the third person singular possessive, but it can also be placed after other pronouns or NPs to indicate the genitive.

The DI-type genitive in all four of the creoles is structurally and lexically similar to Portuguese. It takes the form of a preposition, derived from Portuguese *de* "of", and written *di* to reflect the pronunciation in the creoles.

A third form of genitive is widely found in TC only. This resembles the genitive of Malay and Indonesian structurally. It is formed by stating the possessee followed by the possessor with no intervening morphology. In this sense it has a similar word order as the DI-type genitive, but does not have any particles or prepositions expressing possession.

255.   Malay (native speaker)
rumah saya
house 1SG
"My house"

256.   Malay (native speaker)
orang Melaka
person Malacca
"People of Malacca"

As this type does not use any particle or preposition, I refer to it in this thesis as the **bare genitive**.

All four of the creoles in this study make use of both of these features (and a third in TC). The usage patterns tend to follow the same semantic trends, but there are differences in how often each one occurs. The following sections show how the genitives work in each of the languages.

This is addressed in APICS Chapter 2, "Order of Possessor and Possessum" (Huber 2013). Of the creoles described, some have only the order of possessor-possessum, some only have possessum-possessor, and some have both. For instance, Korlai only has the order of possessor-possessum, which mirrors the SA-type genitive seen in the four creoles in this thesis:

257.   Korlai (Clements 2007: 168)
Pedru su   kadz
Pedru GEN   house
"Pedru's house"

Diu Indo-Portuguese, which has both orders of genitive, has a 90% frequency of possessum-possessor, which resembles the DI-type:

258.   Diu Indo-Portuguese (Cardoso 2009: 175)

kurəsãw	də	makak	dẽt	də-el	mem	korp
heart	of	monkey	inside	of-3SG	EMPH	body

"The monkey's heart is inside his own body"

The remaining 10% are possessor-possessum, although the example given does not show a construction resembling the SA-type, but rather uses an inversion of the DI-type:

259.   Diu Indo-Portuguese (Cardoso 2009: 169)

də	tɛtɛ	kaz	jə	bẽze-w?
of	Tete	house	already	bless-PST

"Has [he] already blessed Tete's house?"

There was no indication that it was a semantic triggering of one type over the other, as is suggested in this chapter, after Baxter & Bastos (2012).

6.6.2 Genitive Patterns in Kristang

Kristang makes use of both of these genitive structures frequently. Baxter and Bastos (2012) laid out where each of them is used and showed that it can be predicted whether the DI-type or SA-type is more likely to be used by the following semantic categories:

a. Kinship
b. Body part
c. Ownership
d. Other interpersonal relationship
e. Classificatory
f. Part-whole
g. Spatial/locative
h. Origin/source
i. Material composition

They found examples in their corpus of the SA-type and DI-type in all of the categories, but one: the DI-type in kinship, as illustrated in the examples below (all are from Baxter 1980, as cited in Baxter & Bastos 2012)

260.   SA-type

   a. Kinship

      bos   sa   fila

      2SG   GEN daughter

      "your daughter"

   b. Body part

      e(li)   sa   garganta

      3SG   GEN throat

      "his throat"

   c. Ownership

      eli   sa   prau

      3SG   GEN boat

      "his boat"

   d. Other interpersonal relationship

      el   sa   dama

      3SG   GEN mistress

      "his mistress"

   e. Classificatory

      isi   nasang sa   pastu

      this   type   GEN bird

      "this type of bird"

   f. Part-whole

      ngua barku sa   kodra

      one   boat   GEN line

      "a boat's mooring line"

   g. Spatial/locative

      yo   sa   kaza   sa   dianti

      1SG   GEN house GEN front

      "the front of my house"

h. Origin/source

aké   mpoku Rome sa   jenti

that   few   Rome   GEN   people

"those few people from Rome"

i. Material composition

papel   sa   pastu

paper   GEN   bird

"a bird (made) of paper"

261.   DI-type

a. Kinship

[no examples]

b. Body part

rosto   di   sa   inyu

face   of   GEN   godfather

"the face of his godfather"

c. Ownership

palasu di   re

palace of   king

"the palace of the king"

d. Other interpersonal relationship

akeli   re   di   demoni

that   king   of   demon

"the king of demons"

e. Classificatory

isti   prispi di   buntal

this   prince of   puffer-fish

"the puffer-fish king (= prince turned into a puffer-fish)"

f. Part-whole

isi   matera di   trigi

this   pus   of   tiger

"the pus of the tiger"

g. Spatial/locative

dianti   di   kaza

front   of   house

"the front of the house"

h. Origin/source

kletura   di   mundu

creature   of   world

"earthly creature"

i. Material composition

albi   di   oru

tree   of   gold

"a golden tree"

When counting the tokens in each category, they showed the tendency for SA-type genitives to be used with possessives that are more inalienable, as well as those which have human possessors, such as kinship and body parts. On the other hand, the DI-type genitive was used more often with alienable possession and non-human possessors, such as locative and origins. Below in Table 5 I reproduce the results from the study.

	DI tokens	SA tokens	DI percentage	SA percentage
kinship	0	500	0%	100%
body part	3	80	4%	96%
ownership	16	250	6%	94%
other interpersonal relationship	5	40	11%	89%
classificatory	6	6	50%	50%
part-whole	11	8	58%	42%
spatial-locative	51	10	84%	16%
origin/source	34	5	87%	13%
material composition	62	4	87%	13%
TOTAL	188	903	17%	83%

Not mentioned in Baxter & Bastos (2012) is the bare genitive, but there were examples of it in the fieldwork I conducted. While not as extensive as in TC, Kristang also has examples, although it was not included in the count above, because it appears to be a variation of the DI-type, not dissimilar to the variation between long and short forms of pronouns.

262.   Kristang (Laub fieldwork, 2017)

Juang   ta   pegah pintura   kwartu watermelon

John   PROG hold   picture   four   watermelon

"John is holding a picture of four watermelons" (after being asked who has four watermelons, me or John?)

In this instance the subject of the sentence is the main focus of the response, and not what the picture contains. When there is no contrast, the bare genitive may also be used.

263.	Kristang (Laub fieldwork, 2017)

Juang	ta	pegah	pintura	paparika	pretu
John	PROG	hold	picture	pepper	black

"John is holding a picture of a black pepper"

This is in response to asking the consultant to describe what John is holding. There is no contrast or contradiction to a question, so the preposition *di* is unstated in this example.

264.	Kristang (Laub fieldwork, 2017)

bos	ta	pegah	pintura	di	kezu
2SG	PROG	hold	picture	of	cheese

"You are holding a picture of cheese" (after being asked if I have a picture of a banana)

265.	Kristang (Laub fieldwork, 2017)

bos	ta	pegah	pintura	di	gatu
2SG	PROG	hold	picture	of	cat

"You are holding a picture of a cat" (after being asked if I have a picture of a dog)

In these examples, the focused elements are what the pictures contain. As the genitive construction is what the speaker wants to bring attention to, the longer DI-type genitive is used.

6.6.3 Genitive Patterns in Makista

Makista also uses both types of genitive, but the genitives found in the data are overwhelmingly DI-type. This could be indicative of heavy influence from Portuguese in the form of decreolization. The SA-type is present, but is mostly used with personal pronouns rather than a full noun. Additionally, most of the examples of the genitive particle in Makista are with the 1SG pronoun, *iou-sua* "my". One feature that sets Makista apart from Kristang is the usage of possessive pronouns *vossa* and *nossa*, which directly come from Portuguese. This could be an example of emerging grammaticalization of the post-nominal genitive particle, into a morpheme that makes a possessive pronoun, as there are so few examples of this particle with a full NP. In addition to being used as a post-possessor particle, *sua* can be used as a 3SG possessive pronoun, which is also the case in the other creoles. This semantic sense is directly from Portuguese.

Using the same methodology as Baxter & Bastos (2012), I counted tokens of DI-type and SA-type genitives in Makista, looking at Santos Ferreira (1967, 1973, 1985), Barreiros (1944) and Pereira (1899). Examples from every category but one (material composition, SA-type) were found. I also expanded the semantic scope of spatial/locative to include temporal.

266.   SA-type

    a. Kinship

        i. iou-   sua   fila

           1SG   GEN   daughter

           "my daughter" (Santos Ferreira 1973: 183)

        ii. ele-   ça   sobrinha

           3SG   GEN   niece

           "his niece" (Barreiros 1944: 482)

    b. Body part

        i. ilôtro-   sua   ôlo   co   nariz

           3PL   GEN   eye   and   nose

           "their eyes and noses" (Santos Ferreira 1985: 53)

        ii. Siára   sua   pê

           Madam   GEN   foot

           "your [formal] foot" (Pereira 1899: 58)

        iii. iou-   ça   braço

           1SG   GEN   arm

           "my arm" (Barreiros 1944: 575)

    c. Ownership

        i. vôs   sua   casa

           2SG   GEN   house

           "your house" (Santos Ferreira 1967: 21)

        ii. elôtro   sua   pataca

           3PL   GEN   pataca

           "their money" (Pereira 1899: 259)

        iii. Dona Ambrósia-   ça   casa

           Mrs Ambrósia   GEN   house

           "Mrs Ambrósia's house" (Barreiros 1944: 483)

d. Other interpersonal relationship

    i.   iou-   sua   amigo Lopes

        1SG   GEN   friend Lopes

        "my friend Lopes" (Santos Ferreira 1973: 145)

    ii.   ele   ça   gente

        3SG   GEN   people

        "his people" (Barreiros 1944: 478)

e. Classificatory

    i.   dóce   sua   carôço

        sweet GEN   lump

        "a piece of candy" (Santos Ferreira 1967: 79)

    ii.   onsong-   ça   gente

        alone   GEN   people

        "people who are alone" (Barreiros 1944: 21)

f. Part-whole

  na   escola- sa   quartinho

  in   school GEN   little room

  "in a little room of the school" (Santos Ferreira 1985: 91)

g. Spatial/locative/temporal

    i.   iou-   sua   diante

        1SG   GEN   front

        "in front of me" (Santos Ferreira 1973: 157)

    ii.   setenta quatro ça   tufão

        seventy-four GEN   typhoon

        "the typhoon of 1874" (Barreiros 1944: 473)

h. Origin/source

  nôs   sua   lingua cristám

  1PL   GEN   tongue Christian

  "our Christian language" (Santos Ferreira 1967: 67)

i. Material composition

  [no examples]

267.   DI-type

 a. Kinship

  i. fila   di Dom Vico

   daughter of Dom Vico

   "Dom Vico's daughter" (Santos Ferreira 1973: 152)

  ii. irmão de Abel

   brother of Abel

   "Abel's brother" (Pereira 1899: 193)

  iii. familia de Julia

   family of Julia

   "Julia's family" (Barreiros 1944: 155)

 b. Body part

  i. ôlo de vôs

   eye of 2SG

   "your eyes" (Santos Ferreira 1967: 28)

  ii. pê de galinha

   foot of chicken

   "chicken's foot" (Pereira 1899: 125)

  iii. ôlo de elôtro

   eye of 3PL

   "their eyes" (Barreiros 1944: 242)

 c. Ownership

  i. bota di capitám

   boat of captain

   "the captain's boat" (Santos Ferreira 1973: 30)

  ii. ropa   de sua pai

   clothing of his father

   "his father's clothing" (Pereira 1899: 60)

  iii. caza de vôsso tio

   house of your uncle

   "your uncle's house" (Barreiros 1944: 34)

d. Other interpersonal relationship

    i. amigo di   Juám

        friend of   Juám

        "Juám's friend" (Santos Ferreira 1985: 57)

    ii. amiga de   mamã

        friend of   mother

        "mother's friend" (Barreiros 1944: 475)

e. Classificatory

    i. unga   xicra   de   chá

        one   cup   of   tea

        "a cup of tea" (Santos Ferreira 1967: 18)

    ii. arve   de   côco

        tree   of   coconut

        "coconut tree" (Pereira 1899: 190)

    iii. brinco de   criança-criança

        toy   of   child-child

        "children's toys" (Barreiros 1944: 242)

f. Part-whole

    i. casca   de   ôvo

        shell   of   egg

        "eggshell" (Santos Ferreira 1967: 68)

    ii. ponta   de   sapato

        tip   of   shoe

        "shoe tip" (Barreiros 1944: 130)

g. Spatial/locative/temporal

    i. Macau de   agora

        Macau of   now

        "the Macau of today" (Santos Ferreira 1967: 69)

    ii. tufão   de   74

        typhoon   of   (18)74

        "the typhoon of 1874" (Pereira 1899: 259)

    iii. Dentro de   greza

        inside of   church

        "inside the church" (Barreiros 1944: 33)

h. Origin/source

    i. águ   di   poço

      water of   well

      "well water" (Santos Ferreira 1985: 31)

    ii. nôsso lingu de   Macau

      our   tongue of   Macau

      "our language of Macau" (Pereira 1899: 194)

    iii. vapor   de   Hong-Kong
      steamboat of   Hong Kong
      "steamboat from Hong Kong" (Barreiros 1944: 250)

i. Material composition

    i. pulséra   di   prata

      bracelet   of   silver

      "silver bracelet" (Santos Ferreira 1985: 38)

    ii. chapeo de   pruma

      hat   of   feather

      "feather hat" (Pereira 1899: 125)

    iii. coração   di   oro

      heart   of   gold

      "heart of gold" (Barreiros 1944: 572)

The table below shows the distributions:

Table 13: Makista Genitive Patterns

	DI tokens	**SA tokens**	**DI percentage**	**SA percentage**
kinship	80	80	50%	50%
body part	133	66	66.8%	33.2%
ownership	131	103	56%	44%
other interpersonal relationship	42	87	32.6%	67.4%
classificatory	665	15	97.8%	2.2%
part-whole	114	1	99.1%	0.9%
spatial/locative/temporal	494	8	98.4%	1.6%
origin/source	283	17	94.3%	5.7%
material composition	65	0	100%	0%
TOTAL	2,007	377	84.2%	15.8%

Table 13 above shows the totals from all corpora, while table 14 below demonstrates that the semantic patterns of the genitive tokens from mid-20th-century Ferreira are not dissimilar from those of the late 19th and early 20th centuries from Pereira and Barreiros.

171

Table 14: Earlier vs Later

	DI tokens earlier corpora		DI tokens later corpora		SA tokens earlier corpora		SA tokens later corpora	
kinship	47	4.1%	33	4%	54	22.1%	26	19.5%
body part	92	8%	41	4.8%	41	16.8%	25	18.8%
ownership	83	7.2%	48	5.6%	70	28.7%	33	24.8%
other interpersonal relationship	19	1.6%	23	2.7%	53	21.7%	34	25.6%
classificatory	376	32.6%	289	33.8%	7	2.9%	8	6%
part-whole	48	4.2%	66	7.7%	1	0.4%	0	0%
spatial/locative/temporal	323	28%	171	20%	6	2.5%	2	1.5%
origin/source	134	11.6%	149	17.4%	12	4.9%	5	3.8%
material composition	31	2.7%	34	4%	0	0%	0	0%
TOTAL	1153	100%	854	100%	244	100%	133	100%

In both the earlier and later corpora, the token counts show that the SA-type genitive is more likely to be found in semantic categories related to a human possessor (the upper rows of the table).

6.6.4 Genitive Patterns in Batavia Creole

BC's genitive construction use is similar to that of Makista. While it uses both, the DI-type ones are more commonly found. The corpus of BC in Maurer (2011) is much smaller than that of Makista, so while not every category has an example, this does not mean that they would not have been possible in BC.

268. SA-type

    a. Kinship

       moler   su   pay

       wife   GEN   father

       "wife's father" (Maurer 2011: 129)

    b. Body part

       ilotër   sua   kabesa

       3PL   GEN   head

       "their heads" (Maurer 2011: 135)

    c. Ownership

       ilotër   sua   chapeo

       3PL   GEN   hat

       "their hats" (Maurer 2011: 134)

    d. Other interpersonal relationship

       [no examples]

    e. Classificatory

       tres   peo   su   fundu

       three   foot   GEN   deep

       "three feet deep" (Maurer 2011: 134)

    f. Part-whole

       [no examples]

    g. Spatial/locative

       [no examples]

    h. Origin/source

       [no examples]

    i. Material composition

       [no examples]

269. DI-type

    a. Kinship

       [no examples]

    b. Body part

       [no examples]

c. Ownership

chang	di	sua	moler	su	pay
land	of	his	wife	GEN	father

"the property of his wife's father" (Maurer 2011: 129)

d. Other interpersonal relationship

[no examples]

e. Classificatory

fula	di	dudaim
flower	of	pandanus

"pandanus flowers" (Maurer 2011: 131)

f. Part-whole

[no examples]

g. Spatial/locative

nu	meo	di	kaminyu
in	middle	of	road

"in the middle of the road" (Maurer 2011: 130)

h. Origin/source

ung	cabalu	di	Bima
a	horse	of	Bim

"a horse from Bima" (Maurer 2011: 129)

i. Material composition

[no examples found]

The distribution of tokens is shown in Table 8 below.

Table 15: Batavia Creole Genitive Patterns

	DI tokens	**SA tokens**	**DI percentage**	**SA percentage**
kinship	0	2	0%	100%
body part	0	1	0%	100%
ownership	2	2	50%	50%
other interpersonal relationship	0	0	N/A	N/A
classificatory	11	1	92%	8%
part-whole	0	0	N/A	N/A
spatial-locative	9	0	100%	0%
origin/source	2	0	100%	0%
material composition	0	0	N/A	N/A
TOTAL	24	6	80%	20%

It is obvious that the sample size from Maurer's text is far too small to make any kind of conclusion. However, it is interesting to note that much like Makista, BC trends more heavily towards the DI-type genitive (four times more than the SA-type), likely as a result of more Portuguese influence. Additionally, where the SA-type genitives occur are closer to the top end of this chart, just as in Makista. These ones are the genitives that are more likely to have inalienable possessions and human possessors, and are where the SA-type genitive is more prominent in Kristang, which uses both structures frequently.

6.6.5 Genitive Patterns in Tugu Creole

TC is an interesting case, as it uses both the genitive forms described in this section, but it also uses a third form. This form is indicated by word order rather than explicit morphology, in which the possessee is followed by the possessor. This feature is common in Malay and Indonesian to indicate possessive.

The following are examples of the genitive forms in TC, organized by semantic category.

270. SA-type

    a. Kinship

       yo   sua   pay

       1SG   GEN   father

       "my father" (Maurer 2011: 193)

    b. Body part

       me   sua   korsang

       1SG   GEN   heart

       "my heart" (Maurer 2011: 178)

    c. Ownership

       yo   sua   sapatu

       1SG   GEN   shoe

       "my shoes" (Maurer 2011: 169)

    d. Other interpersonal relationship

       nos   sua   imigu

       1PL   GEN   enemy

       "our enemies" (Maurer 2011: 183)

    e. Classificatory

       mas   ung   sumang   sua   dura   nobu

       more one   week   GEN   length new

       "after one more week" (Maurer 2011: 161)

    f. Part-whole

       neli   sua   achamentu

       rice   GEN   yield

       "yield of rice"

    g. Spatial/locative

       pilang su   may

       mortar GEN   inside

       "in the mortar" (Maurer 2011: 158)

    h. Origin/source

       nos   sua   orsang

       1PL   GEN   prayer

       "our prayers" (Maurer 2011: 169)

i. Material composition

[no examples]

271. DI-type

a. Kinship

[no examples]

b. Body part

[no examples]

c. Ownership

[no examples]

d. Other interpersonal relationship

[no examples]

e. Classificatory

dia   di   alegria

day   of   joy

"day of joy" (Maurer 2011: 177)

f. Part-whole

mas   di   prendesu

more   of   religion

"more of the religion" (Maurer 2011: 146)

g. Spatial/locative

[no examples]

h. Origin/source

acha   tantu   bondadi   di   siu   gobernemen

get   many   favors   of   their   government

"get a lot of favors from the government (Maurer 2011: 146)

i. Material composition

[no examples]

272. Bare genitive

a. Kinship

[no examples]

b. Body part

[no examples]

c. Ownership

[no examples]

d. Other interpersonal relationship

[no examples]

e. Classificatory

unga   iskola gobernemen

one   school government

"a government school" (Maurer 2011: 146-7)

f. Part-whole

pontu   somenti

tip   seedling

"tip of the seedlings" (Maurer 2011: 153)

g. Spatial/locative

na   riba   furkila

on   top   pole

"on poles" (Maurer 2011: 149)

h. Origin/source

adjiti klapa

oil   coconut

"coconut oil" (Maurer 2011: 164)

i. Material composition

rupia   prata

rupee   silver

"silver rupees" (Maurer 2011: 179)

The table below shows the totals of the tokens in TC genitive constructions.

Table 16: Tugu Creole Genitive Patterns

	DI tokens	SA tokens	bare genitive tokens	DI percentage	SA percentage	bare genitive percentage
kinship	0	3	0	0%	100%	0%
ody part	0	6	0	0%	100%	0%
ownership	0	13	0	0%	100%	0%
other interpersonal relationship	0	3	0	0%	100%	0%
classificatory	4	6	52	6%	10%	84%
part-whole	1	6	6	8%	46%	46%
spatial-locative	0	3	3	0%	50%	50%
origin/source	4	3	29	11%	8%	81%
material composition	0	0	3	0%	0%	100%
TOTAL	9	43	32	11%	51%	38%

TC, with its three-way system of genitives, has near equal instances of the SA-type and the bare genitive constructions. These are both indicative of influence from Malay and other Asian languages, while the much more infrequent DI-type genitive is similar to the Portuguese construction. Unfortunately, there are too few examples in the corpus to make any kind of conclusion as to the patterns of DI-type genitive usage.

As cited by Baxter (1988) and Baxter and Bastos (2012), the SA-type genitive *punya* is in Bazaar Malay, but is not found in standard Malay or Indonesian as a genitive particle. Considering the area of Jakarta where the TC speakers lived, it could be an influence from more native Malay (rather than the Chinese-influenced Bazaar Malay) that introduced the bare genitive into TC.

The origins of the two types of genitive (and three in TC) are obvious; the pre-possessor *di* comes from Portuguese, while the post-possessor *sa* comes from an Indian language with reinforcement from Baba Malay. For TC, the bare genitive has a direct parallel in Malay and Indonesian. What is less clear, however, is where this tendency came from. The

difference between post-possessor and pre-possessor fall along general alienable/inalienable or human/non-human lines, but it is as yet not clear why this is the case and not the opposite.

With the major caveat that the data set for TC is very small, we can see that the bare genitive occurs more often in inalienable and non-human instances, where in Kristang one would be more likely to find the pre-possessor genitive. With only 11% (or 9 total) of the genitive tokens in the TC data set being pre-possessor it is again too small to make a conclusive statement regarding the trends of where it is found, but 6 of the 9 are found in inalienable and non-human categories (the other 3 being part of the ownership class).
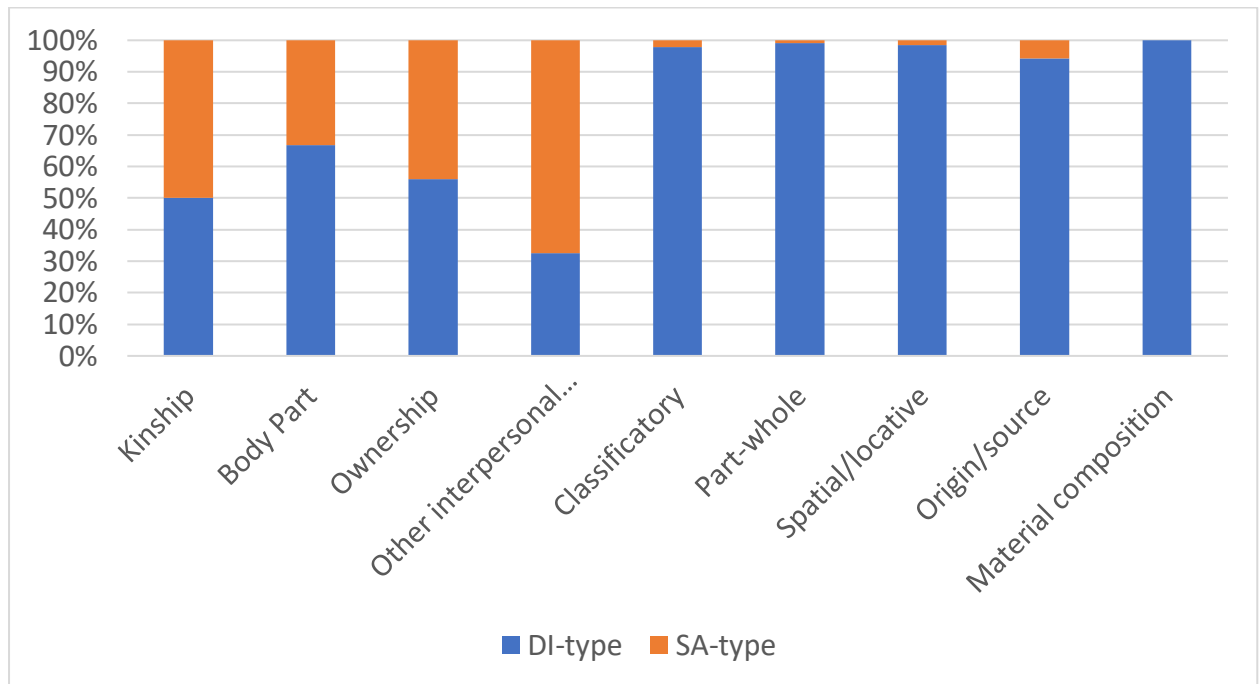
6.6.6 Discussion

The following figures illustrate the differences in distributions of the genitives.

Chart 5: Kristang Genitive Patterns (Baxter & Bastos 2012)
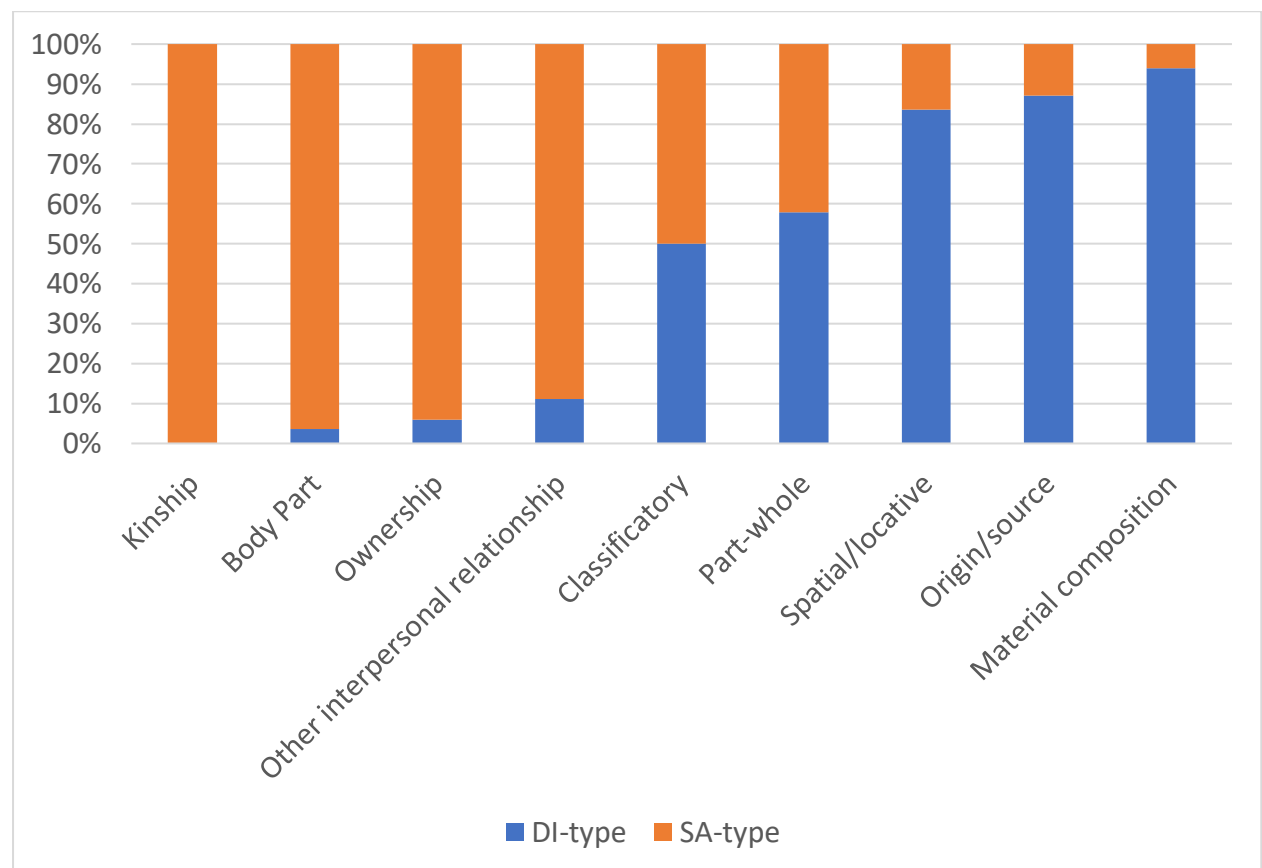
Chart 6: Makista Genitives (Santos Ferreira 1967, 1973, 1985; Pereira 1899; Barreiros 1943)

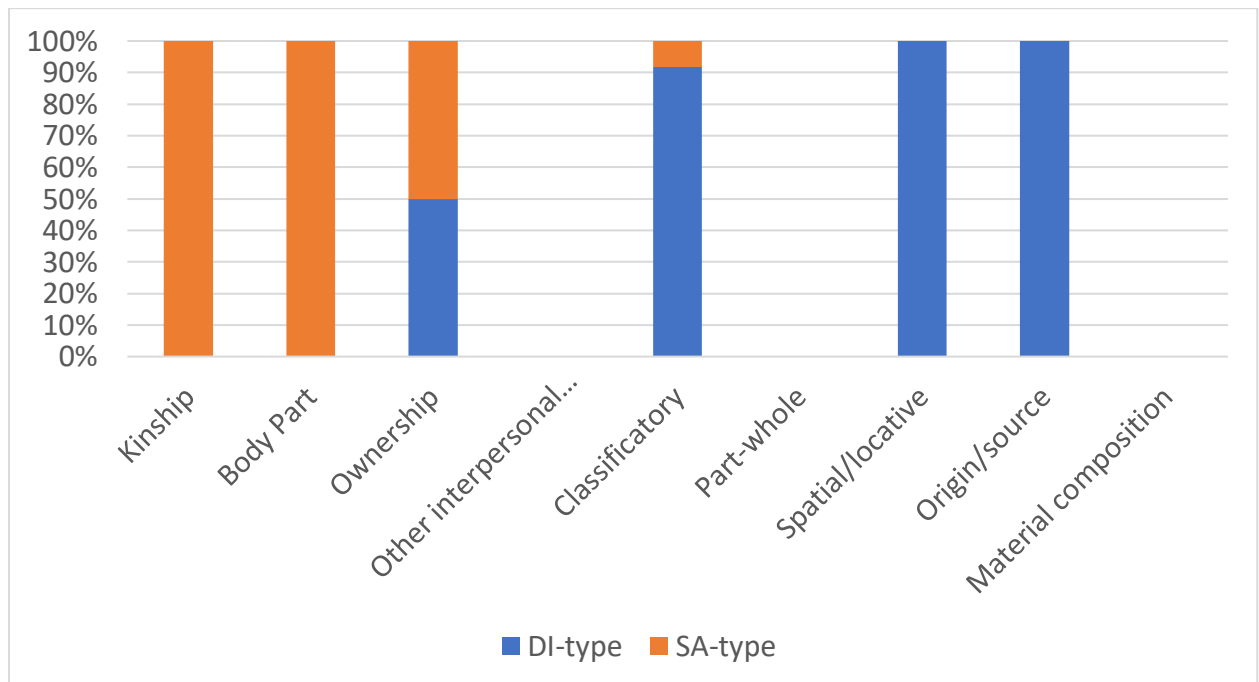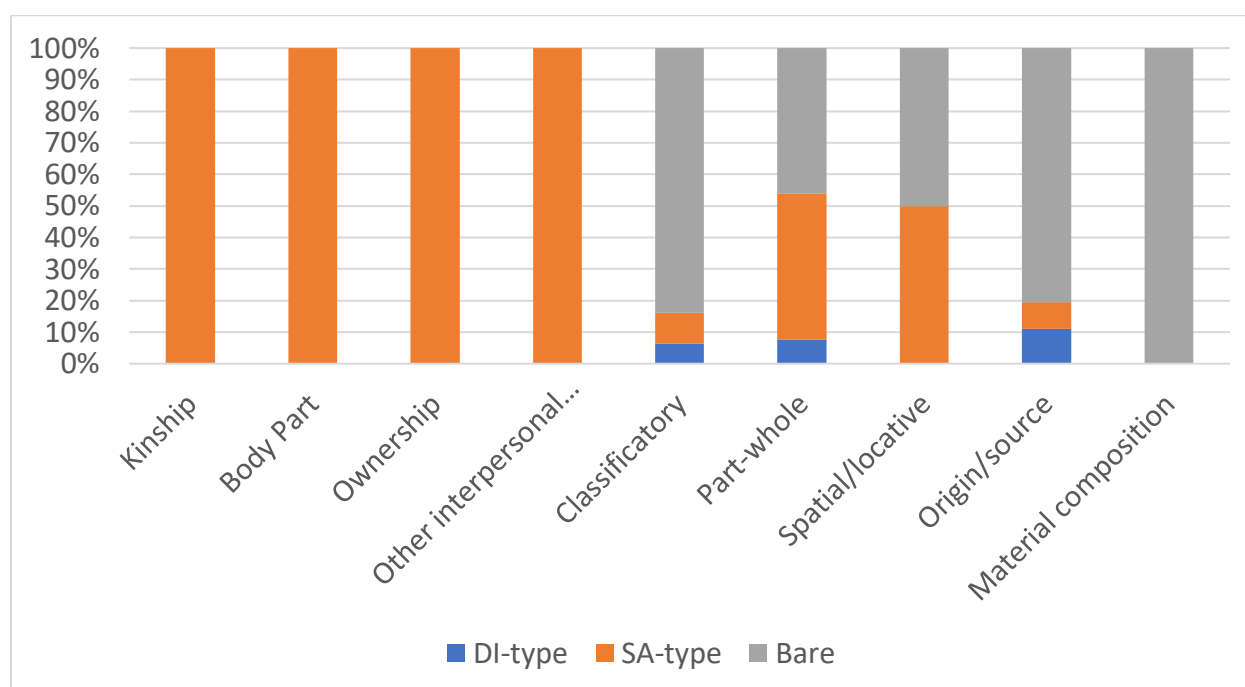*Chart 7: Batavia Creole Genitive Patterns (Maurer 2011)*

Chart 8: Tugu Creole Genitive Patterns (Maurer 2011)

A look into the genitive patterns in these four languages can show us what happens when two competing forms both end up staying within the language. With the methodology used by Baxter and Bastos (2012) we can see that there are semantic patterns that show where a type of genitive tend to be used. This crucially also shows that when the creoles were formed syntactic elements from Portuguese did remain, even when a feature from the substrate was adopted into the language. This is even the case in Kristang, where there was little contact with Portuguese after the 17th Century. Where we do see the biggest difference is in Makista, where continuing Portuguese contact transformed the genitive patterns into one that heavily uses the *di*-form. Change in the other direction occurred in TC, where the *sa*-form was heavily preferred, and also has the presence of the bare genitive form. The change in genitive patterns in both Makista and TC are part of the overall convergences we see with Portuguese and Malay, respectively.

For the purposes of this study, I included animals' body parts in the "body part" category while counting tokens of genitive. In every case, they used the DI-type genitive, and every SA-type in this category was belonging to a human.

273.   Makista (Barreiros 1944: 32)

rabo   de   porco
tail   of   pig
"pig's tail"

274.   Makista (Barreiros 1944: 243)

mão   de   porco

hand   of   pork

"pig trotters"

275.   Makista (Pereira 1899: 125)

pê   de   galinha

foot   of   chicken

"chicken feet"

Conversely, human body parts were found with both SA- and DI-type genitives

276.   Makista (Barreiros 1944: 138)

mão   de   sua   pai

hand   of   his   father

"his father's hand"

277.   Makista (Barreiros 1944: 484)

vosotro-   ça   mam

2PL   GEN   hand

"your hands"

We can see that the overall tendency of SA-type genitives to cluster around the more inalienable and human semantic types occurs in all four languages. The main differences are that in Makista the DI-type has a wider distribution than the SA-type, while in Kristang and BC it is more balanced. In TC the bare genitive is found on the right side of the figures, where the DI-type is preferred in the other languages. This suggests that Malay had a similar influence on TC genitives that Portuguese did on Makista.

We will revisit these patterns in Section 7.5, where we look at the implications of creolization and decreolization on simplicity of morphosyntactic structure. What these charts reveal is a simplification in the distribution of genitives in Makista as compared to Kristang, and is likely from influence by usage of Portuguese in Macau. On the other hand, we have a more complex genitive system in TC where there are three genitive structures. Out of the four, TC has the most influence from the substrates, and what resulted was further complication of the genitive structure.

6.7 Passive

6.7.1 Introduction

The passive is the shifting of the patient of a transitive verb to a subject position. It results in a change of focus. The four creoles in this study have varying types of passive, which have similar structures as in the input languages. Most of the passive morphology is not from Portuguese directly, although there are some instances in Makista that show more direct Portuguese influence. The examples in this section show usage of passive that either came from Portuguese, or do not have an obvious connection to the substrates. However, there are some similarities which will be discussed, especially in the Kristang examples This feature is also interesting in that we have a direct borrowing from a substrate, that of the passive prefix *di* in TC. This has been taken directly from Malay without a Portuguese calque, as is often the case with other morphemes in these creoles.

6.7.2 Passive in the Lexifiers and Substrates

Passives in Portuguese are formed by using the copula and the past participle:

278.   Portuguese (Ackerlind & Jones-Kellogg 2011: 150)

a	porta	foi	aberta		por	José
the	door	was	opened		by	José

"The door was opened by José"

When the agent of the verb is not known, the *por* + NP may be omitted:

279.   Portuguese (Ackerlind & Jones-Kellogg 2011: 150)

a	porta	foi	aberta
the	door	was	opened

"The door was opened"

In an active sentence, the word order is different:

280.   Portuguese (Ackerlind & Jones-Kellogg 2011: 149)

José	abre	a	porta
José	opens	the	door

"José opens the door"

The form of the verb changes as well. We can see that in example 280 the present tense third person singular form *abre* is used, while in the passive voice examples the past participle *aberta* was used.

Aberta is an irregular past participle, deriving from the infinitive *abrir* in Portuguese. Regular past participles in Portuguese depend on the type of verb being used. The following table, from Ackerlind and Jones-Kellog (2011: 143), illustrates these suffixes.

Table 17: Portuguese Past Participles

	-ar	**-er**	**-ir**
ending	-ado	-ido	-ido
infinitive	flar	Comer	partir
past participle	falado/a	comido/a	partido/a

Verbs in the four creoles generally come from the infinitive form. In the Portuguese examples above, the stress comes on the final syllable of infinitive verbs (*falar, comer, partir*). In Kristang these verbs are expressed with the final *r* dropped, but the stress remains (*falá, kumí*). Unlike Portuguese, verbs do not take endings for person or number, nor do they take endings for tense. However, there are some remnants or perhaps reintroductions of these past participle endings, which we will see in examples below.

Nomoto (2006) explains the passive constructions in Malay, which expands upon Voskuil (2000) regarding Indonesian. They classify Indonesian and Malay voice into four categories: morphological active, morphological passive, bare active and bare passive.

281.   Malay (Nomoto 2006) morphological active

dia   sudah membaca   buku   itu

3SG   PFV   read   book   that

"She has already read the book"

282.   Malay (Nomoto 2006) morphological passive

buku   itu   di-baca   (oleh)-nya

book   that   PASS-read   (by)-3SG

"The book is read by her"

283.   Malay (Nomoto 2006) bare active

dia   sudah baca   buku   itu

3SG   PFV   read   book   that

"She has read the book"

284.   Malay (Nomoto 2006) bare passive

buku   itu   sudah dia   baca

book   that   PFV   3SG   read

"She has already read the book"

There is evidence that the passives in some of these creoles comes from Malay influence, especially that of the bare passive.

There are some similarities between Malay and Portuguese passives, in that they both make use of verbal morphology, and they both have the *by + NP* phrase after the verb. However, Malay expresses the passive with a prefix, and it has the bare passive as well.

The passive in Cantonese is marked by the preposition *bei2* "by" before a noun phrase.

285.　　　Cantonese (Matthews & Yip 1994: 149)

ji5chin4	go2	go3	bei3syu1	bei2	ging1lei5	caau2-zo2
before	that	CLF	secretary	by	manager	sack-PFV

"The previous secretary was sacked by the manager"

This construction must precede the verb in a sentence (Matthews and Yip 1994: 149). The agent must also be stated in an overt passive construction. If there is not a definite agent, the phrases *yàhn* "person" or *yéh* "thing" are used.

286.　　　Cantonese (Matthews & Yip 1994: 151)

lei5	bei2	jan4	ngak7-zo2	la3
2SG	by	person	cheat-PFV	PRT

"You've been cheated"

287.　　　Cantonese (Matthews & Yip 1994: 151)

ngo5	jau6	bei2	je5	ngaau5	can1
1SG	again	by	thing	bite	PRT

"I've been bitten again"

Matthews and Yip point out that in spoken Cantonese an object can be topicalized by shifting it to the front of the sentence, and is typically translated into the passive in English.

288.　　　Cantonese (Matthews & Yip 1994: 151)

go2	tou3	hei3	jat1ding6	yiu3	tai2
that	CL	film	must	need	see

"That film has to be seen"

They also mention that the passive in Cantonese generally has an adversative connotation, and that phrases using *bei2* are more restricted than the passive in English which has a more neutral tone.

This differs from the Portuguese and Malay passives, where the *by + NP* phrase is not required. This could be due to the fact that there is no other verbal morphology, as there is in Portuguese and Malay.

In APiCS, Chapter 90 (Haspelmath 2013c) addresses passive constructions, and identifies four types: typical passive, passive without verbal coding, other atypical passive, and absence of passive. The chapter on Kristang (Baxter 2013) notes the presence of typical passive and passive without verbal coding, with the former category having a higher frequency of 70%. Korlai (Clements 2013) and Sri Lanka Portuguese (Smith 2013), have only one type, that of the typical passive construction:

289.   Korlai (Clements 2013)

animal   ʤiko   savad

animal   became   saved

"The animals were saved"

290.   Sri Lanka Portuguese (Smith 1974-5: 4290)

isti   kaaza jaa-fikaa   maraa-du   miɲa   paay   vɔɔnda

this   house PST-become   build-ADJZ   1SG.GEN   father by

"This house was build by my father"

Diu (Cardoso 2013) does not likely have any passive constructions. The four creoles in this thesis resemble that of Kristang, although TC also uses a morpheme *di* which indicates the passive, a direct borrowing from Malay (see Section 6.7.5).

6.7.3 Passive in Kristang

Kristang exhibits two different ways to express the passive. One of which is the adversative passive, mentioned earlier, which uses the verb *toka*. This is a feature that has a parallel in Malay with the verb *kena*, having the same meaning as "touch" or "strike" but also when used with another verb expresses an adversative passive. There are other examples of the passive in Kristang in which there is a simple reordering of the words from the canonical word order.

291.   Kristang (Marbeck 1995)

isti   libru jah   skribeh

This   book   PFV   write

"This book was written [in memory of]"

This example has what would be a passive in English, but in Kristang rather has a reordering of the word order, with the agent not stated. The agent in this case is the author of the book, as it is the dedication page of the volume. Therefore, if an agent were overtly stated, it would be the first person singular pronoun *yo*.

292.   Kristang (Marbeck 1995)

Pasku	lo	selebrah	na	Abril
Easter	FUT	celebrate	in	April

"Easter is celebrated in April"

Again, if this were written as "We celebrate Easter in April", it would be written as:

293.   Kristang (own knowledge)

nus	selebrah	Pasku	na	Abril
1PL	celebrate	Easter	in	April

"We celebrate Easter in April"

Rather than marking the passive with overt morphology, the words are reordered. The patient of the verb "celebrate" is spoken before the verb, while the agent is left unspoken.

294.   Kristang (Pillai 2011)

ke	Ash Wednesday	nus	logu	fasting.	Nungka	nadi
on	Ash Wednesday	1PL	PFV	fasting	NEG	NEG

kumi	karni.	Galinya,	porku,	baka	nadi	kumih
eat	meat	chicken	pork	beef	NEG	eat

"On Ash Wednesday we will fast. We will not eat meat. We won't eat chicken, pork, or beef [Chicken, pork, or beef will not be eaten]"

This is in an explanation of the fasting period during lent. The second and third sentences do not have a stated agent, but it is clear from the preceding one that it is *nus* "we (the people of the Portuguese Settlement)". The last one has the verb *kumi* "eat", a transitive, verb, but the object is not stated after it, but rather before, where the subject is normally stated. The patient *galinya, porku, baka* is also stated before the negative particle. This is similar to the Marbeck example in which nothing intervenes between the pre-verbal particles

and the verb itself. This differs from the Malay bare passive in that the agent is not spoken in the sentence. In Malay, the agent appears between the aspect particle and the main verb, which is not possible in Kristang. This could be an example of topic fronting, rather than actual passive. We can see in examples from the other creoles, as well as from the substrates, where there is actual morphology in the usage of passive. In these Kristang examples we simply see a shift of the patient to a preverbal position, with the agent unstated.

6.7.4 Passive in Makista

Examples of passive in Makista are rare, but when used can be influenced by Portuguese syntax.

295.   Makista (Santos Ferreira 1973: 157)

su	fila	já	ficá	intrujado	pa	dôs	jóvi
3SG.POSS	daughter	PFV	be	deceived	by	two	boy

"His daughter was deceived by two boys"

This does not use the copula *sã*, but rather uses the verb *ficá* "to be, to find oneself". Later in the same paragraph we encounter the verb "deceive":

296.   Makista (Santos Ferreira 1973: 157)

dôs	americano	jóvi	já	intruzá,	cavá	já	dále
two	American	boy	PFV	deceive	then	PFV	hit

"Two American boys deceived her, then hit her"

Other instances of the Portuguese-influenced past participles can be found in the earlier corpora:

297.   Makista (Barreiros 1944: 134)

nunca	olá	más	nada	que	tres	bandera,	pindurado	na
NEG	see	more	nothing	REL	three	flag	hung	on

janela	de	taverna
window	of	tavern

"I saw nothing more than three flags, hung on the window of the tavern"

298.   Makista (Pereira 1899: 323)

governo	nôvo	sam	capaz	e	já	virá	tudo. Mas
governor	new	COP	capable	and	PFV	change	all more

um	pôco	tempo	tudo	lôgo	ficá	virado
one	little	time	all	FI	stay	changed

"The new governor is very capable, and has changed everything. In a bit more time everything will be changed"

Makista exhibits examples of Portuguese verbal morphology through the usage of the past participle suffix as shown above. Based on the Makista texts alone it is not possible to conclude whether *adu* is a productive morpheme or not. However, there are many examples of a past participle-like adjective, and a verbal equivalent, as seen in the examples above. Texts from the earlier corpora exhibit simple, unmarked passive, similar to Kristang.

299.   Makista (Barreiros 1944: 357)

aquelle	tempo,	nôsso	rê	chomá	Dom Manuel
that	time	our	king	call	Dom Manuel

"In that time [of history], our king was called Dom Manuel"

300.   Makista (Pereira 1899: 259)

tudo	casa	de	Macáo	qui	já	vendê	pra	china
all	house	of	Macau	REL	PFV	sell	to	Chinese

"All of the houses of Macau which were sold to the Chinese"

301.   Makista (Pereira 1899: 125)

lôgo	ôvi	tamêm	falá	na	pulpito	que
FI	hear	also	speak	on	pulpit	REL

"[one] will also hear said on the pulpit that…"

6.7.5 Passive in Tugu Creole

TC exhibits various ways to express the passive.

302.   Tugu Creole (Maurer 2011: 148)

ake	albër	neli,	ele	tara	ku	tudu	djenti	bong	ki	riku,
DEM	plant	rice	3SG	plant	with	all	people	good	REL	rich

bong	ki	pobri
good	REL	poor

"This rice plant is planted by everybody, be they rich or poor"

This example of the passive is similar to the Kristang examples above. In this case, the rice plant, the patient, is placed before the verb *tara*, "to plant". Unlike the Kristang examples, however, this one does overtly express the agent, or those who are doing the planting. These people are marked with *ku*, which can either mean "with" or be used as an object marker.

TC also has examples of bare passive in which the agent is unstated, and the theme is stated before the verb.

303.   Tugu Creole (Maurer 2011:155)

tapi	aka	korda	bambu	miste	rusa	djantong	figu
but	DEM	string	bamboo	must	rub	blossom	banana

"But the strings must be rubbed with banana blossoms"

304.   Tugu Creole (Maurer 2011: 159)

sua	tawdu	podi	da	kumi	kabalu
GEN	rice bran	can	give	eat	horse

"Its rice bran can be fed to the horses"

This is an example of the serial verb construction using *da*, "to give". Although this is the only example of *da kumi*, "to feed" in the TC texts, we can see that serial verb constructions using *da* typically have an agent that is actively carrying out an action.

305.   Tugu Creole (Maurer 2011: 73)

isti	belu	da	sabe	kung	ile	ki
DEM	old.man	give	know	ACC	3SG	COMP

"The old man told him that…"

Of the four creoles studied in this thesis, TC has by far the most Malay influence. This is especially seen in the case of the passive, where we see instances that resemble both the morphological and the bare passive in Malay.

306.   Tugu Creole (Maurer 2011: 148)

lama	ki	dja	fay	di-bira	bar	di	garu
mud	REL	PFV	make	PASS-get	earth	of	harrow

"The muddy ground that has been turned into soil by the harrow"

Much like the bare and morphological passives in Indonesian and Malay, we can find both types of passive on the same verb.

307.   Tugu Creole (Maurer 2011: 149)

ake	albër	neli	di-tara	unga	unga
DEM	plant	rice	PASS-plant	one	one

"The rice plants are planted one by one"

Here, we have the verb *tara* "to plant" with the passive prefix *di-* in the front. We can see that it is used in a similar context, but with a morphological passive rather than the bare passive seen earlier.

308.   Tugu Creole (Maurer 2011: 150)

ele	fay	limpu	neli	ake	di	sua	fola	fola;	ki	toka
mara										
3SG	make	clean	rice	DEM	of	GEN	leaf	leaf	REL	touch
bind										

"He separates the rice from its leaves; [the ears] which have to be bound"

Maurer here notes: "The antecedent of the relative clause is absent in the Creole version; the Malay version has *padi* 'unhusked rice', and the German version offers *Ähren* "ears". In this example, we have usage of *toka*, which is the same as in the other creoles, found before the main verb in the sentence. It also in this case has a meaning of obligation.

Serial verb constructions are also used in TC to express a passive:

309.   Tugu Creole (Maurer 2011: 147)

mas	unga	iskola	gobernemen	undi	filu filu	acha	prende	dikausa
more	a	school	government	where	children	get	teach	of.cause
les								
read								

"Furthermore a government school where the children are taught reading"

In this example, the agent of the verb *prende* "to teach" is unstated, while the theme, *filu filu* "children" is placed before the verb.

6.7.6 Discussion

We can see that the input languages have a diverse way of expressing the passive, and some of these features have been passed onto the creole languages. Makista shows the heaviest usage of Portuguese verbal morphology, which can be linked to its linguistic ecology in Macau. On the other hand, we can see that TC has adopted the Malay passive prefix *di-* directly, due to the fact that it was heavily influenced by Malay and Indonesian. This is another example of the language ecology playing a role in the formation of a language's syntactic structure.

6.8 Reduplication

6.8.1 Introduction

This section looks at the phenomenon of reduplication. Reduplication is a process whereby a word is repeated in part or in whole, indicating an inflectional or derivational change. In line with Inkelas & Zoll (2005), I will consider reduplication to be a morphological process. In this section I will take a qualitative method to look at the ways reduplication manifests itself in the four creoles, as well as in the substrate languages. Previous studies of reduplication in these creoles often divides them by grammatical category (Ansaldo & Matthews 2004, Avram 2015). I will also use this method as reduplication even in the same language can have different functions depending on the grammatical category.

Inkelas & Zoll (2005: 14) note that reduplication, while occurring across many languages and serving a morphological function, varies from language to language: "Iconic semantics is not, however, the general rule. Reduplication, especially partial reduplication, is associated cross-linguistically with all sorts of meanings, both inflectional and derivational, whose degree of iconicity is often negligible." Reduplication occurs in many language families of the world and does not always mean the same thing. This section will first investigate reduplication patterns in Sinitic and Malay languages, and then we will see how they compare with the patterns in the four creole languages.

Inkelas & Zoll (2005: 22) lay out four criteria for morphological reduplication:
a. Morphological reduplication serves a morphological purpose
b. Proximity
c. Morphological reduplication involves repetition of a morphological constituent, "potentially truncated to a prosodic constituent"
d. Morphological reduplication involves semantic identity

In the four creoles of this thesis, as well as the substrate languages, reduplication serves a morphological function. The exact type of function can vary between language and grammatical category, but they all go through a morphological change. In the other features in this chapter, I argue that many of the morphosyntactic features that occur in Kristang, Makista, BC, and TC are a result of influence from their input languages, and reduplication is no different. The patterns of reduplication in the input languages have similarities to those found in the creoles. This appears to be a morphological transfer from one language to another, rather than a phonological one.

In this section we will see examples of nominal, verbal, numeral, adjectival, and pronominal reduplication, which is found across many language families, including Sinitic and Austronesian. Shi (2002) describes verbal reduplication in Mandarin to express the delimitative aspect, or for a short time (187).

310.   Mandarin (Shi 2002: 187)

wo3   xiang3 wen4wen   lao3shi1

1SG   want   ask-ask   teacher

"I want to give a try and ask my teacher"

In this example the reduplication of the verb *wen4* to *wen4wen* is an example of inflectional morphology, where the verb retains its grammatical category, and the aspect is affected. It also appears in the same position a non-reduplicated verb would.

Mandarin also uses aspect particles that appear after the verb. In a reduplicated verb, the aspect particle appears in the middle.

311.   Mandarin (Shi 2002: 187)

wo3   zuo2tian1   wan3shang4   kan4-le-kan4   dian4shi4

1SG   yesterday   evening   watch.PFV.watch   TV

"I watched TV for a while yesterday evening"

In this case it has the delimitative aspect, as in the person watched TV for a short while, and the perfective aspect in the middle of the verb expresses the event has finished. Mandarin also has adjectival and nominal reduplication, which act as intensifiers.

312.   Mandarin (Shi 2002: 191)

na4   shi4   yi4ben3   hou4-hou4   de   shu1

That   COP   one.CLF   thick-thick   CLF   book

"That is a very thick book"

In this example the reduplicated adjective *hou4* "thick" retains its grammatical category as an adjective, and its meaning is intensified by the reduplication. Its position in the sentence does not differ from that of a non-reduplicated adjective.

There is also nominal reduplication:

313.   Mandarin (Shi 2002: 197)

ren2-ren2   dou1   zhi1dao4   zhe4-jian4   shi4

person-person all   know   this-CLF   matter

"Everybody knows this matter"

According to Shi, this nominal reduplication gives a universal quantification. Note that in addition to the reduplication of *ren2* "person", *dou1* "all" is also present.

In Cantonese, reduplication can occur on various syntactic categories, but it is particularly important with adjectives (Matthews and Yip 1994: 44). For adjectival reduplication, they identified three types: **adjective-adjective**, **adjective-adjective-*dei***, and **A-B-B adjectives** (Matthews & Yip 1994: 163-5).

Adjective-adjective is reduplication which results in intensifying the meaning of the adjective:

314.   Cantonese (Matthews and Yip 1994: 163)

go3   kau4ceung4   co5   dou3   mun2-mun4   ge3

CL   ball-stadium   sit   till   full-full   PRT

"The seats in the stadium are all filled up"

They also note that bisyllabic adjectives reduplicate in the form AABB:

315.   Cantonese (Matthews and Yip 1994: 163)

syu1fuk9 "comfortable" → syu1-syu1-fuk9-fuk9 "pretty comfortable"

Adjective-adjective-*dei2* is a second type of adjectival reduplication which also changes the degree of intensity in the meaning, but "whereas [adjective-adjective] reduplication confers and emphatic or vivid meaning, the form [adjective-adjective-*dei2*] serves to qualify or moderate the meaning of the adjective, like the English suffix *-ish*" (Matthews & Yip 1994: 164).

316.   Cantonese (Matthews & Yip 1994: 164)

ni1   zeung1 j2   waai6-waai5-dei2   ge3

this   CLF   chair   broken-broken-ish   PRT

"This chair is a bit broken"

A-B-B adjectives are a third type of adjectival reduplication which is less productive, and only occur in the reduplicated form. These often refer to adjectives of perception, like how things look or feel. As the name suggests, only the second component is repeated in this type of adjectival reduplication.

317.   Cantonese (Matthews & Yip 1994: 165)

chi1-lap6-lap6

glue-stick-stick

"sticky, gooey"

*chi1-lap6

glue-stick

Reduplicated adjectives can also form an adverb in Cantonese:

318.   Cantonese (Matthews & Yip 1994: 184)

ngo5	ming4-ming4	tai2	dou2	lei5	tung4	go3	leui5jan2	jat1cai4
1SG	clear-clear	see	V-PRT	2SG	with	CL	woman	together

"I clearly saw you with a woman"

Reduplication is not the only structure that results in adverb formation in Cantonese. One of the other ways is the postverbal particle *dak8*, which is "the most general adverbial construction" in Cantonese (Matthews and Yip 1994: 179).

319.   Cantonese (Matthews & Yip 1994: 179)

keui5	hok9	dak8	hou2	faai3
3SG	learn	ADV	very	fast

"He learns fast"

The other way is the usage of *gám* "in this way" after the adjective and before the verb.

320.   Cantonese (Matthews & Yip 1994: 181)

ngo5	m3hai6	jing4zan1	gam2	hok9
1SG	not-be	genuine	thus	learn

"I'm not studying (it) seriously"

Reduplicated classifiers are used to express quantification

321.   Cantonese (Matthews & Yip 1994: 266)

go3-go3	(jan4)	dou1	seung2	jiu3	do1	di1	mahn1jyu2
CLF-CLF	(person)	all	want	need	more	some	democracy

"Everyone wants a bit more democracy"

Note that it is used in combination with *dou1* "all".

Reduplicated verbs in Cantonese can form an adverb:

322.　　　Cantonese (Matthews & Yip 1994: 183)

lei5　tau1-tau1-dei2 zing2　je5　　sihk9　a4?

2SG　steal-steal-ish　make　things　eat　　PRT

"You've been secretly preparing food, have you?"

They can also have the meaning of a continuous or repetitive action, when followed by *hah5*.

323.　　　Cantonese (Matthews & Yip 1994: 208)

ngo5　lam2-lam2-hah5,　　dou1　hai6　m3hou2　　bun1　uk1.

1SG　think-think-DECL　also　is　　not-good　　move　house

"I've been thinking, it's best not to move house"

324.　　　Cantonese (Matthews & Yip 1994: 208)

go3　bì6bi1　sihk9-sihk9-ha5　　laai5　fan3-zeuk9-zo2

CLF　baby　eat-eat-DECL　　　milk　fall-asleep-PFV

"The baby fell asleep while drinking milk"

This example above shows how a reduplicated verb can be used to form the progressive, where the verb that denotes the interruption (*fan-jeuhk* "fall asleep") is not reduplicated while the verb that denotes the continuous action (*sikh-sikh* "eating") is.

Nouns referring to time in Cantonese make sentential adverbs with a habitual meaning.

325.　　　Cantonese (Matthews & Yip 1994: 266)

keui5　jat6-jat6　　heui3　jau3-seui2

3SG　day-day　　　go　swim-water

"She goes swimming every day"

Reduplication is also present in Malay and Indonesian, and can occur with many different grammatical categories.

Mintz (1994) mentions that nominal reduplication in Malay is often seen as having a pluralizing function, but this is not the case. Plain, non-reduplicated nouns in Malay can either be interpreted as a singular item or multiple items, and it is the context that indicates whether it is one or the other.

326.　　　Malay (Mintz 1994: 281)

Singapura	menjadi	pelabuhan	yang	utama	di	Asia	Tenggara
Singapore	become	port	REL	main	in	Asia	southeast

"Singapore became the main port in Southeast Asia"

327.　　　Malay (Mintz 1994: 281)

pedagang	dari	seluruhdunia	datang	ke	pelabuhan	Malaysia
trader	from	whole world	come	to	port	Malaysia

"Traders from all over the world came to Malaysian ports"

In these examples, the word *pelabuhan* "port" does not change, but the context indicates that the former is a single port, while the latter refers to multiple ports.

Mintz says that "the true function of noun reduplication is to show individuality within a group" (281). The example above in which *pelabuhan* refers to multiple ports within Malaysia implies that they are all treated as a unit. The function of reduplication is to show that they are not considered one unit, and it serves to show the individuality.

328.　　　Malay (Mintz 1994: 282)

pedagang	dari	seluruhdunia	datang	ke	pelabuhan-pelabuhan	Malaysia
trader	drom	whole world	come	to	port-port	Malaysia

"Traders from all over the world come to (the various different) Malaysian ports"

This example emphasizes the individuality of each Malaysian port, while the previous one does not do so.

Reduplicated nouns are generally not followed by words of quantity such as *banyak* "many" or *semua* "all". It is ungrammatical to follow a number with a reduplicated noun. Interrogative pronouns can be reduplicated in Malay, whereby they become indefinite. They may also be used within a negative sentence. See the following table, summarizing Mintz (1994: 289-90).

Table 18: Malay Interrogative Pronominal Reduplication

Malay phrase	English gloss	English translation
siapa	who	who
siapa-siapa	who-who	whoever, anyone
tidak siapa-siapa	NEG who-who	no one
apa	what	what
apa-apa	what-what	whatever, anything
tidak apa-apa	NEG what-what	nothing
mana	which/where	which/where
mana-mana	which-which/where-where	whichever, any which one/wherever, anywhere
tidak … mana-mana	NEG … which-which/where-where	none/no where
bila	when	when
bila-bila	when-when	whenever

These pronouns are used in the following way:

329.　　　Malay (Mintz 1994: 289)

mahu　minum apa?

want　drink　what

"What do you want to drink?"

apa　　apa　　boleh

what　what　can

"Anything will do"

saya　tidak　mahu　minum apa　　apa

1SG　NEG　want　drink　what　what

"I don't want to drink anything"

Adjectival reduplication in Malay and Indonesian can serve the purpose of showing emphasis or intensifying the meaning.

330.   Malay (Mintz 1994: 291)

teh   Cina   dihidangkan   panas/panas-panas   dan   dituangkan ke dalam

tea   China   PASS.serve   hot/hot-hot   and   PASS.pour in inside

mangkuk   kecil

cup   small

"Chinese tea is served hot/very hot and poured into small cups"

In Indonesian the reduplication of an adjective can also show individualizing within a group, in the same way nominal reduplication does.

331.   Indonesian (Mintz 1994: 292)

penduduk   asli   menenun   kain   sarung yang   cantik/cantik-cantik

inhabitant   original   weave   cloth   sarong REL   beautiful/beautiful-beautiful

"The original inhabitants weave beautiful sarongs/various types of sarongs that are beautiful"

The reduplication of *cantik* "beautiful" here serves the same purpose as the nominal reduplications above. Rather than conveying "very beautiful" it shows individualization of the different sorts of beautiful sarongs that are woven. The Malay version, according to Mintz, still has the meaning of intensification.

332.   Malay (Mintz 1994: 292)

rang   asli   menenun   kain   sarung yang   cantik/cantik-cantik

people original   weave   cloth sarong REL   beautiful/beautiful-beautiful

"The aborigines weave sarongs that are beautiful/very beautiful"

Reduplication is addressed in APICS Chapter 26: Functions of reduplication (Haspelmath 2013a). All creoles listed in APICS have iconic reduplication, in that it expresses "intensity, iteration, plurality, or distributivity". This has indeed been found in all four of the creoles in the present study, and I take a qualitative approach to looking at the various usages of reduplication.

6.8.2 Reduplication in Kristang

Reduplication is possible in a number of grammatical categories in Kristang. Kristang exhibits full and partial reduplication. As in Malay, a reduplication of a noun after a numeral is ungrammatical (Baxter 1988: 103).

333.   Kristang (Baxter 1988: 103)

*tres   krenkrensa

three   children

Baxter also points out that there is a connection between specificity and the reduplication of plurals. If a plural noun is specific, it will be reduplicated. Compare the following two examples:

334.   Kristang (Baxter 1988: 103)

eli   gostá ku   aké   krenkrensa

3SG   like   A   that   children

"She likes those children"

335.   Kristang (Baxter 1988: 103)

eli   gostá krensa

3SG   like   child

"She likes children"

He also demonstrates that in possessive constructions the reduplication of plural possessees is required "because the referent of the possessee is specific" (Baxter 1988: 103).

336.   Kristang (Baxter 1988: 103)

John   sa   kachoru

John   GEN   dog

"John's dog(*s)"

337.   Kristang (Baxter 1988: 103)

John   sa   kachoru kachoru

John   GEN   dog dog

"John's dogs"

When non-specific, reduplicated nouns can imply "different types of", "all sorts of" (Baxter 1988: 104).

338.   Kristang (Baxter 1988: 104)

yo   sa   sogru   gadrá   pastu

1SG   GEN   father-in-law   keep   bird

"My father-in-law keeps birds"

339.   Kristang (Baxter 1988: 104)

yo   sa   sogru   gadrá   pastu pastu

1SG   GEN   father-in-law   keep   bird bird

"My father-in-law keeps all kinds of birds"

This type of nominal reduplication resembles that of Indonesian and Malay, where it expresses individualization within a group.

Although nominal reduplication may not occur following a numeral in Kristang, it can be used with some quantifiers.

340.   Kristang (Pillai 2011)

yo   sa   kren-krensa   tudu   papiah Inggeris

1SG   GEN   children   all   speak English

"All my children speak English"

This has *tudu* "all" with the reduplicated form of *krensa* "child". It conveys the meaning of each and every child speaks English.

Adjectival reduplication can result in intensifying in Kristang (Baxter 1988: 107)

341.   Kristang (Baxter 1988: 107)

nus   sibrí   aké   pesi   kanikaninu

1PL   use   that   fish   small small

"We use the very small fish"

However, Baxter mentions that when an adjective is reduplicated to intensify, it only comes from adjectives related to dimension, as in *kaninu* "small" above.

Other classes of adjective may be reduplicated, but this results in the transformation into an adverb.

342.   Kristang (Baxter 1988: 107)

nu   muí   ku   eli   finu finu

1PL   grind   ACC   3SG   fine fine

"We grind it (the fish) very fine"

Verbal reduplication occurs on Kristang active verbs:

343.   Kristang (Baxter 1988: 60)

aké   má   ngua   yo   olá   ku   eli   ta   remá remá

that   more   one   1SG   see   ACC   3SG   PROG   row row

"That other one, I saw he was rowing and rowing"

Verbal reduplication has the meaning of a repetitive or durative action (Baxter 1988: 60).

6.8.3 Reduplication in Makista

Reduplication in Makista is attested to occur in various grammatical categories.

Nominal reduplication, as in the other creoles, can indicate a plural. It is optional and often a plural will be unmarked.

344.   Makista (Santos Ferreira 1967: 19)

rua-   rua   di   Macau

street street of   Macau

"The streets of Macau"

There are, however, examples of a numeral with a reduplication in the Makista corpora.

345.   Makista (Santos Ferreira 1973: 152)

cinco   nhu-nhúm   ta   repimpado   na   dentro

five   boys   PROG   seated   in   inside

"Five guys are seated inside"

346.   Makista (Pereira 1899: 190)

com   mas   dôs   ou   tres   sium-sium

with   more   two   or   three   gentleman-gentleman

"With two or three more gentlemen"

347.   Makista (Santos Ferreira 1973: 153)

siara   co   oito   filo-filo

lady   with   eight   son-son

"[He has] a wife and eight children"

The earlier corpora contain many instances of reduplication, especially those of nouns. For the examples from Santos Ferreiro, it must be noted that he likely used reduplication more often than would have actually been spoken at the time of his writing. Arana-Ward mentions that Makista speakers, who were Santos Ferreira's contemporaries, say that reduplication had gone out of use by the mid-20th Century. Additionally, the restriction on numerals before a

reduplicated noun in Kristang is a parallel with Malay restrictions. The Makista-speaking community would have had little opportunity in Macau to communicate with Malay speakers, so the usage of reduplication could have expanded as a result. Despite the instances of nominal reduplication with a numeral, they are extremely rare in both the earlier and later corpora and instances with numerals tend to be followed by non-reduplicated nouns. Although numerals before reduplicated nouns are rare, the determiner *tudo* "all" (alternatively spelled *tudu*) is found throughout the corpora:

348.   Makista (Pereira 1899: 131)

partí   pra   tudo   caza-caza
leave   from   all   house-house
"left from all their houses"

349.   Makista (Pereira 1899: 324)

saude   para   vôs,   marido   e   tudo   criança criança
health   for   2PL   husband   and   all   child-child
"health for you, your husband, and all of your children"

350.   Makista (Barreiros 1944: 241)

Torná   jâ   espalá   pra   tudo   caza-caza   de   Macáo
turn   PFV   spread   to   all   house-house   of   Macau
"[the news] spread to all the houses of Macau"

In certain contexts a reduplicated noun becomes an adjective, and are an example of derivational morphology.

351.   Makista (Pereira 1899: 60)

cô   este   calôr   cérto   lô   vên   sópa-sopa
with   this   heat   certain FI   come   soup-soup
"with this head you will certainly become soaking wet"

Santos-Ferreira (1978: 315) gives an example of *mulado sópa-sópa* "wet like soup", however we can see in the above example that it can function on its own as an adjective. Another common example is *laia-laia*. While *laia* on its own is "type, manner", when reduplicated it takes the meaning of "all sorts of".

352.   Makista (Pereira 1899: 58)

trezê   cusa   laia-laia
bring   thing   type-type
"bring all sorts of things"

353.   Makista (Barreiros 1944: 33)

ung-a	botle	de	achar	laia-laia
a	bottle	of	pickles	type-type

"a bottle of all sorts of pickles"

Adjectival reduplication can function as an intensifier.

354.   Makista (Santos Ferreira 1985: 41)

sã	êle	acunga	dia	cedo-cedo	vêm
COP	3SG	one	day	early-early	come

"It's he who came very early that day"

355.   Makista (Barreiros 1944: 33)

mas	cham	pra	fazê	palacio	grande-grande	pra	oficial
more	land	for	make	palace	big-big	for	official

"more land to build large palaces for the officials"

356.   Makista (Pereira 1899: 780)

oze	pramicedo,	sentado	omsong-omsong
today	early in the morning	seated	alone-alone

"Early in the morning today, seated all alone"

Adjectival reduplication can also serve as derivational morphology, not unlike those of the examples above. The following examples feature a reduplicated adjective *pobre* "poor" used as a noun:

357.   Makista (Pereira 1899: 259)

pôde	dá	pra	pobre pobre
can	give	to	poor-poor

"could give to the poor"

358.   Makista (Barreiros 1899: 32)

tudo	pobre-pobre	vae	pará	pra	casinha de campo
all	poor-poor	go	stop	by	cottage

"all the poor people went to stop by the cottages"

Verbal reduplication is less common than other categories. It is used to describe a repetitive action.

359. Makista (Avram 2015)

pê   mâm   tremê-tremê

leg   hand   tremble-tremble

"His legs and hands were trembling"

360. Makista (Santos Ferreira 1985: 18)

testa   suado- mulado   coraçám   batê-batê

forehead   sweaty wet   heart   beat-beat

"His forehead was wet with sweat, his heart beating"

361. Makista (Barreiros 1944: 357)

tem   tantu dinhero,   pra   sentá-sentá

have   much   money   for   sit-sit

"You have so much money, so you can just sit around"

Another usage of verbal reduplication indicates a gentle way to make a request:

362. Makista (Barreiros 1944: 468)

virá-virá   vem   dá   unga escuta pa   iou

turn-turn   come give   one   listen to   1SG

"come by when you can to give me a listen"

Santos-Ferreira (1978: 324) describes the construction *virá-virá vem* as a request to "appear from time to time" or to "come when you can"[11].

363. Makista (Barreiros 1944: 249)

dá   quanto pancada   pra   quem pôde ôvi-ôvi

give   some   strike   for   who   can   hear-hear

"give some strikes [of a bell] for those who can hear"

The number of examples of verbal reduplication in Makista is low compared to nominal and adjectival, so it is difficult to conclude whether it has any productivity beyond the sense of repetitive actions.

Numeral reduplication is also attested in Makista

364. Makista (Pereira 1899: 125)

pra   comprá   unga-unga   ancusa

for   buy   one-one   thing

"to buy things one by one"

---

[11] "apareça de vez em quando ou venha quando puder" in the original.

365.   Makista (Santos Ferreira 1996: 86)

Êle   largá pê   pulá dôs-dôs   degrau

3SG   leave leg   jump two-two   step

"He left jumping two steps at a time"

Numeral reduplication in Makista indicates how items are grouped.

6.8.4 Reduplication in Batavia Creole

The BC texts contain some reduplication, mostly nominal, with one example each of verbal and adjectival reduplication. All of the examples of nominal reduplication in BC express a plural.

366.   Batavia Creole (Maurer 2011: 130)

esta   albër albër ki   teng por   lumia

this   tree-tree   COMP COP   PURP list

"The following trees, which are going to be listed…"

367.   Batavia Creole (Maurer 2011: 133)

e   otër   sorti di   albër fula fula

and   other sort   of   tree   flower-flower

"And other sorts of flower-bearing trees"

Note in this example that the repition of *fula* "flower" here implies there are multiple flowers in the trees. Although *albër* "tree" is also implied as multiple, it is not reduplicated. There are no examples of two reduplicated words in a row in BC.

The BC texts have one example of verbal reduplication.

368.   Batavia Creole (Maurer 2011: 129)

ile   ola ola tudu lay   sorti di   fula

3SG   see-see all   kind   sort   of   flower

"He saw all kinds of flowers"

In this case the verb in its reduplicated form has the same grammatical category. In contrast to the Chinese examples above, in which reduplication implies that it is for a short while, this one has a meaning of "looking around".

There is one example of an adjectival reduplication in BC.

369.   Batavia Creole (Maurer 2011: 137)

dretu dretu   sua   predju kantu?

Right-right   GEN   price   how.much

"What is your best price?"

In this example the reduplicated adjective has the same grammatical category as the non-reduplicated form. It also serves as an intensifier, by asking which is the best price.

6.8.5 Reduplication in Tugu Creole

Reduplication is also present in the TC texts, to a larger extent than in BC. Most of the tokens are nominal reduplication, but there are also examples of verbal, adjectival, numeral, and pronominal reduplication.

Most examples of nominal reduplication in TC express a plural.

370.   Tugu Creole (Maurer 2011: 193)

filu filu   e   fila fila   prende na   skola   dominggu

boy-boy   and   girl-girl   learn   in   school Sunday

"The boys and girls study in Sunday school"

In this case the reduplication of the nouns simply reflects the fact that there are multiple boys and girls. Their grammatical category is the same and their position in the sentence is the same as they would be if they were not reduplicated.

371.   Tugu Creole (Maurer 2011: 145)

perbidapos pos   tantu ki   su   agu   salgadu

because   well-well   many   that   GEN   water   salty

"Because there are many wells whose water is salty"

This example also shows how there can often be a classifier in addition to the reduplication of the noun, expressing how much there are.

There is one example where the reduplication of the noun does not express a plural.

372.   Tugu Creole (Maurer 2011: 173)

io   kere   papia ung   taninu ung   taninu

1SG   want   speak a   bit   a   bit

"I want to speak a little bit"

In this case, the article and noun are reduplicated in full. As the article implies that there is only one, the reduplication in this case is not a plural, but rather a diminutive for "a little bit". This could also be considered an intensifier, emphasizing the "littleness" of "little bit".

Of the four examples of verbal reduplication in TC, three are reduplication of the verb *lembra* "remember".

373.   Tugu Creole (Maurer 2011: 145)

tera   kampong   Tugu   ting   lembra lembra ung senta trinta seti djenti

land   village   Tugu   COP   think-think   one hundred thirty-seven people

"In the village of Tugu is more or less 137 people"

374.   Tugu Creole (Maurer 2011: 149)

sua   londji unga   otër   lembra lembra unga   pio   sua   londji

GEN   length one   other   think-think   one   foot   GEN   length

"The distance between them is more or less one foot"

375.   Tugu Creole (Maurer 2011: 156)

sumenti   lembra lembra korenta   dia   nos   rangka   mara

seed   think-think   forty   days   1PL   tear.out   bind

"The seeds, after more or less forty days, we tear them out and bind them together"

This is a reduplication of the verb *lembra* "think", but this is a form of derivational morphology, in which the verbal root when reduplicated becomes an adverb in this case. The Malay translation in the text has a similar instance of reduplication.

376.   Malay (Maurer 2011: 156)

bibit   kira kira   ampat poeloe hari   kita   tjaboet   ikat

seedling   guess-guess   forty   day   1PL   remove   bind

"The seeds, after more or less forty days, we tear them out and bind them together"

This has the exact same word order as the TC example above. In addition, the word *kira* on its own in Malay is a verb "to guess", while when reduplicated has the same effect as an adverb, just like in TC.

The other example of verbal reduplication in TC is more inflectional, as it has an effect on the aspect of the verbs.

377.   Tugu Creole (Maurer 2011:179)

bi   pasa pasa   na   bentu

come   go-go   in   wind

"Let's go walking in the fresh air"

This reduplication of *pasa* "to go" does not change the grammatical category of the non-reduplicated verb, as it is still a verb in this case. Its reduplication affects the telicity of the verb, as it does not imply any certain goal or endpoint to the walking. This is also a common phrase in Malay, and can be found in the Malay translation in this text.

378.	Malay (Maurer 2011: 179)

mari	djalan djalan	di	agin

come	walk-walk	in	wind

"Let's go walking in the fresh air"

Again, this Malay translation has the exact same word order as the TC example above, including the verbal reduplication.

Adjectival reduplication occurs more often in TC than it does in BC. While the BC example serves as an intensifier for the adjective, the TC examples are more derivational, in particular with the reduplicated adjective being used as an adverb. There is also an ambiguous example of reduplication in TC:

379.	Tugu Creole (Maurer 2011: 193)

iste	dia	chua	pichi pichi

this	day	rain	drizzle drizzle

"Today it's drizzling"

It is unclear from the translation in the text itself whether *pichi* is the noun or the verb "drizzle", although the glossary refers to it as a verb. I argue that this example can be considered an example of adjectival reduplication serving the purpose of both being an adverb and intensifying. The word *chua* "rain" in these examples is a noun. Another example sentence has *chua* in a situation where being a noun is clearer.

380.	Tugu Creole (Maurer 2011: 162)

sertedja	teng	chua	djarang djarang

certainly	COP	rain	rare rare

"Certainly it rarely rains"

A more literal translation would be "Certainly there is rarely rain", with the reduplicated adjective *djarang* "rare" modifying the copular predicate *teng chua*.

Another example shows a similar structure, although without the copula *teng*.

381.	Tugu Creole (Maurer 2011: 193)

iste	anoti	chua	kabrola

this	night	rain	hard

"Tonight there's a hard rain"

In this case the adjective *kabrola* "hard" modifies the noun *chua* "rain". With this in mind it calls the translation of the sentence above ("Today it's drizzling") into question, at the very least the gloss of *pichi* as "drizzle", a verb, in a position that is occupied by adjectives in similar sentences.

Adjectival reduplication is also present in TC.

382.	Tugu Creole (Maurer 2011: 163)

podi	nos	acha	ponta	pertu	pertu	kal	grandi

can	1PL	get	shoot	close	close	REL	big

"We can shoot big ones from very close"

383.	Tugu Creole (Maurer 2011:171)

para	obida	bong	bong

for	listen	good	good

"Listen very attentively"

These examples of adjectival reduplication indicate an intensification of the meanings of the adjectives.

There are three examples of numeral reduplication in TC.

384.	Tugu Creole (Maurer 2011: 149)

ki	ake	albër	neli	ditara	unga	unga

REL	here	tree	rice	PASS.plant	one	one

"Here the rice plants are planted one by one"

385.	Tugu Creole (Maurer 2011: 154)

kaba	djenti	barisin	singku	singku	arma

finish	people	line.up	five	five	arrange

"Then one lines it up in rows of five, and arranges it"

386.	Tugu Creole (Maurer 2011: 177)

tudu	dos	dos	omi	muler

All	two	two	male	female

"All in pairs of males and females"

This numeral reduplication indicates how objects are grouped.

There are two examples of pronominal reduplication in TC.

387.   Tugu Creole (Maurer 2011: 175)

irimang   bi   komi   aros,   tapi   nute   ki   ki   dangki

brother come   eat   rice   but   NEG.EXIST   what   what   thank

"My brother, come and eat rice, but there is nothing to thank for"

388.   Tugu Creole (Maurer 2011: 193)

yo   nunte   doy,   nunte   ki   ki

1SG   NEG.HAVE   money NEG.HAVE   what   what

"I don't have any money, I don't have anything"

These examples of pronominal reduplication are very similar to that used in Malay. When not reduplicated, *ki* has the maning of "what" and can be used as a question word. When reduplicated, it takes on the meaning of "whatever" or "anything". These examples both use different negative particles, but the word order is the same as in Malay, in which the negative particle precedes the reduplicated pronoun.

6.8.6 Summary and Discussion

The types of reduplication in Kristang, Makista, BC, and TC are Nominal, Verbal, Adjectival, Numeral, and Pronominal.

Table 19: Summary of Reduplication

	Kristang	**Makista**	**BC**	**TC**
nominal	pluralization/types of things	pluralization	pluralization	pluralization/types of things
verbal	repeated action	repeated action	duration	duration
adjectival	intensification, adverb	intensification	intensification	intensification
numeral	---	grouping	---	rouping
pronominal	---	---	---	indefinite

Reduplication is a common feature throughout the world's languages, not just in creoles, so the mere presence of them does not indicate anything related to its potential relationship with a language ecology. However, we do see many parallels between the patterns of Malay and Cantonese reduplication that could have been an influence on the reduplication patterns in the four creoles.

Reduplication in these languages, as well as in the substrates, is a morphological, rather than a phonological, process. The act of reduplicating a word, in part or in whole, has a productive effect on the meaning or the function of the word itself, and is not just a phonological feature. For one, a reduplication of a nominal is, at its most basic definition, an optional pluralizer. However, there are many more similarities with Malay beyond just being the indication of a plural. As Mintz mentions, it is an individualizing effect, that is, it points out that each item belongs to a set, but on its own. We can see in Kristang examples where reduplication has the same effect. The reduplication of *pastu* "bird" indicates that there are all kinds of birds, not just an indefinite amount of birds. Another parallel is the prohibition of reduplication of nouns after a numeral. Nouns may be reduplicated after a quantifier, but not after a numeral in either Malay or Kristang.

The analysis of this feature is similar to that of serial verb constructions, in that it is a wide category that occurs cross-linguistically, so a look into the patterns of usage are what indicate the influence from input languages, rather than its presence alone. Reduplication patterns in Kristang particularly show similarities with Malay, where a doubling of a noun indicates different types of an object, rather than just plural. Another important similarity is the prohibition of reduplication after a numeral, which has occurred in the Makista data.

6.9 Constituent Order and Object Marking

6.9.1 Introduction

This section looks at the issues of constituent order within a sentence in Kristang, Makista, TC and BC. As these languages are more analytic, word order can play an important role in the conveyance of meaning. Information structure also plays a role in this section. Important concepts in information structure include **focus vs background, topic vs comment, and given vs new** (Zimmerman & Féry 2009). Information structure can explain how a language will express something that is new to the listener, or something that is already known to both speaker and listener, either through the context of the conversation or through common knowledge. It also is utilized in situations where one wants to contrast one thing with another, and the way they express this will be different from a standard, non-contrasting, utterance.

In English, this is often expressed with word order. Consider the sentence "I like blueberries, but oranges I don't like." Although the standard word order of English is SVO, in this contrasting example the object is moved to the front of the clause.

The four creoles in the present study are all more analytic than synthetic, and because of this their information structure is more to do with word order than with morphology. This

section will only look at how word order works, whether by information structure or other means. The previous Section 6.3 also deals with information structure in its discussion about cleft constructions. Word order is also important in Malay grammar. In standard Malay, there is no morphology to express possession, it is rather expressed with word order (bare genitive).

This is addressed in APICS Chapter 61: Order of recipient and theme in ditransitive constructions (Haspelmath 2013b). Kristang (Baxter 2013) is shown to have both subject-verb-recipient-theme order and subject-verb-theme-recipient order, with the former likely having a higher frequency. Rather than the order that influences markings of the recipient and theme, the usage of prepositions tends to be used, as will be seen in the following sections.

6.9.2 Constituent Order and Object Marking in Kristang

In Kristang, adjectives follow the noun which they are modifying, in the same way as in both Portuguese and Malay.

Fronting is used in Kristang to put a constituent into focus.

389.   Kristang (Pillai 2011)

ki	mas	bos	papiah?
what	more	2SG	speak

"What else do you speak?"

Malayu,	Inggeris.	Yo	ngkah	papiah	Cina,	Moru,	Malayu
Malay	English	1SG	NEG	speak	Chinese	Tamil	Malay

ngkah	ah	Malayu	yo	papiah
NEG	PRT	Malay	1SG	speak

"Malay, English. I don't speak Chinese, Tamil, Malay, no ah Malay I speak"

A prepositional phrase in Kristang that modifies the verb can come in between the subject and the verb.

390.   Kristang (Pillai 2011)

yo   papiah Kristang   ku   yo   sa   kren-krensa,   kauzu   yo

1SG   speak Kristang   to   1SG   GEN   children   because   1SG

di   keninu papiah Kristang

from   small   speak   Kristang

"I speak Kristang to my children, because I've been speaking it since I was a child"

This appears to occur when the modifying phrase is related to tense or aspect.

391.   Kristang (Pillai 2011)

yo   kustumi   papiah Kristang

1SG   be.used.to   speak   Kristang

"I am used to speaking Kristang"

This type of modification can also occur after the verb.

392.   Kristang (Pillai 2011)

beng   nakih   papiah kada   dia

come   here   talk   every   day

"I come here to talk every day"

393.   Kristang (Pillai 2011)

so   nus   botah padi   anu   per   anu

so   1PL   put   for   year   by   year

"So we put [the boat here] every year"

In this case, the modifying phrase comes after the verb, and there is nothing intervening between the subject and the verb (when there is an overt subject).

Bruyn, Muysken, and Verrips (1999) identify two ways to express a verb phrase with two objects, or a dative verb. The first is the prepositional dative construction (PDC). This exists in English in the sentence *Jane gave the book to Mary*, with the indirect object, *Mary* being in a prepositional phrase. The second is double-object construction (DOC). This is also present in English, as in *Jane gave Mary the book*, with the indirect object *Mary* being stated before the direct object *book*. They noted that the romance languages that were major lexifiers for Atlantic and Asian creoles, Spanish, French, and Portuguese, there are no DOCs, but only PDCs. However, many of the romance-lexified creoles have DOCs.

394.   Haitian (Bruyn, Muysken, and Verrips 1999: 330)

li   rakonte   papa- li   istwa sa-a

3SG   tell   father 3SG   story   this

"He told his father this story"

395.   Papiamentu (Bruyn, Muysken, and Verrips 1999: 330)

bo   a   duna   mi   e   buki

2SG   ASP   give   1SG DET   book

"You have given me the book"

396.   Guinea-Bissau Kriyol (Kihm 1994: 54-56, as cited in Bruyn, Muysken, and Verrips 1999: 336)

mininu   manda si   mame un   karta

boy   sent   3POS mother ART   letter

"The boy sent his mother a letter"

397.   Principense (Günther 1973: 87, as cited in Bruyn, Muysken, and Verrips 1999: 336)

pw´ɛ   sa   dá   mínu dyó

father ASP give   child   money

"The father gives the child money"

They note that the only group of romance-lexified creoles that does not have DOCs is the group of Malaysian and Indonesian creoles.

      Kristang uses *ku* "with" to mark the indirect object in a dative verb phrase. This is also used to mark the accusative for human objects.

398.   Kristang (Pillai 2011)

dos   yo   dah   ku   bos

two   1SG   give   with   2SG

"I gave you two [cards]"

399.   Kristang (Pillai 2011)

eli   ja   dah   isti   libru   per   nus   jenti di   Melaka

3SG PF   give   this   book   to   1PL   people of   Malacca

"He gave this language to us, the people here in Malacca"

400. Kristang (Pillai 2011)

porkih	kauza	akeh	mai	papiah	ku	eli	sa	familia	Portugis	di
Because		this	mother	speak	with	3SG	GEN	child	Portuguese	of

Melaka
Malacca

"Because the mother speaks Malacca Portuguese to her children"

401. Kristang (Pillai 2011)

isti	yo	falah	ku	siara	siuris
this	1SG	speak	with	lady	gentleman

"This is what I say to you, ladies and gentlemen"

402. Kristang (Pillai 2011)

eli	logu	da	ku	nus	mpoku	doi	dos	mil
3SG	PFV	give	with	1PL	a little	money	two	thousand

"They [the government] will give us a little bit of money, about two thousand"

403. Kristang (Pillai 2011)

yo	papiah	Kristang	ku	yo	sa	kren-krensa
1SG	speak	Kristang	with	1SG	GEN	children

"I speak Kristang to my children"

404. Kristang (Pillai 2011)

ah	yo	misti	papiah	yo	sa	krensa	Kristang
ah	1SG	must	speak	1SG	GEN	child	Kristang

"Ah, I must speak Kristang to my children"

405. Kristang (Pillai 2011)

otru	ngka	papiah	Kristang
other	NEG	speak	Kristang

"With others they don't speak Kristang"

406. Kristang (Pillai 2011)

eli	nggeh	olah	ku	eli	sa	mai
3SG	NEG	see	with	3SG	GEN	mother

"He doesn't want to see his mother"

407.   Kristang (Pillai 2011)

yo	sa	kambradu	ja	chomah	ku	yo
1SG	GEN	friend	PFV	call	with	1SG

"My friend called me"

With verbs of direction, a preposition or marker is not used:

408.   Kristang (Pillai 2011)

jenti	di	ta	bai	mar	sa	jenti
people	of	PROG	go	sea	GEN	people

"People that go to the sea"

409.   Kristang (Pillai)

yo	tokah	beng	kaza	lah
1SG	touch	come	house	*lah*

"I have to come back home"

In the spoken corpus of Kristang there are many instances of datives, with different expressions. In the above examples we have the following ways of expressing dative:

a. DO S V *ku* IO
b. S V DO *per* IO
c. S V *ku* IO DO
d. S V DO *ku* IO
e. S V IO DO
f. IO V DO

These examples show a large variety of expressing the dative. While we do have examples with a prepositional phrase, there are some without. The ones without prepositions in this selection always end with the direct object, however. There appears to be a large variety of word order as well, especially when we have a prepositional phrase. This normally begins with *ku* but there is at least one instance that begins with *per*. Semantically, these verbs do differ from each other. We have *da* "give", but we also have *papiah* "speak". This does not seem to have an effect on the word order within dative constructions.

There is also evidence that in certain cases, non-human animates may be marked with *ku*. In the "animal birthday party" activity (see Section 5.4) participants were given picture cards of animals and gifts to give to them. In this case *ku* is used to mark an animal in the same way it can mark a human object.

410.		Kristang (Laub fieldwork, 2017)

ku		kuelu	bedri,	bo	ki	ke		dah?

to		rabbit	green	2SG	what	want	give

"What do you want to give to the green rabbit?"

In this situation *kuelu* is marked the same way as a human object would be.

411.		Kristang (Laub fieldwork, 2017)

yo			dah	lus	ku	bos

1SG			give	light	to	2SG

"I give a light[bulb] to you"

A reason that *kuelu* is marked with *ku* is that the situation presented is treating the animals as if they are human, in a birthday party. Aissen (2003) notes that in differential object marking there is a hierarchy of animacy: human > animate > inanimate. This also explains why *kuelu* can be marked in this situation. The imaginary situation itself deals with treating animals the way humans are treated, and animates are only one step below humans in the animacy hierarchy. Compare this with inanimates:

412.		Kristang (Laub fieldwork, 2017)

e		ta	dali	kosi	dah	kadera

3SG		PROG	hit	kick	give	chair

"He is kicking the chair"

413.		Kristang (Laub fieldwork, 2017

subi	martelu		dali	dah	kareta

Lift	hammer		hit	give	car

"[She is] hitting the car with a hammer"

In these examples the objects are inanimates: a chair and a car. The objects are also in the focus of the answer. He is kicking a chair, not the other furniture. In this case the verb *dah* "give" is used instead of *ku* which would be used when the object is human. Looking at Aissen's hierarchy, there are two steps between inanimate and human, so this is less likely to use *ku* to mark the object.

6.9.3 Constituent Order and Object Marking in Makista

Makista makes use of a dative marker *pa* and its variants, which are synonymous to Kristang *per*.

414.	Makista (Santos Ferreira 197: 152)

fazê	favô	pa	quim	pedí

Do	favor	for	who	ask

"[He] does favors for those who ask"

415.	Makista (Santos Ferreira 1973: 155)

dá	órdi	pa	Dios

Give	order	to	God

"Give orders to God"

416.	Makista (Barreiros 1944: 357)

já	dá	quatru	naviu	pr'	elle

PFV	give	four	ship	to	3SG

"…gave four ships to him"

417.	Makista (Pereira 1899: 57)

elle	preguntâ	pâ	Mai

3SG	ask	to	mother

"she asked her mother"

There is also usage of *com* and its variants (synonymous to Kristang *ku*):

418.	Makista (Barreiros 1944: 129)

eu	amestê	pedi	ung-a	cuza	com	vôs

1SG	must	ask	one	thing	with	2SG

"I must ask you something"

419.	Makista (Barreiros 1944: 358)

n'.aquelle	tempo	tudu	portuguez-portuguez	custumado	faze

guéra

in.that	time	all	Portuguese-Portuguese	used to	make	war

co	môro-môro

with	Moor-Moor

"In that time all of the Portuguese were used to making war with the Moors"

420.	Makista (Barreiros 1944: 468)

iou	tem	unga	historia	muito	cumprido	pa	contá	com	vôs

1SG	have	one	story	very	long	for	tell	with

2SG

"I have a very long story to tell you"

The use of *com* as a dative marker in Makista is not as widespread as in Kristang. As seen in the examples above, they were only found in the Barreiros corpus, which was from an earlier source than Santos-Ferreira's work.

Pa can also be used as a preposition of movement:

421.　　　Makista (Santos Ferreira 1973:155)

passá　pa　　vánda　di　　　trás
Pass　　to　　alley　of　　　back

"[They] went to the alley in the back"

422.　　　Makista (Santos Ferreira 1973: 157)

Fred　ta　　vai　　pa　　　fora
Fred　PROG　go　　to　　　outside

"Fred was going outside"

423.　　　Makista (Pereira 1899: 193)

fuzí　　tudo　azinha azinha　pra　　Hongkong
flee　　all　　quick-quick　　to　　Hong Kong

"quickly flee to Hong Kong"

424.　　　Makista (Barreiros 1944: 481)

afinal　eu.ça　　　　Fulgencio　já　　parti　pa　　Cantan
finally　1SG.GEN　　Fulgencio　PFV　leave　for　Guangzhou

"finally my Fulgencio left for Guangzhou"

As in Kristang, there are also instances of null marking for verbs of movement:

425.　　　Makista (Barreiros 1944: 475)

vós　pode　vai　Ongkong　ô　　Cantan,　　　mas　　mamã mas
gostá
2SG　can　　go　　Hong Kong　or　　Guangzhou　but　　mother more
want
vôs　vai　Cantan　　ficá　cu　　tio　　Tone
2SG　go　　Guangzhou　stay　with　uncle　Tone

"You can go to Hong Kong or Guangzhou, but mother prefers that you go to Guangzhou to stay with Uncle Tone."

The presence of *pa* for verbs of movement is optional as can be seen in the following example (null object marking and preposition highlighted):

426.   Makista (Barreiros 1944: 358)

Purtuguez	já	vem	Ø	Macau na		1557.	Cavá	non	tem
niunga									
Portuguese	PFV	come		Macau in		1557	then	NEG	have
none									

ôtro	naçam	qui	já	vem	**pra**	China	fazê	negócio,		cuza	di
other	nation	REL	PFV	come	to	China	make	business		thing	of

80	annu
80	years

"The Portuguese came to Macau in 1557. After that there were no nations which came to China to conduct business for 80 years."

This example illustrates that the variation between marking an indirect object for *vai* "go" or not would not have varied person-to-person, as we can see that examples of both are found in the same sentence.

Pa can also be used optionally as a marker after the verb *olá* "to see, to watch"

427.   Makista (Santos Ferreira 1973: 166)

estunga	demónio	falá	qui	lôgo	vêm	olá	pa	iou
this	devil	say	REL	PFV	come	watch	to	1SG

"This devil said that he would come watch me"

428.   Makista (Santos Ferreira 1973: 166)

vosôtrota	bêm	di	ispantado	olá	iou	aqui
2PL	PROG well	of	startled	watch	1SG	here

"You really startled me by seeing me here"

We must note that the examples used from Makista here are from a written source, while the Kristang data come from spoken recordings. In spoken language, we are more likely to see shifting of word order to change the focus of the utterance.

Makista has examples of null direct object marking for humans:

429.   Makista (Pereira 1899: 125)

já	cavá	tempo	de	Eva	enganá	Adão,	agora	sã	macho macho
PFV	finish	time	of	Eve	fool	Adam	now	COP	man-man

que	enganá	femea femea
REL	fool	woman-woman

"The time that Eve deceived Adam is in the past, now it's men who deceive women"

The above examples show the relative lack of diversity of dative constructions in Makista than what we find in Kristang. Kristang uses two different prepositions to indicate an indirect object, while Makista heavily favors one of them, with the other not found in the 20th Century corpus. Also, there are Kristang examples where the word order indicates which is the direct, and which is the indirect object, and there is no preposition being used to indicate the indirect object. In Makista, there are no such examples in the Santos Ferreira data.

What is notable is the usage of *ku* in Kristang to mark an indirect object, as it is also used to mark a direct object. This is limited to human objects, but there is also an example of *per nus* "to us", which is also a human object but used with a different preposition. The examples in Makista with *pa* have a human indirect object, and an animate one, *Dios* "God". We do find *pa* being used as a directional preposition, and in this case it is used with non-human objects. This appears to be a general dative marker, whereas in Kristang there is more diverse usage of prepositions.

6.9.4 Constituent Order and Object Marking in Batavia Creole

BC and TC exhibit datives that resemble Kristang in structure. The object marker, related to *kung*, is used to mark human direct and indirect objects.

430.   Batavia Creole (Maurer 2011: 131-2)

isti	belu	da	sabe	kung	ile	ki	esta	teng	lugar…
DEM	old.man	give	know	OBJ	3SG	COMP	DEM	COP	place

"The old man told him that this was the place where…"

431.   Batavia Creole (Maurer 2011: 142)

eo	teng	amor	keng	da	amor	kung	eo
1SG	have	love	REL	give	love	OBJ	1SG

"I love the person who gives love to me"

432.   Batavia Creole (Maurer 2011: 144)

fala   kung   ile   ki   eo   dja   teng   aki

tell   OBJ   3SG   COMP 1SG   PFV   COP   here

"Tell him that I have been here"

In these examples the word order differs between stating the direct object or the indirect object first. It could be due to the semantics of the verb, as *fala* "tell" is followed by *kung* IO, then a relative clause, whereas *da* "give" is followed by DO *kung* IO.

Many ditransitive clauses in BC leave the indirect object unstated, and understood from the context.

433.   Batavia Creole (Maurer 2011: 140)

da   otër   piring

Give   other   plate

"Give [me] another plate"

434.   Batavia Creole (Maurer 2011: 140)

eo   lo   manda minya kareta

1SG   FUT   send   my   carriage

"I will send [him] my carriage"

435.   Batavia Creole (Maurer 2011: 142)

obi   ung   bes.   Eo   lo   konta alung kudja.

Listen one   time   1SG   FUT   tell   some   thing

"Listen here, I'll tell [you] something"

BC datives are expressed in the following ways:

    a. S V *ku* IO DO
    b. S V DO *ku* IO

6.9.5 Constituent Order and Object Marking in Tugu Creole

436.   Tugu Creole (Maurer 2011: 194)

baklay   namas

fight   only

"You do nothing but fight"

437.	Tugu Creole (Maurer 2011: 194)

bos	nang	rikadu	namas,	koma	djenti	mal
2SG	NEG	scold	only	like	person	bad

"Don't just scold, like a bad person"

438.	Tugu Creole (Maurer 2011: 146)

bobernemen	dja	da	ku	nos	unga	gredja	per	konserta	doti
government	PFV	give	OBJ	1PL	one	church	PURP	exert	?

"The government gave us a church to practice our religion"

439.	Tugu Creole (Maurer 2011: 146)

unga	mestër	indjil	ki	prende	djenti	kampong	mas	di	prendesu
one	teacher	gospel	REL	teach	people	village	more	of	religion

"A gospel teacher who teaches the people of the village more of the religion"

TC datives are expressed in the following ways:

a.	S V *ku* IO DO
b.	S V IO DO

6.9.6 Summary and Discussion

The following table shows the patterns of marking of the direct and indirect objects (Adapted from Maurer 2011: 80)

			direct object	**indirect object**
pronouns	1SG	Kristang	---	ku
		Makista	pa / Ø	---
		BC	kung	kung
		TC	parmi	koyo, parmi, por yo
	2SG	Kristang	ku	ku
		Makista	---	---
		BC	per	---
		TC	---	---
	3SG	Kristang	ku	ku
		Makista	---	---
		BC	kung	kung
		TC	ka ~ ku, per / Ø	ka ~ ku / Ø, ku ele
	1PL	Kristang	---	per / ku
		Makista	---	---
		BC	---	---
		TC	ka / Ø	ku
	2PL	Kristang	---	---
		Makista	---	---
		BC	---	---
		TC	---	---
	3PL	Kristang	---	---
		Makista	---	---
		BC	---	---
		TC	---	---
nouns	humans	Kristang	---	ku / Ø
		Makista	---	pa
		BC	kung / Ø	---
		TC	ka / Ø	ka ~ ku / Ø
	animates	Kristang	---	---
		Makista	---	pa
		BC	Ø	---
		TC	Ø	Ø
	inanimates	Kristang	---	---
		Makista	---	---
		BC	Ø	---
		TC	Ø	---

In the case of object marking, all four varieties generally follow the same patterns. All of the languages use *per* and *ku* (or a cognate), and they are used to mark human objects (the only instances where the object is not strictly human is *Dios* "God" in Makista, and *kuelu* "rabbit" where the wider context was a birthday party for animals). This is an instance where there is little difference between the four varieties with a particular feature. It appears to have

been an internal development within Malacca creole before the split, as it is present in all four and the patterns of usage resemble each other.

6.10 Summary of Features

The following table is a summary of the features found in this chapter.

Table 21: Summary of Features in Chapter 6

Language	Kristang	Makista	BC	TC
serial verb constructions	resent	present	present	present
cleft constructions	*teng…ki* reminiscent of Portuguese	sinitic-influenced cleft with copula	---	relativizer *ki* or Ø
short/long pronouns	present, likely variation due to information structure	absent	absent	absent
verbal ellipsis after TAM marker	grammatical after *ja, lo,* ungrammatical after *ta*	grammatical after *ja, lo,* no examples of *ta* fragment	possibly after aux verb, unknown after TAM	possibly after aux verb, unknown after TAM
genitive	balanced patterns of DI-type and SA-type in semantic categories	DI-type much preferred	DI-type much preferred	third structure, bare genitive, from Malay influence
passive	bare passive; usage of *tokah* as adversative passive	usage of Portuguese verbal suffix *-ado*	---	usage of Malay-derived verbal prefix *di-*
reduplication	nominal, verbal, adjectival	nominal, verbal, adjectival, numeral	nominal, verbal, and adjectival	nominal, verbal, adjectival, numeral, and pronominal
constituent order and object marking	*Per* and *ku* for indirect and human direct objects; focus fronting	*Pa* or Ø marking object	*Kung*, *per*, or Ø marking object	*Par, per,* or *ku* marking object

Part Three: Discussion and Conclusion

Chapter 7: Creole Exceptionalism and Language Ecology

7.1 Introduction

This chapter revisits the debate regarding creole exceptionalism, and if creoles exhibit features due to rapid contact-induced change (McWhorter 2005, Bakker et al 2011) or due to the languages from which they descend and their language ecology (Mufwene 2002, DeGraff 2003, Ansaldo 2009). As I have shown in the previous chapter, differences between Kristang, Makista, BC, and TC can be attributed to the various languages the speakers would have come into contact with.

I also explain how much of creole literature looks at comparisons with the substrate (Muysken & Smith 1986), but an extensive undertaking of comparing creoles with each other should also be considered when looking at the origins of creole structure. When we look at creoles that happen to be related, such as the four in this thesis, we can attribute their similarities to be those features that have not changed between them, and the differences can be highlighted to show how a creole language evolves.

One issue that has been brought up in the study of creoles is the issue of complexity. McWhorter (2001) dedicates an entire chapter to this discussion, and other scholars have debated the issues of complexity in creoles. Although McWhorter (2001) argues that creole grammars are the world's simplest, I will argue that it is not necessarily the case. We will see examples of features from Kristang and Makista that were described in Chapter 6. I do not argue the opposite of McWhorter, that creoles are more complex, but that creoles are not necessarily the world's simplest grammars. In some cases creolization can lead to more complexity, and decreolization can lead to more simplicity.

Mufwene (2000) argued that creolization is a social, rather than a structural process. This next section will bring this issue into the discussion. Much of creolistics can feel like circular reasoning, that language X can be considered a creole, because it contains features that are deemed to be creole-like because those features were identified in languages that were once deemed creoles.

The identification of creoles as a separate language type must have its origins in the idea that these mixed languages that arose in colonial circumstances were not inferior versions of their European lexifiers, but rather languages in their own right with distinct

syntactic systems. Identifications of features found in creole languages came out of work that examined the structures of languages that had been identified as a creole, but this identification originally derives from the social circumstances from which they emerged.

7.2 Creole Exceptionalism, Substratism, and Superstratism

The identification of creoles as distinct languages from their lexifiers gave rise to the questions of whether they exhibit some kind of universal features. Bickerton (1981, 1984) in his Language Bioprogram Hypothesis argues that as languages contact and a new one is rapidly formed, as in a pidgin and creole, the parameters are reset to unmarked settings as found in the human language organ. He then argues that because of this, creoles will share universal features, and their input languages will not exert much effect on the structure of these languages.

Bakker et al (2011) further demonstrated the exceptionality of creole languages by quantitatively measuring features of creole languages against those of other non-creoles. Using phylogenetic mapping, they showed that creoles tend to cluster amongst themselves, apart from other languages. The features they looked at in their study were taken from Holm and Patrick (2007), a comparative creole work.

The notion that creoles are exceptional within the world's languages has been met with contention, in particular by DeGraff, who argued that the very idea of creole exceptionalism has its roots in history and colonialism (DeGraff 2003, 2005). He argues that creoles should be considered as their own group of languages from a sociohistorical standpoint, rather than a linguistic one. DeGraff (2008) points out similarities between Haitian Creole and English, such as the negation marker preceding the verb, despite both languages having evolved from languages with a negation marker after the verb. Since Haitian is viewed as a creole language to begin with, this feature can be used as evidence that Haitian is a "typical" creole language, where the same cannot be said about English.

The large transmission of lexical items from the lexifier to the creole suggests that a vocabulary is not the only way a creole is influenced by the lexifier. Mufwene (2004) argues that creoles should be classified in the same way that other languages are, and that they share a genetic relationship with their lexifiers in a similar way to how language families and sub-families group together. He suggests that English- and Dutch-lexified creoles should be classified as Germanic languages, and French-, Spanish-, and Portuguese-lexified creoles should be classified as Romance. This thought has been continued by DeGraff (2005), who compared Haitian Creole with its lexifier, French. French, being descended from Latin, is

230

grouped in the romance languages along with Portuguese, Spanish, and others. However, he argues that defining how close a language is to its relatives depends on the way in which we investigate the structure. He argues that along certain parameters, French is closer to Haitian Creole than it is to Latin.

Decades of work on creoles (Bickerton 1984, Holm 1988, DeGraff, 2003) has led to the idea that they are a subset becoming ingrained in the field; even amongst those who disagree with creole exceptionalism a barrier already exists before any research is taken out. (Ansaldo & Matthews (2007) argued that the creole/non-creole divide is an "artificial dichotomy" (3). I acknowledge that this thesis is yet another contribution to the idea that creoles form a group of languages, whether structurally or (as I argue) socially. One reason for this is that creoles have historically been viewed as lesser or broken versions of their lexifiers. While this is certainly not the case in modern linguistics and creolistics anymore, and has not been for a long time, the relationships between creoles and the languages in their ecologies must be taken into account.

7.2.1 The Cafeteria Principle and the Feature Pool

Bickerton (1981) uses the term "cafeteria principle", saying that languages do not simply pick up features from various sources, as a diner would pick up different food from a cafeteria. This is in criticism of some of the substratist viewpoints that creole features are taken from the substrate languages in a seemingly arbitrary way.

Siegel (2008b) defended the "cafeteria principle" in a description of Hawaiian Creole. It is notable among creoles in that there were many languages that contributed to its creation. The superstrate language was English, being the language of the Americans who took over the archipelago in the 19th Century. Among the substrates are the indigenous language Hawaiian, as well as the languages immigrants brought with them from Japan, China, the Philippines and Portugal. As there were many languages that were in the backgrounds of those who spoke Hawaiian Pidgin and what would become Hawaiian Creole, we can find features with parallels to the various input languages, without pointing to a single substrate with overarching syntactic influence. For instance, relative clauses may occur pre-nominally or post-nominally, reflecting the features found in the substrates. We can see a parallel with the genitive in the Luso-Asian creoles, where pre- and post-nominal particles indicate input from more than one source language.

Siegel also describes the maintenance of features from substrate languages. Hawaiian Creole uses one word *got* as a possessive as well as an existential. This existed before contact

with Cantonese speakers, but the fact that Cantonese has the same word for possessive and existential likely reinforced this structure – this can also be said of Makista. The same word *ada* is used in Malay for possessive and existential, and is also found in Kristang as *teng*. The similar usage in Makista, although parallel with the Cantonese structure, would not have had its origin in Macau, but from the Malay-Portuguese creole that had arrived in the colony from Malacca. Its continued usage could partially be attributed to the influence of Cantonese on speakers of Makista.

While Siegel (2008b: 77-78) embraces the cafeteria principle, he notes that it is not without constraints. Although a creole may pick up features from any number of sources, it must have a place to which it can transfer. We can see this example in Makista usage of the copula *sã*. The copula in Cantonese has many of the same functions as its equivalent in Portuguese, which can explain why we can find examples using *sã* in Makista that would be near-identical to a Portuguese structure. However, the transfer from Cantonese included other functions of the copula that are not found in Portuguese, that of the cleft (see Chapter 6, Section 3).

The cafeteria principle also has parallels with the Feature Pool hypothesis, which in turn is related to the theory of Linguistic Ecology. In the case of Makista *sã*, the introduction of a copula would be something that had emerged in Macau, as there is no equivalent in Kristang. As there was no existing copula in the language upon arrival to Macau, there was room for a copula to enter the language, either from Portuguese or Cantonese. It is likely that the copula transferred from Cantonese because it carried with it not only the properties of equating one thing with another but the usage of the introduction to a cleft clause, something that does not occur in the other Luso-Asian creoles.

7.2.2 Causes of change

Migge & Goury (2008: 313), in the comparison of TAM markers in Surinamese creoles, explained that there are four processes that can lead to change: substrate influence, internal change from a substrate calque, superstrate influence, and leveling. These processes can be applied to Kristang, Makista, BC, and TC and we will look at the influence of each of them on the four languages.

    a.   Substrate influence

We can see differences in the degrees of substrate influence in relation to the linguistic ecologies present in each of the locations where the languages were spoken. Malay and Indonesian influence are found to be the strongest in Kristang and TC, where there was little

Portuguese influence after the Dutch took over Malacca. The types of serial and auxiliary verbs found in Kristang have similarities with Malay, especially Kristang *tokah* with Malay *kenah* (See Section 6.2.2).

TC has very strong substrate influence, to the point where its genitive system had picked up a third type, when the other two languages only had two. This bare genitive is identical to the Malay and Indonesian structure, and it must have been through substrate influence that it was introduced, since the language already had two genitive structures in its repertoire. TC passives are another strong suggestion for substrate influence, where a verbal prefix *di-* was incorporated directly from Malay without any translation or calquing into a Portuguese-derived equivalent (See Section 6.8.5).

b. Internal change from calque

The variation in pronouns we find in Kristang is a candidate for a calque from Malay, which also has a short/long variation unrelated to case. There are important differences between the Kristang and Malay pronoun variation, however. Malay short pronouns are clitics that must be attached to the main verb in the VP. Kristang, on the other hand, has multiple intervening elements between the subject pronoun and the main verb, and this rule overcomes that of any type of cliticization of the pronoun. It is because of this that I propose the origins of short/long pronouns in Kristang are likely a calque of Malay, but have their own functional rules that have developed Kristang-internally.

c. Superstrate influence

This is most apparent in Makista, where the superstrate Portuguese remained in use throughout Macau's history. We can see the influence from the superstrate in the usage of morphemes such as the verbal past participle suffix *-ado* that also occurs with Makista verbs (see Section 6.6.2). Portuguese also exerted its influence on Makista with the genitive patterns, where the DI-type has had much more prominence and distribution than in Kristang (see Sections 6.7.2-3). BC has similar patterns of genitives to Makista, although its socio-political history was different. This would be a case of a lack of substrate influence that diminishes the superstrate's impact on the structure. In the case of Makista there is instead exertion of the superstrate into the morphosyntactic structure of the language.

d. Leveling

Leveling is a process in which there are two competing features within a similar semantic or pragmatic role, and one of the two becomes the dominant feature (Siegel 1997). The two

types of genitive structures are a good example of this, in that they are two different ways to express the same thing. In Kristang, the two forms have their own semantic domains where one feature is more likely to be found than the other. Makista, however, overwhelmingly uses the DI-type genitive, with a remnant of the SA-type system still used in Kristang. While this is an example of leveling, we must also see why one of the features had been preferred over the other. Leveling itself is a process that takes place in many languages, so we cannot attribute that alone to the linguistic ecology of Macau. We can, though, see that the continuing influence of Portuguese on Makista led to the choice of DI-type genitives becoming the dominant morpheme over the other.

7.3 Comparative Creole Literature

The notion of creoles forming their own separate group is strong, even among those who argue against creole exceptionalism. One of the products of this notion is the publication of comparative creole literature. Holm & Patrick (2007) looked at features of creole languages from around the world, and from various lexifiers. The features that they found to be most common among the languages chosen were used in Bakker et al (2011). While this was a very thorough study of creoles, the main commonality among these languages is that they are all considered creoles. A similar resource for comparative creole studies is the *Atlas of Pidgin and Creole Structure* (APiCS). Similar to Holm and Patrick (2007), the main commonality between the languages is that they are considered creoles, and have various input languages due to their worldwide distribution.

Cardoso, Baxter, and Nunes's 2012 volume takes a look at Ibero-Asian creole languages. This narrows the scope of comparative studies to creoles that have Spanish or Portuguese as a lexifier, and are spoken in Asia. Due to historical reasons, these are limited to South and Southeast Asia. The similarity between these languages has become more significant, as the Luso-Asian creoles share a common lexifier, and their relative geographical proximity to each other also has sociohistorical significance. Due to the history and connections between these languages, similarities in their structures cannot be solely due to having typical creole features. It also becomes important to look at the differences between these languages as well.

Alleyne (1986: 305) pointed out the importance of looking at differences in comparative creole studies:

> We therefore treat differences between substrate languages as weakening the
> possibility of substrate influences. Then we treat differences between substrate

languages and creoles as a further weakness of the substratum hypothesis. But

then we may overlook differences between creoles, preferring to treat them as

one in order to support other hypotheses.

What previous creole studies tended to investigate, especially those that look at creoles in general with little regard to their histories, were similarities between creoles in order to develop a theory on creole universals. One aspect that is not given enough prominence is that creoles do not develop within a vacuum, and are created by speakers who actively create ways to communicate. Ansaldo (2009) argues that it is the language ecology that leads to the creation of these creoles, and therefore will affect their features.

Muysken & Smith (1986: 2) touch upon the issue of substrate influence versus internal development and creole universals:

Although it may well be the case that such parallels are the result of substrate

influence they cannot prove the substratist case if the same phenomena are also

claimed by the universalists to represent the unmarked settings of various

parameters.

[T]o *prove* substrate influence we have to look for 'marked' structures

appearing in both languages – the language potentially subject to substrate

influence and the potential substrate language. (Muysken & Smith 1986: 4)

These quotations address one issue that is common within creolistics, whether a common feature between the substrate and the creole language is due to influence, or mere coincidence. Universalists argue that if a feature is not "marked" then it is due to language universals rather than contact influence. The very notion of markedness here is problematic; what is considered to be marked depends on what one considers to be universals. It also must take into account how a language's features differ from other languages it is related to or has a contact relationship with. What may be a marked feature in a West African context could very well be an unmarked one in East Asia. Up until this point, the common features identified across creole languages are simply those that have been identified and described, and this is something that will change as more research is done.

When we look at the environment in which a creole language develops, we can see that the children who would go on to nativize a creole language would not have done so with influence from the pidgin only. This is another important point that the idea of language ecology can contribute to, in which the language takes its structures from the surrounding landscape. Siegel (2007: 176-6) notes that "at no stage did a group of people suddenly start speaking an unexpanded pidginized variety as their first language—even in rapid

nativization. Rather, the pidginized variety was one of the many sources in the context environment, and its simplified features entered the 'pool of variants'." In the case of Malacca, the communication between the Portuguese traders and Malay locals would have likely been some kind of pidgin, but this does not prevent them from speaking their own languages in the same way they would at home. As the Eurasian community expanded and developed its own distinct identity, their language came together as well. This language would have used the pidgin as a skeleton upon which to build a full language, and would have been informed by the previous linguistic knowledge of Malay and Portuguese. The features we find in Kristang today could not have come from nothing, when we consider that the city had been a busy port at the time of Portuguese colonization. It is highly unlikely that an entire community would have formed its own language without any contact or influence from the majority language of the city, Malay. We can see its legacy in the structure of the language today, as well as the other creoles in this thesis.

Some of the criticisms of the substratist idea, and that of language transfer, and why we have many similarities between creoles, have been laid out in previous studies. Bickerton (1986) criticized the framework for not coming up with any constraints to language transfer, an argument furthered by Mufwene (1990). Since then, there have been developments in how language transfer would work in a substratist system. Siegel (2007: 180) argues that for transfer to occur, the most important factor is for an item to have somewhere to transfer to. A lexical item from the lexifier must map onto a corresponding morpheme in the substrate. We can see how this affected the differences between the different creoles. Cantonese has an equative copula, while Malay does not, and we can see how Cantonese influenced the structure of Makista. Not only do we have the usage of such a copula, *sã*, but the way it is used also marks it as particularly Cantonese. The cleft in Makista is almost identical to that of Cantonese, and is not a feature of standard Portuguese.

A similar phenomenon is described in Lefebvre (2008) called relabeling. When a creole is created, the parameters from the substrate language are transferred to the creole, with the phonological shape of the lexifier (Lefebvre 1998). This is attested in lexical transfer into creoles from the substrate. Muysken (1988) and Lumsden (1999) argued that relabeling is strictly lexical, but Lefebvre argues that it can be functional as well.

To illustrate lexical relabeling we will consider the Haitian Creole word *vyann*. *Vyann* has a phonological shape derived from the French *viande*. While both have the semantic meaning of "meat", Haitian *vyann* also has the meaning of "edible animals" as a complement of the verb "to kill", a meaning that is not part of French *viande*. Fongbe, Haitian's substrate,

236

has the word *làn*, which does carry both the meaning of "meat" and "edible animals" (Lefebvre 2008: 199). This is an example of how, when transferred from the lexifier, the semantics of a particular lexical item change within the creole.

Lefebvre (2008: 199) describes this process in three representations of a lexical item. Part (a) represents the lexical entry as it is in the lexifier:

a. /phonology/$_i$

[semantics]$_i$

[syntax]$_i$

The next part (b) is the assignment of a second phonological representation to the semantic and syntactic features of the item:

b. /phonology/$_i$ /phonology/$_j$

[semantics]$_i$

[syntax]$_i$

The final part (c) is where the phonology from language *i* is abandoned, leaving only one phonological shape. It is the phonology that has changed but the underlying syntax and semantics that remain the same.

c. /phonology/$_j$

[semantics]$_i$

[syntax]$_i$

This results in "hybrid lexical entries that have the same semantic and syntactic properties of the original ones and phonological representations that are derived from phonetic sequences from another language." (Lefebvre 2008: 199).

Going back to the instance of the copula in Makista, we can see the following steps:[12]

a. /são/$_{Por}$

[to be, describes permanent state]$_{Por}$

[present tense, third person plural]$_{Por}$

b. /hai6/$_{Yue}$, /são/$_{Por}$

[to be]$_{Yue}$

[no tense/person distinction, can introduce a cleft]$_{Yue}$

---

[12] I use the following ISO 639-3 codes: Por for Portuguese, Yue for Cantonese, Mzs for Makista

c. /sã/*Mzs*

    [to be]*Mzs*

    [no tense/person distinction, can introduce a cleft]*Mzs*

This is an instance of lexical and functional relabeling. The Portuguese copula *ser*, from which Makista *sã* derives, has similarities with the Cantonese copula in conveying the meaning of "to be", but the function of *são* and *sã* do not match perfectly. The syntactic properties of Cantonese *hai6* would have had to transfer over to Makista along with the semantic ones, or we would not have examples of the copular cleft in Makista as we do. Portuguese also makes a distinction between the permanent and non-permanent copula, respectively *ser* and *estar*. For non-permanent states, Cantonese instead uses stative verbs rather than an overt copula. This is another reason why the lexical relabeling took the phonological shape of *sã* rather than one derived from *estar*.

The above examples look at when two different lexical items have differing semantics, but sometimes the substrate can have many words for something where the lexifier only has one. Lefebvre notes that in Fongbe there are many words for "to cut", but only one equivalent in French. Therefore, the phonological shape of *couper* took on the semantics of all of the different words for cut in Fongbe. Another example is the word for "rice" in the Luso-Asian creoles. Malay has two different words for cooked and uncooked rice, which Portuguese lacks. We can see a similarity in Kristang, where there is only one word for both cooked and uncooked rice. This can be attributed to the Potuguese lexical item *arroz* corresponding to both cooked and uncooked. However, TC has two different words. The TC word *neli* came to mean uncooked rice, while the Portuguese-derived *aros* shifted to the semantics of the Malay *nasi* "cooked rice".

a. /arroz/

    [rice]

b. /nasi/, /arroz/

    [cooked rice]

c. /aros/

    [cooked rice]

Lefebvre's description of this relabeling is in reference to radical creoles, that is, creoles that have a structure most similar to their substrates. The types of relabeling found in the Luso-Asian creoles are not as widespread as those in the more radical creoles. Lefebvre illustrates the pronominal paradigm in Haitian and Solomon Pidgin, where the pronouns

almost perfectly match with those of the substrate. Standard Malay has many more pronouns than Kristang, and the paradigm did not expand to match it.

The connection between demographics and language structure is an important theme in this thesis, and has been in previous literature. Parkvall (2000: 186) noted that although Bickterton (1981) advocated the language bioprogram hypothesis, he also suggested that the proportion between the speakers of a substrate and those of a lexifier needs to reach a certain point to allow pidginization to occur. Goodman (1985) doubted such connections between demographics and pidgin- and creolization, and McWhorter (1999) took it further to suggest that it was motivation to acquire the target language rather than demographics that played a role in the formation of creoles.

Parkvall (2000) took this into account when comparing creoles based on a list of structures deemed to be common amongst prototypical creoles. He also looked at the various demographic situations in the areas where these creoles are spoken. It was through a quantitative method that he scored how "prototypically creole" a language was. My method, while mixed quantitative-qualitative, bears some resemblance to Parkvall, in that I investigate certain features in the four languages. Due to space constraints and in the interest of qualitatively describing and comparing the features, my list of features explored is smaller than those looked at in Parkvall's work.

Parkvall (2000: 199-200) listed 45 features to mark the amount of creoleness a language had. For most of these features, they were either present in all four creoles or absent from all four creoles studied in this thesis. I will highlight the features where there were differences between the creoles.

a. Personal pronouns derived from superstratal oblique forms

Personal pronouns in Kristang, BC, and Makista are derived from the subject forms. TC has both oblique-derived and subject-derived forms.

b. Conjunction marking "and" derived from "with"

This is true of Kristang and Makista, using the co-ordinating conjunction derived from the Portuguese *com* "with". TC uses *o* as a conjunction.

c. Juxtaposition of nominal possessor and possessum (Peter book ~ book Peter)

Genitives in all four creoles are expressed with the genitive particle *sa* (and its equivalents) or the preposition *di*. The one exception to this is TC, which in addition to the other two genitives has a possessee-possessor word order to express genitive.

d. A PP [prepositional phrase] (other than that of the lexifier) is used to express the (absolute) possessive (book for Peter)

The genitive particle *sa* in isolation is the 3SG possessive pronoun. Makista has examples of *vossa* and *nossa*, derived from the Portuguese.

e. Zero complementizer (where one would be required in the lexifier)

Kristang, Makista, and BC have a complementizer. TC has examples that allow for no complementizer, but there are examples that do have it.

f. Zero equative copula

Makista makes much use of equative copula *sã*, while Kristang does not express an equative copula.

g. Diachronic variation or merger of /v/ and /b/

Kristang exhibits a merger of /v/ and /b/, while Makista does not.

The case of TC is interesting in that there are features where it differs from the other three. Some of these differences are due to it being more creole-like, according to the model, and some are due to it being less creole-like. Two of these features are where TC is more like the substrate than the other languages. One of them is that TC uses a co-ordinating conjunction that is not derived from the lexifier preposition *with*. Malay also has a different form for the conjunction *and* than it does for the preposition *with*. This similarity with the substrate also makes it less creole-like, according to the model. The other one is the juxtaposition of the possessor and possessee in genitive constructions. Although TC uses the genitive particle *sa* as well as the preposition *di* in some of its genitive constructions, it also uses the form of possessee-possessor. This is another instance where TC is more like Malay than the other three languages, which exclusively use both of the other genitive forms. As with the usage of a co-ordinating conjunction, this makes TC more like the substrate. However, it also makes it more creole-like.

The features where Makista differs from the rest are where it is deemed less creole-like according to the model. One of these is the usage of the equative copula. Makista makes usage of its copula more so than the other creoles. According to the model, this makes it less creole-like. It also makes it more similar to the substrate, which in Makista's case was Cantonese, as well as to the lexifier. Another feature where Makista differs is the usage of possessive pronouns. This also makes it less creole-like, and at the same time makes it more similar to the lexifier, but less like the substrate.

We can see similarities in other language contact situations with Chinese languages, also in Malaysia. The Peranakan people are descendants of Chinese merchants who settled along the Straits of Malacca on the Malay Peninsula. While retaining some of their traditions, they had children with local Malays and this community eventually became a Malay-speaking group, distinct from other Chinese and Malay communities. Their language is also a distinct variety of Malay, called Baba Malay. Ansaldo, Lim, & Mufwene (2007) note that although it may be considered a mixed language rather than a creole, it does exhibit some features that are reminiscent of creole languages. One reason that it is considered a non-creole is that the social dynamics of the creation of a Peranakan culture is very different from the situations one finds in the plantations of the Caribbean or the Eurasian communities in South and Southeast Asia. However, we do find a similar lexifier-substrate relationship within Baba Malay, with Malay and Hokkien, respectively.

One feature of note is the usage of existentials. Baba Malay makes usage of existentials in the same way as Malay:

440.	Baba Malay (Ansaldo & Matthews 1999: 57, as cited in Ansaldo, Lim, & Mufwene 2007: 216)

ada	manyak	temput lagi

have	much	time	more

"There is plenty of time"

441.	Malay (Ansaldo, Lim, and Mufwene 2007: 216)

ada	koran	di	meja

have	newspaper	on	table

"There is a newspaper on the table"

This has parallels with the usage in Hokkien.

442.	Hokkien (Ansaldo, Lim, and Mufwene 2007: 216)

Ě-mng ŭ	lû-tiăm

Amoy	have	hotels

"There are hotels in Amoy"

They note the heavier usage of existantials in Baba Malay than in standard Malay, with frequencies similar to that of Chinese (Ansaldo, Lim, & Mufwene 2007: 216).

The above features are chosen by Parkvall because they have been deemed to be creole-like, but we must also look at whether the substrate languages also exhibit these features. For example, Malay possessives are expressed by juxtaposition between the

possessor and possessee. This is identical to one of the ways TC expresses the genitive. Additionally, Malay has no equative copula, which is reflected in the syntax of the creoles in this thesis. The one creole that sticks out in this feature is Makista, which has Cantonese as a main substrate, which does exhibit usage of the copula.

One thing the creolist must be careful about is the creation of prototypes or other types of features or phenomena that are deemed to be typical of creoles. Labeling a group of languages as creoles, and then finding similar features between them, and then decreeing that these features are what make a language a creole is circular logic at best. Creoles are a diverse group of languages that are best defined through social histories. We could similarly group languages that spread through empires, such as English, Latin, or Russian.

These generalizations, while problematic when applying them to creoles as a whole, can be useful in other contexts. Faraclas et al (2007) note that generalizations made about creoles are more appropriate if we apply them to a restricted group of languages. This is part of the current study, in which I reach conclusions and generalizations based on Kristang, Makista, BC, and TC only. I do compare them to other creoles, as well as their input languages, but I only focus on the consequence of the interaction between languages in the Luso-Asian context. The generalizations I make from this have wider implications in the study of Luso-Asian creoles, general creolistics, and language contact. Some of the findings in this thesis may apply to other similar situations of language contact and creole creation, but I lay out the circumstances that led to the structures in these four creoles to make clear that there are many factors that must be considered in the study of these four languages. These languages did indeed get their structures from the interaction of a lexifier and a substrate, like every other creole language. But what has influenced the actual phenomena we find in the languages, and how they work in each one, are consequences of additional factors beyond the simple fact of language contact and pidginization taking place.

Faraclas et al (2007) go on to note that some of the opposing viewpoints within the fields of creolistics and language contact can be mitigated by such changes in perspective: "Studying these similar groups of languages one by one rather than considering Creole origins categorically can reveal that seemingly contradictory assertions are less opposed to one another than has often been assumed" (Faraclas et al 2007: 228). With this in mind, I do not just want to present this work within the context of creole studies, but also within language contact as well, and the implications thereof. The differences in results from language contact situations from a single-origin group of languages can reveal what can

result from centuries of contact in differing circumstances, and how this may impact other languages.

Along with Mufwene (2000), DeGraff (2003), Alleyne (1986), Ansaldo (2009) and many others, I hope to help break down the artificial barrier between creoles and other languages. However, Faraclas et al (2007) also note that there are further implications to breaking this down. If we can prove that creoles do not form a typologically distinct class of languages, but rather a group of languages that have similar socio-historical origins, then we must also analyze every other language with the social and political dynamics within the communities in mind. Faraclas et al (2007: 233) point out that "The sharp distinction between 'linguistic' and 'non-linguistic' … must be abandoned, not just in the study of Creoles but in the study of any human language." Although I feel the need to distinguish "linguistic" (such as syntactic structure and morphology) from "non-linguistic" (such as history and demographics)—this is evident in my separation of discussion of the histories of Malacca, Macau, and Jakarta, from the bulk of this thesis, which focuses on morphosyntactic structure—I think that this divide is more conceptual, and is permeable, as we have seen in the evidence from the four Luso-Asian creoles. What happens on the non-linguistic side of the barrier has direct implications on the linguistic side, and these four creoles are just one small example within all human language representing how the non-linguistic and the linguistic interact, and how it can have a profound effect.

7.4 Comparing Kristang, Makista, Batavia, and Tugu Creoles

With the previous studies in mind, a look at differences between creoles can show how their environments can determine how their structures will develop. This is what makes a comparison of Kristang and Makista not only useful but necessary in the study of creole development. For instance, a difference between structures between a Portuguese-lexified creole in India and an English-lexified creole in the Caribbean may likely have always been there. However, Kristang and Makista have a common origin, so the differences found between them would not have been there at the outset.

As we have seen in the discussion of features of these two languages, some of the more prominent differences can easily be connected to the language ecologies of Malacca, Macau, and Jakarta. The post-nominal genitive marker -sa has far more widespread usage in Kristang than in Makista, with Makista demonstrating a preference for the Portuguese-influenced structure. However, remnants remain of the post-nominal system, showing that there was likely a shift from a system more similar to the one currently used in Kristang. The

pervasiveness of Cantonese in Macau also has left its mark on Makista, in particular the usage patterns of the copula that is not found in Kristang. The usage patterns of serial verb constructions as well reflect a Cantonese influence that would not have been present in Malacca.

McWhorter (2018) makes the following claims regarding creole exceptionalism, the first being about the copula: "Typically, creole languages do not preserve their lexifier's copular morphemes, or use them only variably and in fashions quite different from in the lexifier." (McWhorter 2018: 10) While it is true that Makista and Kristang use the copula in ways other than Portuguese does, they exhibit similarities with their substrate languages. Chapter 6 explains the usage of the Makista copula *sã* to make a cleft construction, very similar to that of Sinitic languages. The Makista copula is also used as an equative, in the same way as the Portuguese *ser* (of which *são*, the origin of Makista *sã*, is a form). Another issue is that of case, in that Creole creators eliminate case distinctions in lexifier pronouns (McWhorter 2018: 12). McWhorter argues that creoles have eliminated case distinction in pronouns, and that creoles lexified by English, French, and Portuguese use as the 1st Person singular pronoun the form derived from the object. The use of case here is problematic, as there are languages that have variation in pronunciation of pronouns that exist aside from case. Malay, for example, has clitic forms of pronouns that change depending on the information structure of a sentence. This appears to have an influence on the pronouns of Kristang as well. None of the creoles in this thesis use case distinctions in the way that Portuguese does, that is, with a separate form of pronoun. This does not mean that case has ceased to exist, but it is expressed in ways different from that of Portuguese, as another restructuring of the language to resemble that of the substrate or a new innovation. One such example of the marking of case is the marker *ku*, which is a differential object marker that is used on human objects.

As we have seen in Section 6.4, Kristang does have variation in pronoun forms. While this is not a case distinction as one can find in Portuguese and other European languages, it is a distinction that serves a function. Again, this is a restructuring of language, rather than a simplification. Fon Sing (2018, 45) notes that "Creoles are considered special because they are somewhat closer (compared to non-Creole languages) to the universal system of features that shapes linguistic systems before they get "trapped" by historical change." Kristang is still spoken 500 years past initial contact with certain features that some would deem typical of a creole. However, many of these features are also present in Malay (verb-initial aspect, lack of verbal morphology, lack of overt copula). We do find an overt copula in Makista, where the

input languages (Portuguese and Cantonese) use them. The patterns of usage of the copula in Makista also resembles that of Cantonese.

Many of the features that McWhorter cites as prototypical of creoles are also found in Austronesian and Sinitic languages, languages that have helped form these four creoles into what they are. Although they certainly have been subject to language contact, Malay and Chinese are not considered creoles, due to their less mixed cultural origin, relative to the Eurasians of Southeast Asia. Such features common to the four creoles, but also to Malay and Chinese are the usage of free morphemes to indicate tense and aspect, the lack of case in pronouns, and reduplication. These are strikingly non-Portuguese-derived structures, but by looking at the linguistic ecology of Southeast Asia we can see that they do not come from nowhere.

7.4.1 The Post-Creole Continuum

The Post-Creole Continuum is the notion that post-creolization, there exists a continuum with two ends: the basilect, the least prestigious variety; and the acrolect, the most prestigious (Stewart 1965). This represents a continuum between the "most creolized" variety at the basilect, and the standard variety at the acrolect. Such basilect/acrolect continua exist in Haiti (Haitian Creole/French) and Jamaica (Jamaican Creole/Jamaican English) among other speech communities. While a useful tool for analyzing the varieties of language spoken in former European colonies of the Americas, it is less useful for looking at Kristang. For one, it supposes that the standard or prestigious variety of the superstrate still has an official status and/or is spoken by a large number of community members. The variety of speech on the continuum may also depend on the addressee, and many may be competent in both basilectal and acrolectal varieties. This is not the case with Kristang, where the languages of prestige are Malay and English, the former originally being the substrate of Kristang, and the latter playing no role in the creation of the language. Since the 17th Century, standard Portuguese has hardly been spoken in Malacca, which left no space to even make a continuum.

Despite this lack of a continuum per se, there appears to be influence on Kristang by Malay, in particular the short/long pronoun distinction. Kristang also retains many of the features that would be considered typical of a creole, such as a lack of verbal morphology, usage of serial verb constructions, and the copula being used much less than in Portuguese. While Portuguese influence was effectively cut off in the centuries after the Dutch takeover of Malacca, Malay continued as a main language in contact with Kristang. The features mentioned above, listed as part of a prototypical creole structure, are also features found in

Malay. Arguments in favor of creole exceptionalism state that upon creolization the features are set to that of a typical creole, and as contact with the lexifier continues we get patterns of decreolization and a creole continuum. As the lexifier language left Malacca, Kristang speakers found themselves speaking mostly with Malay- and Chinese-speaking people. Despite the potential for stronger influence from Dutch or English, there was little influence on the syntax and it retained its "creole-like" structures. As these structures closely resemble those of Malay, they could have remained in the language due to constant Malay contact, rather than a universal creole system influencing the grammar.

Makista, on the other hand, did have some contact with Portuguese in the past centuries. What this has led to is that many of the differences between Makista and Kristang can come down to influence from Portuguese. Among these are the genitive system, a larger usage of the copula, and verb phrase morphology in Makista. While Macau has some of the elements that can contribute to a post-creole continuum, the actual presence of Portuguese in the government, media, and education is smaller than English in Jamaica, for instance.

We also have an interesting case with BC and TC. In looking at which languages resemble the lexifier, BC and Makista can be grouped together, while TC and Kristang can be grouped into being more like the substrate. As opposed to the greater geographical distance between Malacca and Macau, BC and TC were spoken in Jakarta, with the former spoken in a more central area. Despite being closer to each other geographically, they were distinct and separate communities.

In many of the Atlantic creole communities, the space was shared by both basilectal and acrolectal speakers, leading to the continuum. Such concepts are problematic in the case of the creoles studied in this thesis, as there was variation, but more distance between groups of speakers than would be the case in, for instance, a Caribbean island with a majority of the population speaking a Creole.

We must also consider the idea of multiple substrates. This is an issue that has arisen in other creole studies, such as on Jamaican Creole (Kouwenberg 2008). In the early English period in Jamaica, slaves were brought from various areas in West Africa, and there were a multitude of languages spoken by the slave population, including Akan, which is sometimes seen as the dominant substrate. However, there was no one demographically dominant African ethnic group in the early slave era, so there were multiple substrates that could have contributed to Jamaican creole.

Makista has the distinction of having a Malayo-Portuguese background but a history of being spoken in a majority Cantonese-speaking city. We have also seen that while Malay

is the dominant language in Malacca, there have been Chinese speakers (mostly Hokkien) from before the European period. While it may not be prudent to go so far as to deem Hokkien a substrate, we cannot rule out influence to Kristang either. Baba Malay has some similarities with Kristang that are not found in standard Malay.

7.5 Are Creole Grammars Simpler?

One major argument for creole exceptionalism is the notion that creole grammars are the simplest grammars, and that this simplicity is due to the genesis as a pidgin and nativization as a creole language. Simplicity and complexity are notions that are at once intuitive and hard to define and quantify. The focus of this thesis is the morphosyntax of Luso-Asian creoles, and as such this will not include a discussion of simplicity in phonological or sociolinguistic terms.

The idea that every language has the same amount of complexity is often cited in discussions of linguistic complexity (McWhorter 2001b, Szmrecsanyi & Kortmann 2012). This idea claims that if one feature of a language simplifies over time another feature will complexify as a result, and the language as a whole is no more simple or complex than before. This idea is known as the ALEC (All Languages are Equally Complex) statement (Deutscher 2009), or the linguistic equi-complexity dogma (Kusters 2003), which states that simplicity in Domain A must result in Complexity in Domain B.

There are also distinctions in what type of complexity one will measure. Global complexity looks at the complexity of a language as a whole, whereas local complexity looks at the subfeatures of a language in terms of complexity, such as phonological, morphological, syntactic, semantic & lexical, and pragmatic (Miestamo 2008). There is also the notion of relative versus absolute language complexity. Relative complexity deals with language learning, and is a comparison of the native language of a particular speaker and the target language(s). The complexity of an L2 depends on its relationship with the complexity of L1. Absolute complexity, on the other hand, is concerned with the number of parts in a system, and the connections between the different parts (Miestamo 2008). Fully investigating the absolute complexity of a language requires both quantitative methods, including counting the different parts of a system and their frequencies relative to each other; and qualitative methods, explaining how each component works and some of the nuances each one may provide to meaning.

With these in mind, I want to look at the four creoles here in terms of absolute and local complexity. Although I am comparing the languages in relation to each other, I am not

concerned with the ease of a speaker of one language learning the other. This section will concern the number of features and the relationships between them. I will also focus on local complexity for three reasons: first, a full description of the global complexity of Makista, Kristang, BC, and TC is such a large undertaking as to be beyond the scope of this thesis; second, this thesis focuses on morphosyntax, and any discussion of phonological complexity would be a distraction; and third, this discussion of complexity is only part of the argument for substratism. I hope that this argument can provide a framework for future studies of complexity in languages, creole and non-creole.

Determining a language's complexity also depends on how much of the language one is looking at. As this is about the morphosyntax of creole languages, we can focus on comparing the complexity of the morphosyntactic systems of different languages. It is quite difficult to come to a conclusion of how complex a language's syntax is, as this depends on what one is looking at as well as the methodology and viewpoints one is coming from. Chapter 6 of this thesis looks at different features found in the four creoles, and one could easily judge complexity based on the categories discussed. However, this is by no means an exhaustive list of features within these languages, nor is it an exhaustive list of all grammatical features which may not even be present in these languages

I am not setting out to make a conclusive statement on the complexity of any Luso-Asian creole, or any other language for that matter. My goal here is to show how complexity can be measured, and how we can apply these measurements to certain features of Luso-Asian creoles.

7.5.1 Paradigmatic simplicity

This refers to the number of forms one finds within a language. This can be found in such paradigms as verbal morphology. See the following examples in Mandarin, English, and Portuguese.

Table 22: Mandarin Verbal Paradigm

wo3 chi1	wo3men chi1
ni3 chi1	ni3men chi1
ta1 chi1	ta1men chi1

Table 23: English Verbal Paradigm

I eat	we eat
you eat	you eat
he eats	they eat

Table 24: Portuguese Verbal Paradigm

eu como	nós comemos
tu comes	vós comeis
ele come	eles comem

These tables illustrate the present tense forms of the verb *to eat* in Mandarin, English, and Portuguese. In Mandarin there is only one form of the verb, while in English there are two, one with the suffix *-s* for 3rd person singular, and the other for all other pronouns. Portuguese, however, has a different form for each pronoun. Looking at this paradigm, one can conclude that in the case of verb conjugations, Mandarin is the simplest of the three with one form, and Portuguese is the most complex with six. English is more simple with two forms, but is not the simplest. Paradigmatic complexity can be analyzed on a continuum, as I will show with the following examples, in Mandarin, English, Portuguese, and Russian.

Table 25: Mandarin Personal Pronouns

number	person	pronoun
singular	1st	wo3
	2nd	ni3, nin2
	3rd	ta1
plural	1st	wo3men
	2nd	ni3men
	3rd	ta1men

Table 26: English Personal Pronouns

number	person	subject	object
singular	1st	I	me
	2nd	you	you
	3rd	he, she, it	him, her, it
plural	1st	we	us
	2nd	you	you
	3rd	they	them

Table 27: Portuguese Personal Pronouns

number	person	subject	object of verb (direct)	object of verb (indirect)	object of preposition
singular	1st	eu	me		mim
	2nd	tu	te		ti
	3rd	ele, ela	o, a	lhe	ele, ela
plural	1st	nós	nos		nós
	2nd	vós	vos		vós
	3rd	eles, elas	os, as	lhes	eles, elas

Table 28: Russian Personal Pronouns

number	person	nominative	genitive/accusative	dative	instrumental	prepositional
singular	1st	ja	menja	mne	mnoj	mne
	2nd	ty	tebja	tebe	toboj	tebe
	3rd	on, ona, ono	ego, ejo	emu, ej	im, ej	njom, nej
plural	1st	my	nas	nam	nami	nas
	2nd	vy	vas	vam	vami	vas
	3rd	oni	ix	im	imi	nix

These tables, ordered from the least to most forms, are the following: Mandarin with 4 individual forms (not counting plural pronouns, which are formed only by adding the suffix *-men*), English with 12, Portuguese with 15 (not counting 3rd person plural pronouns, which are formed only by adding the suffix *-s*), and Russian with 32. Looking at personal pronoun forms alone shows how we can compare these features based on paradigmatic complexity. This just shows a slice of a language, and to get a better idea of its paradigmatic complexity one would need to look into other features, including formation of plurals, verb inflection for person and number, or verb inflection for tense and aspect.

7.5.2 Distributional complexity

Another type of complexity relevant to the present study is that of distributional complexity. Languages indeed differ in the number of forms they can have, but the distributions of forms also differs from language to language. This is a functional parameter that is less concerned with paradigms but rather where forms are found and the patterns which can be found in their distributions.

For examples of distributional complexity I am continuing with the usage of personal pronouns. If we look at 2nd person pronouns alone, Mandarin is more paradigmatically complex, with three separate forms (singular familiar *ni3*, singular formal *nin2*, and plural *ni3men*), while English only has one form (*you*). There are strategies for conveying the plural second person pronoun in English, which are not obligatory.

443.   English

Bill$_i$, may I speak with [you$_i$/*you all$_j$/*you guys$_j$]?

444.   Mandarin

Bi3er3	wo3	ke3 bu4 ke3yi3	gen1	[ni3/*ni3men2]	jiang3jiang3hua4
Bill	1SG	may-NEG-may	with	2SG/*2PL	speak

"Bill, may I speak with you?"

In these examples, the distribution of grammatically singular and plural pronouns is identical between Mandarin and English. Where the pronoun's referent is semantically a singular, the plural (and in English overtly plural) pronoun is ungrammatical.We see a difference when the referent is semantically plural:

445.   English

"Bill and Mike, may I speak with [you/you guys/you two/?you all]?

446.   Mandarin

Bi3er3,Mai4ke4   wo3   ke3 bu4 ke3yi3   gen1
[*ni3/ni3men2/ni3men2lia3]

Bill,   Mike   1SG   may-NEG-may   with [*2SG/2PL/2PL
two.people]

jiang3jiang3hua2

speak

"Bill and Mike, may I speak with you/you two?"

The English example allows for more types of pronouns than Mandarin does. More specifically, the English pronoun *you* has a wider distribution than Mandarin *ni3*. Another example of distributional complexity can be seen in the third-person pronouns of English. While English does not have as much grammatical gender as other European languages, there are masculine and feminine pronouns (*he* and *she*), in addition to the neuter pronoun (*it*) and the indefinite pronoun (*one*). *It*, while gender neutral, is also most commonly used for non-human referents, and when referring to a human is often very semantically marked. When one does not know the gender of the referent, or does not want to explicitly state the person's gender, the third person plural pronoun can be used:

447.   English

I'm going to see a doctor₍ᵢ₎ this afternoon but I don't know what they₍ᵢ₎ will say to me.

They, normally being a plural pronoun, only refers to one person in this instance. *They* can also be used as a pronoun for those whose gender identity falls outside of the he/she binary (Balhorn 2004). In such cases, *they/them* is used identically to *he/him* and *she/her*, except in the case of verbal inflections. While it is semantically a singular, it still is plural grammatically. Compare the utterances "They are a nice person" vs "She is a nice person". We can therefore say that the usage of third person pronouns in English is distributionally complex, because of the ways that different pronouns can be used to express third person singular.

When looking at the complexity or simplicity of a language, distributional complexity should be given as much importance as paradigmatic complexity. This can be seen in other subfields of linguistics, where phonology looks at the number of phonemes within a language but is also concerned with the distribution and patterns of usage of the phonemes.

7.5.3 Overspecification

Overspecification is another factor in measuring the complexity of a language. As McWhorter (2008) points out, overspecification increases complexity. Using English as an example, we can look at the relativizers used in the language, as explained by Huber (2012). First, English has the distinction between who(m) and which, the former being reserved for animate and the latter for inanimate.

448.   English

   a.   The man [who/*which] I saw

   b.   The table [*who/which] I built

There is also the instance of the relativizer *that*, and the zero relativizer. These two cut across the distinction of animate and inanimate.

449.   English

   a.   The man [that/Ø] I saw

   b.   The table [that/Ø] I built

Huber points out that many of the world's languages have only one type of relativizer. This makes English relativizers especially complex, with one reserved for animates, one for inanimates, and one that works with either one, as well as the possibility of no relativizer.

7.5.4 Simplicity in non-creole languages.

Gil (2007) points out that McWhorter (2001, 2005) concedes that Riau Indonesian (not considered by Gil to be a creole) is as simple as other creole languages, but that it is a creole itself. McWhorter (2005) also claims that colloquial varieties of Malay and Indonesian are creoles as well, partly because of their simplicity.

McWhorter (2018) claims that non-creole languages are less complex than other. He also cited Chinese as having simplified over the course of restructuring. In the current study, this raises the question: what happens when creole languages are in turn restructured into another creole? If McWhorter's assumption is taken into account, the creole-like features of (Bazaar) Malay would have undergone further restructuring after initial contact with Portuguese.

Farquharson (2007) points out how the notion of complexity within languages, especially creoles, is problematic. One has the tendency to begin with a viewpoint that looks for what a creole language lacks, especially in relation to its lexifier. What we should do, instead, is to focus on which features are actually there, and how the different features and

phenomena in each of the creoles works. Farquharson (2007: 30) also argues the issues of aspect markers in Jamaican creole. The progressive marker is a pre-verbal element, which he argues to be an affix rather than a free morpheme. This is because it cannot receive independent stress and nothing can come between it and the stem of the verb. This is also the case in Kristang, where the progressive marker *ta* is unstressed, has no short and long forms (as does *lo(gu)*), and nothing can come between it and the verb (although this third point is also true of the other TAM particles). This raises the question of whether it truly is a free morpheme, or perhaps a clitic or an affix, as Farquharson argues is true of the Jamaican particle.

He also points out that reduplication is often overlooked in the literature regarding simplicity and complexity in creole languages. Farquharson (2007: 31) argues that "there is no *a priori* reason why adding an affix to a stem is more complex than reduplicating the stem—both may have morphophonemic repercussions." This is certainly true of the four creoles in this thesis. We can also find further structural similarities with the aspect markers and with verbal reduplication. Both of these are found adjacent to the verb itself, and both have a morphological implication. In addition, there can be no intervening words between the aspect marker and the verb.

450.   Kristang

bos   ja   kumi
2SG   PFV   eat
"Have you eaten?"
*ja   bos   kumi
PFV   2SG   eat

In a similar way, reduplicated verbs (and not just verbs, but any grammatical category) are always adjacent to each other.

451.   Kristang (Baxter 1988: 60)

eli   ta   remá remá
3SG   PROG row row
"He was rowing and rowing"
*eli   remá   ta   remá
3SG   row   PROG row

One reason that we see a parallel between pre-verbal particles and verbal reduplication is that both have some effect on a verb, with reduplication in this case having a modal effect.

Paradigmatically, such instances of reduplication might be more simple than Portuguese, which has a number of affixes, nominal and verbal, to use to express certain things that are expressed with reduplication in the four creoles. However, this does show that there are indeed verbal morphological processes taking place in these languages. As Farquharson argues, I too argue that reduplication is a feature of verbal morphology in Kristang and the other creoles, and is a system that is on par with Portuguese affixes, or Chinese aspectual markers.

With the evidence shown in Chapter 6, we can see that not just the existence of a certain feature, such as reduplication or serial verb constructions, has been passed down to the creoles, but the patterns of usage have as well. As Aikhenvald (2006) notes, serial verb constructions are a very wide category and do not constitute a single phenomenon. Inkelas & Zoll (2004) also point out that although reduplication is, at its most simple definition, a reduplicating of part or all of a word or morpheme, beyond that their function and appearance varies widely from language to language. Therefore, it is not enough to say that the simple presence of reduplication indicates that a language had undergone creolization at some point in its history. We must also consider the patterns that we can find within these features that indicate any kind of restructuring.

Take for example the restriction on reduplicating nouns next to a numeral. This is a restriction that also exists in Malay, and this is the most likely source of the identical restriction in Kristang. If the feature of reduplication had been simplified through the act of creolization, or if Kristang had begun from nothing and developed reduplication as a language-internal evolution, we would wonder what the chances are that the same kind of restriction exists in Kristang as in Malay. It would appear that the introduction of a new rule within the syntax of reduplication would complicate it. For instance, the English plural is required when it expresses a semantic plural, whether there is a numeral present or not.

7.5.6 The Genitive in Kristang and Makista

This section returns the discussion to genitive patterns of Kristang and Makista. As outlined in Section 6.6, Luso-Asian creoles have two ways to express the genitive: the SA-type and DI-type. SA-type genitive uses a possessive particle stated after the possessor. This particle is derived from the Portuguese possessive pronoun *sua*, but the syntactic structure resembles that of other languages, such as Baba Malay and Hokkien, more than Portuguese. The DI-type genitive uses the preposition *di* in front of the possessor. This is nearly identical to the Portuguese usage of the preposition *de* (Baxter & Bastos 2012).

Of the four languages focused on in this thesis, all of them have both SA-type and DI-type genitives. TC also has a third, the bare genitive, which the other three do not. Paradigmatically, we can conclude that TC is more complex (with three types of genitive) than BC, Kristang, and Makista (each with two).

To review, tokens of the genitive were counted in the following categories, as originally outlined in Baxter & Bastos (2012)

    a. Kinship

    b. Body parts

    c. Ownership

    d. Other interpersonal relationships

    e. Classificatory

    f. Part-whole

    g. Spatial/locative

    h. Origin/source

    i. Material composition

The categories near the top of the list are more likely to have a human possessor, whereas the ones near the bottom are more likely to have a non-human and non-animate possessor. The ones near the top are also more likely to be an inalienable possession. The results from Baxter & Bastos (2012) can be found in Section 6.8.2.

The SA-type and DI-type are both found in nearly all of the categories (the one exception being DI-type in kinship), but there is a clear pattern that emerges from the chart. The semantic categories that are more likely to be human or inalienable are mostly SA-type, while the opposite is true in the categories more likely to be non-human.

As Makista has the same paradigmatic pattern of genitives as Kristang, I looked at the Makista data for tokens of the genitives along the same semantic categories.

As can be seen in Section 6.8.2, the DI-type type is much preferred in Makista compared to Kristang usage. Indeed, all of the examples with SA-type are part of *iou-sua* "my". Another notable feature of this chart is that the SA-type genitives are still prominent in the areas where they are prominent in Kristang, although much less so. Considering that Kristang and Makista have the same origins, the genitive system in Makista resembles an erosion of the Kristang system, where the prominence is similar but the number of tokens is much less.

Again, the number of genitive structures is identical in Makista and Kristang, so paradigmatically they have the same complexity. However, a look at the distribution of the two types by semantic categories reveals a distributional complexity in Kristang that does not exist in Makista. The reason for this difference can be seen with a look at the linguistic ecologies of the two languages. Since they came from a single origin, and the patterns of the two distributions have similarities, we can assume that the systems had similar origins. The type of genitive that took over in Makista was the DI-type, which is the structure derived from the Portuguese form. Unlike Malacca, Macau remained a Portuguese colony for centuries, and while few people actually spoke it in the colony, there remained official schools and administration run by the Portuguese. As the patterns of Makista drifted away from those of Kristang, they came to resemble standard Portuguese more and more, resulting in more decreolization than in Kristang.

Paradigmatically, Kristang and Makista have the same amount of complexity in terms of their genitive structures. In looking for genitives, I found both DI-type and SA-type genitives in both languages. I also did not find other genitive structures, as is the case with TC. Using these criteria we can conclude that the genitive patterns in these two languages are of equal complexity.

The investigation into the distributional complexity of the patterns of genitive highlight some of the limitations of looking at complexity through the paradigmatic lens only. Since these two languages come from a common origin only centuries old, it may seem quite obvious that they share many paradigms between themselves. However, as seen in previous chapters one can see that these two languages do have some differences in their syntactic structure. A look from a distributional point of view is also required to further highlight and investigate the differences in complexity between Kristang and Makista.

The main argument in McWhorter (2001) is that creole grammars are simpler, and the reason for this simplicity is the fact that they had undergone creolization from their input languages, thus stripping them of any "unnecessary" features. This then streamlines the medium of communication which helps people from disparate backgrounds communicate with each other. The relative youth of creole grammars is also a factor, with older grammars having picked up more features due to being around for longer. Part of his argument compares the Northeast Caucasian language Tsez with the Surinamese creole language Saramaccan. Using the former as a representative of older grammars and the latter as a representative of newer grammars, he looks at different features in the grammars in which Tsez is more paradigmatically complex than Saramaccan. Other features that exist in Tsez

but not Saramaccan, such as the marking of evidentiality in past tense marking, are also listed as evidence that older grammars are more complex than those of creoles. One feature he mentions in his argument is that of possession, and he states that "[no] creole known to [him] makes a grammaticalized alienable/inalienable possessive distinction". The possessive distinctions in Kristang are not strictly an alienable/inalienable binary, but there is a very clear pattern as to which forms are preferred, and where.

Kristang and Makista provide an opportunity to look at two creole languages that have a common origin and went on separate trajectories. Rather than comparing one language as a representative of non-creoles and another unrelated one as a representative of creoles, this looks at how creoles themselves can change in relation to each other and other languages. We also have the opportunity to look into the implications of such change onto factors such as the morphosyntax, and indeed the relative complexity of languages. McWhorter is correct in demonstrating that the only way one can determine the complexity or simplicity of a language is by comparing it to another, and by looking at two related creoles this is useful in determining the origins of complexity in syntactic structure.

Looking at the history of the two languages, we can see that Makista speakers would have had many more opportunities for contact with Portuguese speakers than speakers of Kristang would have. This difference in circumstances has had an effect on other features of Makista syntax (see Pinharanda Nunes 2012). The two main types of genitive in Luso-Asian creoles also have a clear origin, the SA-type being more similar to languages that acted as substrates, and the DI-type type being more similar to Portuguese. Of the two types of genitives within the inventory of Luso-Asian creoles, the one that was more similar to that of the lexifier became more prominent in Makista. This should come as no surprise due to the longer presence of standard Portuguese in Macau. The patterns of usage of SA-type genitives in Makista suggest that the genitive system in Kristang could be its predecessor. The existence of both types in BC and TC is further evidence of this system having existed in the original language.

It is important to look at the various usages of the types of genitive in Luso-Asian creoles, because simply counting the types within a paradigm does not give us the full picture. Mühlhäusler (2012: 123) states that "in my view it makes little sense to ask questions about the simplicity of any tool (and languages are tools employed in the business of communication) unless one considers the tasks in which it is employed". This is relevant to the current study and how we look at the distributional complexity in Luso-Asian creoles. The ways in which SA-type genitives are used vary across the four languages greatly, even

though the form and the meaning are the same. Counting the tokens of each one can give us an idea of how they are used differently, and this reveals how much the SA-type genitive has died out in Makista and BC. As we can see, it depends on the method in which we look at complexity to give a clearer view of each structure. Personal pronoun inventory is another important part of this study. We can look to previous studies of creole languages and their pronouns, and how they are related to or different from the input languages. Mühlhäusler (2012) investigated Norf'k pronouns through the lens of linguistic complexity. He mentioned that there are types of pronouns found in Norf'k that do not have any obvious parallel in the three input languages: English, Tahitian, and West Indian Creole. The short/long distinction in Kristang appears to be a local innovation. Although Kristang does not have the same subject/object distinction that Portuguese does, the short/long distinction is something that does not exist in Portuguese. This has the effect of increasing complexity.

One thing that substratists and creole universalists alike can agree upon is that creolization leads to change – otherwise they would not be considered separate languages. The disagreement is based upon how these changes take place and what comes out of them as a result. One of McWhorter's main arguments is that creolization will more often than not lead to a simpler syntax, relative to the input languages. The above case of the genitive structures in Makista and Kristang show that Makista's genitive system is simpler than Kristang's, and this is due to decreolization. Thus, the ancestor genitive structure of these Southeast Asian Portuguese creoles is more distributionally complex than what is found in Makista. I do not argue whether the grammars of Kristang and Makista are more or less simple than other languages. They are arguably more simple than standard Portuguese with its verbal inflections and other morphology. However, we can compare the two languages with each other, and look at their relative "simplicity". There are indeed counterexamples to this, in which Makista has a feature that is more complex than one found in Kristang – the cleft construction using a copula, for example. However, what makes these grammars more simple or complex in their various features appears to be a symptom of language contact, rather than whether one language is more of a prototypical creole or not.

When regarding the issue of simplicity one must also consider the origins of the scholars and the points of view through which the creoles are being investigated. As I will argue in the next section, creoles have been seen as typologically distinct because of their sociohistorical backgrounds. We must be careful not to argue that the absence of a feature from one language, where the native language of the researcher does have it, indicates simplicity. Gil (2007: 76) notes that "in practical terms, when linguists say things like 'the

grammar doesn't encode it, so speakers have to figure it out from the context,' the 'it' in question is, more often than not, a distinction that is absent from the target language but one that the linguist expects to find, because it is prominent in English and other European languages."

This is especially relevant in creole studies, where the lexifier has provided enough words for the creole to have a very strong resemblance to them, but the grammar has been restructured such that many of the morphosyntactic features that indicate whether a language is Portuguese, English, French, or others have largely disappeared. We therefore have two biases that shade the arguments. The first is, as Gil points out, the language of academia, or the language of prestige, usually English or another European language. The second is the ultimate European origin of the creole languages. When looking at a non-European analytic language, such as Mandarin or Malay, there are no "missing" pieces that one might notice in a Portuguese- or French- lexified creole, such as verbal suffixes, or markers for number and gender on nouns. Ultimately, the question of languages' relative simplicities is not the most important one in the present study. However, as a part of McWhorter's overarching argument, I felt that it must be addressed within the context of this thesis.

7.6 Creoles as a Social Construct

Creolistics is recognized as a distinct field by linguists on both sides of the universalism versus substratism debate. McWhorter (2001) emphatically makes the argument that they are a typologically distinct set, and states that all creoles will fulfill a number of criteria that he laid out:

    a. A lack of inflectional morphology

    b. A lack of tone on monosyllabic words

    c. A lack of semantically opaque word formation

He claims that these criteria are how one can predict what languages are creoles, and are what set creoles apart from non-creoles. He contends that this set of conditions arises from a language having been formed from a pidgin. McWhorter's arguments in favor of the creole prototype are about the emergence of a creole from a pidgin, but also arise from the fact that they are new languages. He says that many of the features that can be found in older languages are a consequence of their age, such as the presence of tone on monosyllabic words in Vietnamese (McWhorter 2000). He does, however, note that the criteria of a creole prototype can be measured on a continuum, and that creoles can and do change much like non-creole languages. He further notes:

A creole honing to the Prototype today is most likely to be found in situations offering two crucial conditions. The first is that the initial social context limited learners' ability or desire to acquire the lexifier to such a degree that a pidgin variety of that lexifier developed (unlike, for example, Réunion). The second condition is that the expansion of the pidgin into a natural languge did not include borrowing from the lexifier, but instead occurred mostly or entirely via the recruitment of language-internal resources. (McWhorter 2000: 110)

This quote refers to circumstances that deal with the creation of a pidgin and the later expansion into a creole. Kristang, Makista, BC, and TC all developed after this process, but it can be useful to look at the circumstances surrounding their subsequent histories and relationships with the lexifiers and substrates based on these conditions that McWhorter laid out.

Kristang is the creole that is spoken by the descendants of the creators of the pidgin spoken in Malacca, and the creole that arose from it, living continuously in the same city. Portuguese rule lasted for about a century in Malacca, so the main source of contact was during this period. After Portuguese rule ended, they were cut off from a large amount of communication with standard Portuguese, although not entirely.

Makista was brought to Macau and remained spoken there throughout Macau's time as a colony of Portugal. In this case, the language was severed from its original substrate, Malay, and was placed in a city that provided contact with another substrate, Cantonese. Access to Portuguese was much more easily available, as the Portuguese remained in control of the colony until the end of the 20th Century.

BC did not emerge until after the Dutch had taken over Malacca, so their contact with Portuguese lasted the same amount of time as Kristang did. However, they were moved to a city where Portugal had never made settlements, so they were further removed from Portuguese than if they had remained in Malacca. Contact with Malay remained unbroken, as Malay was the lingua franca of the Indonesian archipelago.

TC, just like BC, did not emerge until after the Dutch had taken over Malacca and brought Eurasians to Jakarta. Compared to BC, TC was in even more contact with Malay speakers. As a result, many of the lexical items in this language are from Malay, while equivalent terms in other creoles are from Portuguese. TC also has the distinction of directly borrowing a morpheme from Malay without a Portuguese lexical origin, the passive prefix *di*.

Semantically opaque word formation is another major component of McWhorter's argument. He often cites English words like *understand* or *make up*, as well as verbal prefixes such as Russian *na-* and German *ver-*. Some features in the creoles may be considered to have opaque word formation. In particular the Makista asymmetric serial verb construction with the major verb *virá* "turn" comes to mind. Other verbs of motion used in asymmetric serial verb constructions are used to literally convey the direction of an action, especially with *come* and *go*. This verb, however, does not convey any movement but rather is used to move a narrative forward, in much the same way English uses *then*. I have found no examples of *virá* being used as the major verb in a serial verb construction with the meaning of literally turning, although it is used in such a way in isolation.

We can also consider phrases with the Kristang verb *abri* "start, undo, open". Some of these are more obvious in their meanings, even for someone unfamiliar with the language. This includes *abri korsang* "to reveal one's feelings (lit. *open heart*)", which clearly is a metaphor. However, other phrases are less clear in their meanings. One is *abri mang* "open hand", which can either mean "to be thankful" or "to ask for something". This is semantically not quite the same as the English phrase "with open hands", which conveys a meaning of generosity.

Some of McWhorter's own reasoning is problematic, in that he argues that noncompositional morphology is an indicator of an older language. He makes an example for English: "The prefix re- in English is quite productive, and yet there are many uses of it which have drifted semantically into noncompositionality (having actually done so within the European languages they were borrowed from), such as represent and repose." (McWhorter 2000: 92). If such words had become noncompositional in a European language before having been adopted into English, then the word itself was borrowed in full, and the evolution of English has little to do with how *repose* became a monomorphemic word.

Those that argue for substratism also do not dispute the distinctness of creoles, but look at them from a social and historical point of view (DeGraff, Mufwene, Ansaldo). The rise of European colonies and slave plantations led to a mass movement and resettlement of people unknown to the world before. Amongst many changes, this also led to the creation of new languages in these colonies.

> "Aside from the fact that creoles have no structural prototypes, an important problem with distinguishing creoles from non-creoles on structural grounds lies in the fact that creoles do not share their features universally." (Mufwene 2000: 70)

I want to add to this conversation by providing evidence that one creole language gave birth to four distinct but related creoles; and that the differences in these creoles' structures can be attributed to their distinct sociopolitical histories. Some of the features that meet McWhorter's criteria for creoleness can also apply to other languages. It is true that Kristang, Makista, BC, and TC diverged after initial creolization. However, they are all identified as creole languages today, despite their differences. As with Luso-Asian creoles, the creole languages of the Caribbean have been spoken for multiple centuries, and certainly have had many opportunities for contact with other languages in that time. McWhorter argues that features that stray from the Creole Prototype are a function of time. Since all four of the creoles in this study had origins at the same point of time, we can see how they adhere to or stray from this prototype. The differences between them, therefore, cannot be attributed to time. Farquharson (2007:34) points out that from the perspective of language acquisition and language learning, all languages are equal in age. As someone learns or acquires a language, they are using their own cognition to form a system based on what they are learning from the input. While it is absolutely true and necessary that there be consistencies from what was input to what system forms in the learner, the result is a completely new system in a person, that had not existed before. It is with this in mind that Farquharson argues that the idea that complexity results from older languages is flawed.

Another significant trait of these creoles is that they are all urban minority languages. This makes them doubly susceptible to change from other languages, and, unfortunately, decline or even death in the case of BC and TC. As some point out (Mufwene 2000), language change does not happen in a vacuum, and this is especially the case when a certain language is spoken in a large, often multicultural city. As we have seen in Chapter 6, many of the differences between these languages can be attributed to the geographic and demographic situations in which they found themselves. Certain traits stick out, such as the use of the copula cleft in Makista that is notably absent in Kristang, BC, and TC. This difference cannot be attributed to a difference in time, but place. Out of the three cities, Macau has by far the highest concentration of Sinitic speakers, in this case Cantonese. This copular cleft feature resembles that of Cantonese to such a degree that it is more likely it emerged out of contact than by simply language-internal change. We can see that the effects certainly are structural, but the cause is social. Just as the instance of pidginization, then creolization, was social in the first place, much of the subsequent change can be attributed to that as well. As brought up by DeGraff (2005) and restated by Ansaldo and Matthews (2007), the ideas of creole exceptionalism can be broken down into three "creole myths": the myth of simplicity, the

myth of decreolization, and the myth of exceptional diachrony. Simplicity has already been brought up in this chapter, with evidence that different features in these four creoles have differing degrees of complexity, despite McWhorter's claim that complexity is a function of time. I also pointed out an instance of a language becoming less complex due to its becoming more similar to the lexifier. The myth of decreolization touches on the idea that a creole can only move in directions away from the lexifier (creolization) and towards the lexifier (decreolization). "In this respect DeGraff notes how, while all languages undergo change, only within creolistics do we find concepts applying the *de-* prefix (e.g. 'de-Latinization,' 'de-Africanization' are not used with respect to other languages). They rightly point out its absence from other fields of linguistics (was English de-Germanizing following the Norman invasion?). This is the kind of thinking that can arise from the notions of prototypes of a category of languages. Rather than conforming or differing from a prototype, the languages have reacted to their respective ecologies and contacting languages. This is something that is relevant in the third myth.

Matthews & Ansaldo (2007: 13) also point out this assumption about decreolization, that "whenever Creole languages are spoken alongside their lexifier, Creole languages, being impoverished, would naturally tend to assimilate 'back' to it." In respect to verbal morphology and the usage of DI-type genitives, Makista is the most "Portuguese" of the four creoles in this study. In the case of the genitive system, this assimilation to Portuguese is actually the result of an erosion of a feature, and thus a simplification. In addition, Makista was spoken in a city with a) comparatively more contact with the lexifier and b) contact with a new substrate, Cantonese. While some of the original Malay features did continue on in the structure of Makista, we can see the features that it has are often parallel to similar structures found in Cantonese, and not Portuguese. It is unclear how a language which both assimilated in some respects to Portuguese but also to Cantonese would be considered in the creolized/decreolized scale. In this case the similarities that have moved towards Portuguese have distinguished Makista from Kristang, but so have the similarities with Cantonese. This makes it difficult to put it on a creole/non-creole scale considering the complex relationship it has had with different languages in contact. The language ecology theory can help explain why Makista differs from Kristang. For instance, Makista has been spoken in a city with Portuguese linguistic influence, alongside Cantonese. In this way, a language ecology, rather than a linear scale this represents more dimensions in which a language can gain or lose features, or qualities of such features.

The third myth is that of exceptional diachrony. While it is true that many creoles are interesting in that they have a lexicon coming from one language but a grammar from some other source, the notion of exceptional diachrony is that they adhere very strongly to the language bioprogram, more so than any other language that undergoes change. As I noted earlier, these four creoles have their own structures as a reaction to the environments in which they developed. This thesis is a synchronic study, so I do not look deeply into how the languages have possibly evolved over time. However, looking at these four creoles, and the similarities and differences between them, gives us an insight into the result of four centuries of contact in different situations.

The role a community plays within its greater language ecology is something that is important as well. As mentioned before, Haitian is a creole that has many millions of speakers and has been the vastly dominant language in that country for centuries. On the other hand, the Luso-Asian creoles were always urban minority languages, with speakers regularly coming into contact with speakers of other languages. Despite the small population of the speakers, their role within society would not have been that of slaves, during the Portuguese period, although some Eurasians were taken as slaves during the Dutch period and would eventually form the BC and TC communities. Ansaldo, Lim, & Mufwene (2007: 222) discuss how the role a community plays in its society can also have an affect on the language itself, noting that "population structure… can be a more significant factor than population ratio." Alleyne (1971) made a similar argument, that the demographic proportions do not take prominence in what leads to change, rather the relationship between different groups of people does. This idea is very similar to the notion of super- and substrates, whose definitions come from the relationship the speakers of such languages had with the other.

Of course, ratio is something that must also be taken into account. I laid out the demographic information earlier (Chapter 2, Section 7) to show the percentages of the populations of the various cities have. This is a part of the language ecology, especially as we investigate both the type of influence and change, and the extent of change. As Ansaldo, Lim, and Mufwene (2007) explain, the Peranakans, a Chinese-descended community in Malaysia, are a case of a community that speaks either a mixed language or a creole, but they came to exist out of a different experience than other creole-speaking communities. When looking at creoles, we must also consider the difference between Luso-Asian Creoles, which are spoken by small mixed communities that usually existed as a small part of a large city, versus the Atlantic creoles, which were more based on plantations and the mass movement of people through slavery.

Even though I do not compare Kristang, Makista, BC, and TC to the creole prototype, I do not avoid comparison in the study of these four creole languages. I do not compare them to the creole prototype. In a study that concerns language ecology we must consider the influence other languages have exerted on them. It is when we consider the influence from the languages with which the speakers of these creoles would most likely have contact, that the differences between them, and the origins and motivations of such differences reveal themselves.

7.7 Conclusion

This chapter is meant to be a revisitation of the concepts explained in chapters previous to Chapter 6, with the actual data and analysis in mind. It also connects some of the concepts dealing with creole exceptionalism and substratism to the features of Kristang, Makista, BC and TC.

Comparative creole literature is a valuable resource for the creole researcher, especially one who wants to investigate typological tendencies, or the lack thereof. I present my own work as both an addition to and a contrast to some previous comparative literature. I continue in looking at how certain features in creoles compare and contrast with each other, as well as how they are similar or different from the lexifiers and substrates. What makes this significant is that it looks at four languages that had diverged after a single instance of creolization in Malacca in the 16th Century. Knowing that all four languages have a single common origin is significant in that it leaves any divergence to what had happened after initial contact, and not what might have been differences that existed from the beginning.

Simplicity is a part of the argument for creole exceptionalism. This idea builds upon McWhorter's criteria for a creole prototype. Because there are differences between the four languages in this thesis, we can also see how some features might be more complex in one language than in another. It is with this in mind that I looked at the distributional complexity of genitives in Kristang and Makista, as there is a difference in the distributional patterns between them. It appears that the genitives in Makista were influenced by continuing Portuguese presence in Macau, and that the lack of Portuguese influence could be a factor in the maintenance of a more evenly distributed two-form genitive in Kristang. If we compare the complexity of this feature between Kristang and Makista, we can see that Makista genitives are less distributionally complex than that of Kristang. We can also see that this is due to the influence of Portuguese. It appears that in this feature, decreolization had led to a decrease in complexity in this feature.

I touched upon the idea that creoles are a social construct, building off Mufwene's idea that creolization is a social process. The analysis of features presented in this thesis reflect the histories of the cities in which they are and were spoken, and the types of contact and change that could have resulted in the features we see in the Kristang of the 21st Century, and the Makista of the 20th. Although these four languages diverged post-creolization, they are still considered to be creole languages. I argue that this is because they are spoken by a mixed-race Eurasian community, rather than because of the syntactic structure of the languages, which has many parallels with Malay and Cantonese.

The line that divides creole languages from non-creoles greatly affects how they are described. This has an impact on their description and analysis in that they are placed within a group called "creoles", or describing which characteristics a creole language is more likely to have than a non-creole. This can also have an effect in the way about which creoles are written. Linguists, especially in the West, run the risk of inadvertently looking at languages in comparison to European ones, especially the major ones such as English or French. This is something that often cannot be avoided when looking at European-lexified creoles, since their ancestry is so clearly from (at least partially) a European origin. Creoles are languages that have undergone a significant restructuring in a comparatively short amount of time, while still retaining the lexicon from their lexifier languages enough to superficially resemble and perhaps even be mistaken for their lexifier languages. From this, the concept of absence takes hold. Portuguese uses definite articles, which also indicate number and grammatical gender. Kristang, Makista, BC, and TC do not have such a characteristic. Portuguese makes extensive us of verbal suffixes to express number, person, and tense, while the four creoles in this thesis hardly make use of this at all. If one begins to look at these languages through a Portuguese perspective, one would assume that it had undergone a loss of many features that Portuguese still has. However, the four creoles have a two-feature genitive system (three, in the case of TC), largely based upon human possessors and inalienability, which is something that Portuguese lacks.

I do not wish to argue that for every Portuguese feature these creoles lost, they picked up a creole-like or substrate-like one to preserve equal complexity. However, rather than focus on what is "missing" from these languages, I focus on the features they already have. I also do not argue that the mere presence of some of these features marks them as more or less creole-like, nor do I want to compare them with other creole languages that they are not related to. Instead, I set out to describe the ways these features interact within the languages

and compare with each other, while keeping in mind the patterns one can find in the substrate languages.

Lumsden (1999: 138) points out, in a discussion of the Fongbe contribution to Haitian grammar: "When the substrate and superstrate have contrasting semantic patterns, the Creole follows the substrate language." With the case of Kristang, it is true with many features, especially the expression of aspect, as well as the usage of serial verb constructions. The genitive is one where it sticks out though. It appears that there were two competing systems for the genitive, but both found a place within the language. Another interesting point is the way that they have different word orders, with X sa Y, but Y di X. Bickerton (1999: 58) notes that in the case of genitives in creole grammars, if a possessional morpheme fails to survive transmission from the lexifier, we may get the genitive with a similar word order, but no overt morphology (for example, *book of John* → *book John*). This does appear to be the case in these creoles, where the Portuguese-derived possessive morpheme *di* did survive, but is not found in every genitive construction. This could be a reason for the SA-type genitive structure to have a roughly opposite word order than the DI-type. We also see a similar pattern in TC, where the non-morphemic possessive follows the same word order as the one utilizing *di*. However, it is also important to consider the influence the input languages had had on the creoles. What was the reason for the "survival" of the morpheme *di*? Its "vitality" differs from language to language, so we must look into the reasons why it may have survived transmission into all four, but only thrived in Makista. It is true that the DI-type genitive in all four creoles has the same order as the bare possessive in TC. While this could be a result of a simple disappearance of the morpheme itself, we must also consider the impact that Malay/Indonesian had put on TC. The bare possessive in TC is expressed in exactly the same way as in Malay. This is one feature of many that it had picked up from Malay, to an extent that the other creoles have not. This is one instance where one must be careful not to jump to conclusions if a creole language follows so-called prototypical patterns, if language ecology can also explain such structure.

Chapter 8: Conclusion

8.1 A Case Study of Four Related Creoles

The situation of Portuguese-lexified creoles in Southeast Asia provides us with a good opportunity to see how creoles have evolved, how they relate to each other and to their input languages, and helps us get closer to the answers to questions regarding creole exceptionalism and language ecology. This case study of four related creoles show us how the language ecology and features with which they come into contact can affect their morphosyntactic structures. An important reason why I chose these languages to study is the historical evidence of their relatedness to each other. Much creole literature compares creoles by virtue of their being considered creole languages, or by having the same lexifier language. While this does have its uses and has given us answers about how creoles work, we should take into account languages that are definitely related to each other, as is the case with Kristang, Makista, BC and TC. Much of the previous creolistics literature (McWhorter 2005, Holm & Patrick 2007, Bakker et al 2011) compared different languages by virtue of them being called creoles. It is my hope that this will contribute to such literature, but also motivate future research to consider the relations between two languages other than just being creoles.

The four languages examined in the present study are very similar to each other because they come from the same source. It is when we look at the differences between them that we can see the significance of a comparison of such closely related languages. Makista's relatively high contact with Cantonese can explain some of its unique features in comparison with the other three languages. The isolation of TC speakers to a less urban area left them with more contact with natives to Java than with Europeans; we can see one result of this being the heavy presence of Malay/Indonesian-influenced structures, and even some morphemes directly borrowed from Malay without translation to a Portuguese equivalent. The degree of contact with Portuguese also reveals itself in the differences between Makista, spoken in Portugal's last colony, and Kristang, spoken in a city the Portuguese last controlled in the 17th Century.

A major part of this study is the effect language ecology can have on morphosyntactic structure. The mixed qualitative and quantitative methods have highlighted the similarities between the creoles and their substrates. This is also why the historical background and demographic profiles of the cities in question play an important part in the comparison of the

creoles' structures and those of their substrates. As minority languages in large cities, it is clear that the speakers of these four creoles have had many opportunities to communicate with speakers of other languages.

I wanted to investigate some of the reasons that creole exceptionalists give to explain how creoles are a special set of languages. One of these was simplicity, a topic explored in McWhorter (2001). Looking at the genitive patterns of the four languages (See Section 6.7), we can see that in this case, more "creole-like" languages such as Kristang and TC have more complex patterns than the less "creole-like" ones, Makista and BC.

The social aspect of language change is something that should not, and in the case of these four languages, cannot be ignored. Following in the footsteps of DeGraff, Alleyne, Ansaldo and others, I have demonstrated that the sociopolitical histories of a certain region can have an impact on the structure of a language that is spoken there. A substratist point of view helped inform this study and the analysis of the morphosyntactic features found in this thesis. We can see the influence of the substrate languages in almost all of the features discussed, to one degree or another. However, I also think that a superstratist point of view is warranted in view of the data of the creoles as well as the lexifier, Portuguese. While the language did undergo substantial restructuring to give us the Kristang, Makista, BC, and TC structures we have today, it is not without morphosyntactic influence from Portuguese as well. The preservation of the DI-type genitive is such an example. Even in TC, which largely uses a SA-type particle or a bare genitive, there are some examples of the DI-type prepositional genitive as inherited from Portuguese. Its survival through restructuring is what had let it expand in the Makista system out of continuing Portuguese influence. We can also see Portuguese influence in the TAM particles, where the verb-derived particle cannot occur on its own. The existence of two genitive systems side-by-side—one a continuation of Portuguese structure and one from a likely South Asian source—show how the creole had made room for structures originating in both the lexifier as well as the substrate.

8.2 Implications for Further Research

This thesis is meant to bring together data from four related Luso-Asian creoles, and analyze them through the lens of language contact, feature selection, and language ecology. The features described and analyzed in Chapter 6 are not an exhaustive list, nor is each description. What I have done is to demonstrate the implications that language ecology has on the structure and development of creoles, in a contrast to universalist points of view.

Although BC and TC are no longer spoken we still have data to access and continue to analyze, and there are still speakers of Makista and Kristang alive at the time of this writing. This leaves open the possibility of further research and, more crucially, further opportunities of documentation and revitalization of these endangered languages.

One section that is especially in need of further research is the short and long pronouns in Kristang. As the corpus is still small it is not enough to come to any definite conclusions, but what I have shown are suggestions of what they mean and how they function. It also merits studying the feature in Kristang alone, as it is not present to my knowledge in the other creoles, but such a study would not be appropriate in this comparative thesis. The effects of language ecology can also be further studied in the other languages spoken in Malacca, such as Baba Malay, Standard Malay, Hokkien, Tamil and English. As the Kristang-speaking community is shifting to English we can see how the ecology affects their usage in comparison to how it had affected Kristang.

Information structure likely plays a role into the variation of pronouns and other grammatical items in Kristang, and this is something where there is the opportunity for more research. While the main focus was the difference between long and short personal pronouns in Kristang, we also see a variation between auxiliaries (*kereh* "to want" vs *ke*), TAM markers (*logu* vs *lo*), and genitive structures (possessee-possessor with or without *di*). The recordings I made were on Kristang only, but the same elicitation exercises can be carried out with speakers of the other languages of Malacca to see what patterns are used.

Also related to information structure is ellipsis and sentence fragments, as touched upon in Section 6.5. I have shown that the progressive marker *ta* cannot occur as a fragment, while the perfect marker *ja* and the future marker *logu* may. For future study, a deeper look into this will be required, and a look into fragments that consist of other grammatical categories should be a part of that. This also will require a look into fragments in the other languages spoken in Malacca, as well as in Portuguese. Santos (2002: 82) notes that among native speakers of Portuguese, usage of *ta* "ok" as a fragment answer does not emerge until later on, after other fragment answers (including *já* "already") did occur in their speech. While the lexical features of Portuguese *já* and *ta* do not exactly match those of their Kristang counterparts, it is a striking similarity that *ta* as a fragment answer is not grammatical in Kristang and does not appear as early in Portuguese. Future studies can look into patterns of first- and second-language acquisition and if those align with grammaticality in the creole languages.

There is also opportunity to further explore the effects on language ecology on further Luso-Asian creoles, namely those of South Asia. As pointed out, there is evidence that South Asian languages have had an influence on the Southeast Asian creoles, including the SA-type genitive (see Section 6.7). Cardoso et al (2012) laid out the framework for which these types of studies can take place, and from this case study we can further investigate the impact on language ecology on the creoles' morphosyntax, and to what extent the differing ecologies of Southeast Asia and South Asia led to differences between them.

The methods I used in the investigation of genitive patterns was only part of the methods in Baxter & Bastos (2012). While this thesis applies the methods of showing the distributional patterns of the two or three types of genitive along semsntic lines, Baxter & Bastos (2012) continued to look at the structures of the genitive constructions themselves, again with a quantitative methodology. They showed how many tokens of genitive constructions were found in the categories of: (a) a deleted head; (b) a pronoun; (c) a NP with a pre-nominal modifier; (d) a DP itself determined by a post-nominal genitive; and (e) a bare noun (Baxter & Bastos 2012: 68-9). This is something that can be further explored in the cases of Makista, BC, and TC. It can also help reveal the extent of Portuguese impact on the genitives of Makista, as it had in the expansion of the DI-type genitive.

Another open question is the influence of language-internal development in creole structures. As this thesis looks at the impact of the language ecology, I focused on the languages with which speakers of Kristang, Makista, BC, and TC would have had contact. While it is clear that creole language have had their origins in contact situations, further investigation of their own internal change and development can help us remove the barrier that exists between creolistics and other linguistic fields. A look at the features investigated in this thesis through the lens of reanalysis and extension (Harris & Campbell 1995) can help contribute to the pictures of the structures of Luso-Asian creoles, as well as their similarities and differences.

8.3 Addressing the Research Questions

With these discussions in mind, I will restate my research questions:

1. In what major structural ways are Kristang, Makista, BC, and TC similar, and in what major ways are they different?
2. How did the substrates (Malay for Kristang; Malay and Cantonese for Makista; Malay and Indonesian for BC and TC) affect these differences?

3. How did the lexifiers (mostly Portuguese, some Malay and English for Kristang; mostly Portuguese and Cantonese, some Mandarin and English for Makista; Malay and Indonesian for BC and TC) affect these differences?

Chapter 6 outlined the major differences and similarities between these four Luso-Asian creoles, with a selection of morphosyntactic features. As all four languages have their origins in a single source, that of the Portuguese-based pidgin and creole from Malacca, we will find many more similarities than differences. It is when we investigate the differences that we can see the patterns and their implications in the developments of the languages. In some cases, such as Serial Verb Constructions (see Section 6.2), there are not major differences between all four. Other features, such as the genitive, show the same general structures present, but different functional patterns. Looking into these patterns can show us just how different the creoles can be, even if the underlying structures are identical.

When we can see the historical and current demographics of Malacca, Macau, and Jakarta, we can understand how the languages that speakers of Luso-Asian creoles would have played a part in their lives, and how they have affected the morphosyntactic structures of them. The removal of Malay as a substrate for Makista can be seen in that it does not exhibit a short/long pronoun sytem, which Kristang does. We can also see further Malay influence in TC, where the genitive structure has more aligned with the Malay/Indonesian system. Cantonese plays a role as substrate in Macau, which shows itself in the patterns of usage of the copula, especially in the formation of cleft structures.

An investigation into lexifier and substrate influence in creoles shows us how the feature pool and the language ecology have played a role in the developments of Luso-Asian creoles. As features were transferred from one language to the other, they took with them their own lexical and functional features from the substrate to the creole. Because the four languages had divergent development, we would expect the differences in features they have to align with the differences in ecologies as well. This thesis shows that there are indeed influences that can be traced to the language ecologies where they have developed, and that for Kristang, Makista, BC, and TC, their morphosyntactic structures are a product of their particular environments.

References

Aboh, Enoch. 2009. Competition and selection: That's all! In Enoch Aboh & Norval Smith (eds.), *Complex processes in new languages,* 317-344. Amsterdam: John Benjamins.

Aboh, Enoch & Umberto Ansaldo. 2007. The role of typology in language creation. In Ansaldo, Umberto, Stephen Matthews & Lisa Lim (eds.). *Deconstructing creole*, 39-66. Amsterdam: John Benjamins.

Ackerlind, Sheila R. & Rebecca Jones-Kellogg. 2011. *Portuguese: A reference manual.* Austin: University of Texas Press.

Adams, Kathleen & Kathleen Gillogly. 2011. *Everyday life in Southeast Asia*. Bloomington, IN: Indiana University Press.

Aikhenvald, Alexandra. 2006. Serial verb constructions in typological perspective. In Alexandra Aikhenvald & R. M. W. Dixon (eds.). *Serial verb constructions: A linguistic typology*, 1-68. Oxford: Oxford University Press.

Aikhenvald, Alexandra. 2018. *Serial verbs*. Oxford: Oxford University Press.

Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory* 21, 435-483.

Alleyne, Marvyn. 1971. Acculturation and the cultural mix of creolization. In Dell Hymes (ed.). *Pidginization and creolization of languages*, 169-86. Cambridge: Cambridge University Press.

Alleyne, Mervyn. 1986. Substratum influences – Guilty until proven innocent. In Peter Muysken & Norval Smith (eds.). *Substrata versus universals in creole genesis: Papers from the Amsterdam Creole Workshop, April 1985,* 301-315. Amsterdam: John Benjamins

Alleyne, Mervyn. 2000. Opposite processes in "creolization". In Ingrid Neumann-Holzschuh & Edgar W. Schneider (eds.). *Degrees of restructuring in creole languages*, 125-133. Amsterdam: John Benjamins.

Ansaldo, Umberto. 2009. *Contact languages: Ecology and evolution in Asia*. Cambridge: Cambridge University Press.

Ansaldo, Umberto, Lisa Lim, & Salikoko Mufwene. 2007. The sociolinguistic history of the Peranakans. In Umberto Ansaldo, Stephen Matthews & Lisa Lim (eds.). *Deconstructing creole*. Amsterdam: John Benjamins.

Ansaldo, Umberto & Stephen Matthews. 1999. The Minnan substrate and creolization in Baba Malay. *Journal of Chinese Linguistics* 27(1), 36-68.

Ansaldo, Umberto & Stephen Matthews. 2004. The origins of Macanese reduplication. In Geneviève Escure & Armin Schwegler (eds.). *Creoles, contact and language change: Linguistic and social implications,* 1-19. Amsterdam: John Benjamins.

Ansaldo, Umberto & Stephen Matthews. 2007. Deconstructing creole: the rationale. In Umberto Ansaldo, Stephen Matthews & Lisa Lim (eds.). *Deconstructing creole*, 1-18. Amsterdam: John Benjamins.

Arana-Ward, Marie. 1977. A synchronic and diachronic investigation of Macanese: the Portuguese-based creole of Macau. Master's thesis, University of Hong Kong.

Arcodia, Giorgio. 2017. On Sinitic influence on Macanese. *Language Ecology* 1(2), 158-184.

Arends, Jacques. 1994. The socio-historical background of creoles. In Jacqes Arends, Pieter Muysken, & Norval Smith (eds.). *Pidgins and creoles: An introduction*, 15-24. Amsterdam: John Benjamins.

Arends, Jacques, Pieter Muysken, & Norval Smith (eds.). 1994. *Pidgins and Creoles: An introduction*. Amsterdam: John Benjamins.

Avram, Andrei. 2015. Reduplication in Macao Creole Portuguese and its origins. *Études Créoles* 23, 673-688.

Aye, Khin Khin. 2006. Bazaar Malay: history, grammar and contact. PhD thesis, National University of Singapore.

Aye, Khin Khin. 2013. Singapore Bazaar Malay structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds). Atlas of Pidgin and Creole Language Structures Online. Leipzig: Max Planck Institue for Evolutionary Anthropology. (Available online at http://apics-online.info/contributions/67, Accessed on 2021-12-16.)

Azevedo, Milton. 2005. *Portuguese: A linguistic introduction.* Cambridge: Cambridge University Press.

Bakker, Peter. 2014. Creoles and typology: Problems of sampling and definition. *Journal of Pidgin and Creole Languages* 29(2), 437-455.

Bakker, Peter, Aymeric Daval-Markussen, Mikael Parkvall, & Ingo Plag. 2011. Creoles are typologically distinct from non-creoles. *Journal of Pidgin and Creole Languages* 26(1), 5-42.

Balhorn, Mark. 2004. The rice of epicene *they. Journal of English Linguistics* 32(2), 79-104.

Bao, Zhiming, & Khin Khin Aye. 2012. Bazaar Malay topics. In Umberto Ansaldo (ed.). *Pidgins and creoles in Asia*, 151-167. Amsterdam: John Benjamins.

Barreiros, Leopoldo Danilo. 1943-1944. *O Dialecto português de Macau: Antologia.* Macau: Renascumento.

Batalha, Graciete. 1988. *Glossário do dialecto macaense.* Macau: Instituto cultural de Macau.

Baxter, Alan. 1988. *A grammar of Kristang (Malacca Creole Portuguese).* Canberra: Australian National University.

Baxter, Alan. 1996. Portuguese and Creole Portuguese in the Pacific and Western Pacific Rim. S. A. Wurm, P. Mühlhäusler, & D. T. Tryon (eds.). *Atlas of languages of intercultural communication in the Pacific, Asia, and the Americas*, 299-338. Berlin: Mouton de Gruyter.

Baxter, Alan. 2013. Papiá Kristang structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds). Atlas of Pidgin and Creole Language Structures Online. Leipzig: Max Planck Institue for Evolutionary Anthropology. (Available online at http://apics-online.info/contributions/42, Accessed on 2021-12-16.)

Baxter, Alan. 2018. Malacca Creole Portuguese in the 19th century: Evidence of a wider lectal range? *Journal of Pidgin and Creole Languages* 33(2), 247-279.

Baxter, Alan & Augusta Bastos. 2012. A closer look at the post-nominal genitive in Asian Creole Portuguese. In Hugo Cardoso, Alan Baxter, & Mário Pinharanda Nunes (eds.). *Ibero-Asian creoles: Comparative perspectives*, 47-49. Amsterdam: John Benjamins.

Bergström, Magnus & Neves Reis. 2011. *Prontuário ortográfico e guia da línga portuguesa.* Alfragide: Casa das Letras.

Bickerton, Derek. 1981. *Roots of language*. New York: Karoma Publishers.

Bickerton, Derek. 1984. The language bioprogram hypothesis. *The Behavioral and Brain Sciences* 7 173-188.

Bickerton, Derek. 1999. How to acquire language without positive evidence: What acquisitionists can learn from creoles. In Michel DeGraff (ed.). *Language creation and language change: Creolization, diachrony and development*. Cambridge, MA: The MIT Press.

Bodman, Nicholas. 1987. *Spoken Amoy Hokkien*. Vol 1. New York: Spoken Language Services Inc.

Bort, Balthasar. 1678 (1927). Report of Governor Balthasar Bort on Malacca 1678. (Trans. M.J. Bremmer. *Journal of the Malayan Branch of the Royal Asiatic Society* 5(1), 9-232.

Bowern, Claire. 2008. *Linguistic fieldwork: A practical guide*. London: Palgrave Macmillan.

Bruyn, Adrienne. 2008. Grammaticalization in pidgins and creoles. In Silvia Kouwenberg & John Victor Singler (eds.). *The handbook of pidgin and creole studies*, 385-410. Oxford: Wiley-Blackwell.

Bruyn, Adrienne, Pieter Muysken & Maaike Verrips. 1999. Double-object constructions in the creole languages: Development and acquisition. In Michel DeGraff (ed.). *Language creation and language change: Creolization, diachrony and development,* 329-373. Cambridge, MA: The MIT Press.

Burnet, Ian. 1972. *Spice Islands*. Kenthurst: Rosenberg Publishing.

Bybee, Joan & Östen Dahl. 1989. The creation of tense and aspect systems in the languages of the world. *Studies in Language* 13(1), 51-103.

Cabral, Carlos Alberto Anok. 2016. *Comê qui cuza?* Macau: Agência Comercial Seng Kung.

Cameron, John. [1865] 1965. *Our Tropical Possessions in Malayan India*. Oxford in Asia Historical Reprint. Kuala Lumpur: Oxford University Press.

Cardoso, Hugo. 2009. The Indo-Portuguese language of Diu. PhD thesis, University of Amsterdam.

Cardoso, Hugo. 2012. Luso-Asian comparatives in comparison. In Hugo Cardoso, Alan Baxter, & Mário Pinharanda Nunes (eds.). *Ibero-Asian creoles: Comparative perspectives,* 81-124. Amsterdam: John Benjamins.

Cardoso, Hugo. 2013. Diu Indo-Portuguese structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds). Atlas of Pidgin and Creole Language Structures Online. Leipzig: Max Planck Institue for Evolutionary Anthropology. (Available online at http://apics-online.info/contributions/39, Accessed on 2021-12-16.)

Cardoso, Hugo, Alan Baxter, & Mário Pinharanda Nunes. 2012. Introduction. In Hugo Cardoso, Alan Baxter, & Mário Pinharanda Nunes (eds.). *Ibero-Asian creoles: Comparative perspectives,* 1-14. Amsterdam: John Benjamins.

Chaudenson, Robert. 1979. *Les créoles français*. Paris: Fernand Nathan.

Chaudenson, Robert. 1992. *Des îles, des hommes, des langues*. Paris: L'Harmattan.

Chaudenson, Robert. 2000. Créolisation du français et francization du créole: Les cas de Saint-Barthélemy et de la Réunion. In Ingrid Neumann-Holzschuh & Edgar W.

Schneider (eds.). *Degrees of restructuring in creole languages,* 361-382. Amsterdam: John Benjamins.

Clements, J. Clancy. 1996. *The genesis of a language. The formation and development of Korlai Portuguese.* Amsterdam: John Benjamins.

Clements, J. Clancy. 2007. Korlai (Creole Portuguese) or Nɔ Ling. In Holm, John and Patrick, Peter (eds.), *Comparative creole syntax: Parallel outlines of 18 creole grammars*, 153-173. London: Battlebridge.

Clements, J. Clancy. 2012. Notes on the phonology and lexicon of some Indo-Portuguese creoles. In Hugo Cardoso, Alan Baxter, and Mário Pinharanda Nunes, eds. *Ibero-Asian creoles: Comparative perspectives,* 15-46. Amsterdam: John Benjamins.

Clements, J. Clancy. 2013. Korlai structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds). Atlas of Pidgin and Creole Language Structures Online. Leipzig: Max Planck Institue for Evolutionary Anthropology. (Available online at http://apics-online.info/contributions/40, Accessed on 2021-12-16.)

Coates, Austin. 1978. *A Macao Narrative.* London: Heinemann.

Coelho, F. A. 1881. *Os dialectos romanicos ou neo-latinos na África, Ásia e América.* Lisbon: Casa da Soc. de Geografia.

Crowley, Roger. 2015. *Conquerors: How Portugal forged the first global empire*. London: Faber & Faber.

Cruz-Ferreira, Madalena. 1999. European Portuguese. *Journal of the International Phonetic Association* 25(2), 90-94.

Cumming, Susanna. 1991. *Functional change: The case of Malay constituent order.* Berlin: De Gruyter.

DeGraff, Michel. 1999. Creolization, language change, and language acquisition: A prolegomenon. In Michel DeGraff (ed.). *Language creation and language change: Creolization, diachrony and development*, 1-46. Cambridge, MA: The MIT Press.

DeGraff, Michel. 2003. Against Creole Exceptionalism. *Language*, 79(2), 391-410.

DeGraff, Michel. 2005. Linguists' most dangerous myth: The fallacy of creole exceptionalism. *Language in Society*, 34(4), 533-591.

DeGraff, Michel. 2009. Creole exceptionalism and the (mis)education of the creole speaker. In Jo Anne Kleifgen & George Bond (eds.). *The languages of Africa and the diaspora: Educating for language awareness*. Bristol: Multilingual Matters.

De Silva, Shihan. 2008. *The Portuguese in the East: A cultural history of a maritime trading empire.* London: Tauris.

Detges, Ulrich. 2000. Two types of restructuring in French creoles. In Ingrid Neumann-Holzschuh & Edgar W. Schneider (eds.). *Degrees of restructuring in creole languages*, 135-162. Amsterdam: John Benjamins.

Deutscher, Guy. 2009. "Overall complexity": a wild goose chase? In Geoffrey Sampson, David Gil, & Peter Trudgill (eds.). *Language complexity as an evolving variable*, 243-251. Oxford: Oxford University Press.

Diffie, Bailey W. & George D. Winius. 1977. *Foundations of the Portuguese Empire, 1415-1580*. Minneapolis: University of Minnesota Press.

Direcção dos Serviços de Estatística e Census. 2011. Results of 2011 Population Census. Macau.

Dóci Papiaçám di Macau. 2013. *Dóci Vai Rua – Ficá Vivo na Macau*. <https://www.youtube.com/watch?v=93YUQCwejNY>

Dóci Papiaçám di Macau. 2014. *Dóci Culpa*. <https://www.youtube.com/watch?v=YVy2Z0lHd6c&t=>

Donohue, Mark. 2011. Papuan Malay of New Guinea: Melanesian influence on verb and clause structure. In Claire Lefebvre (ed.). *Creoles, their substrates, and language typology*, 413-435. Amsterdan: John Benjamins.

Eberhard, David M., Gary F. Simons, & Charles D. Fennig (eds.) 2019. *Ethnologue: Languages of the World*. Twenty-second edition. Dallas: SIL International.

Encyclopædia Britannica. Treaty of Tordesillas. Retrieved on 2019/12/04 from <https://www.britannica.com/event/Treaty-of-Tordesillas>

Eredia, Godinho de. 1613 (1930). Eredia's description of Malaca, Meridional India and Cathay, 1613 (Trans. J.V. Mills, 1930). *Journal of the Malayan Branch of the Royal Asiatic Society* 8(1), 1-228.

Faraclas, Nicholas, Don E. Walicek, Mervyn Alleyne, Wilfredo Geigel, & Luis Ortiz. 2007. Context – the complexity that really matters: The role of political economy in creole genesis. In Umberto Ansaldo, Stephen Matthews, & Lisa Lim (eds.). *Deconstructing creole*, 227-264. Amsterdam, John Benjamins.

Farquharson, Joseph. 2007. Creole morphology revisited. In Umberto Ansaldo, Stephen Matthews, & Lisa Lim (eds.). *Deconstructing creole*, 21-37. Amsterdam: John Benjamins.

Farquharson, Joseph. 2013. Jamaican structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath, & Magnus Huber (eds.). Atlas of Pidgin and Creole Language Structures Online. Leipzig: Max Planck Institute for Egolutionary Anthropology. (Available online at http://apics-online.info/contributions/8, Accessed on 2021-12-16.)

Fattier, Dominique. 2013. Haitian Creole structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath, & Magnus Huber (eds.). Atlas of Pidgin and Creole Language Structures Online. Leipzig: Max Planck Institute for Egolutionary Anthropology. (Available online at http://apics-online.info/contributions/49, Accessed on 2021-12-16.)

Fernandis, Gerard. 2000. Papia, relijang e tradisang: The Portuguese Eurasians in Malaysia: Bumiquest, a search for self identity. *Lusotopie*: 261-268.

Fernando, Radin. 2004. Metamorphosis of the Luso-Asian Diaspora in the Malay Archipelago, 1640-1795. In Peter Borschberg (ed.). *Iberians in the Singapore-Melaka Area and Adjacent Regions (16th to 18th Century)*, 161-184. Wiesbaden: Harrassowitz Verlag.

Ferreira, Susannah. 2015. *The crown, the court and the Casa da Índia: Political centralization in Portugal 1479-1521*. Leiden: Brill.

Florey, Margaret. 2004. Countering purism: confronting the emergence of new varieties in a training program for community language workers. *Language Documentation and Description 2*.

Fon Sing, Guillaume. 2018. Creoles are not typologically distinct from non-Creoles. *Language Ecology* 1(1), 44-74.

França, António Pinto da. 1995. *Portuguese influence in Indonesia*. Lisbon: Calouste Gulbenkian Foundation.

Frankopan, Peter. 2015. *The Silk Roads: A new history of the world*. London: Bloomsbury.

Ganho, Ana Sofia & Timothy McGovern. 2003. *Using Portuguese: A guide to contemporary usage*. Cambridge: Cambridge University Press.

Gil, David. 2001. Creoles, complexity, and Riau Indonesian. *Linguistic Typology* 5, 325-371.

Gil, David. 2007. Creoles, complexity and associational semantics. In Umberto Ansaldo, Stephen Matthews, & Lisa Lim (eds.). *Deconstructing creole*, 203-226. Amsteram: John Benjamins.

Goodman, Morris. 1985. Review of Roots of language, by Derek Bickerton. *International Journal of American Linguistics* 51, 109-137.

Hall, Robert. 1958. Creole Languages and Genetic Relationships. *Word* 14, 367-373.

Hall, Robert. 1966. *Pidgin and creole languages*. Ithaca, NY: Cornell University Press.

Hancock, Ian. 1969. The Malacca Creoles and their language. *Afrasian* 3, 38-45.

Hancock, Ian. 1973. Malacca Creole Portuguese: A brief transformational account. *Te Reo* 16, 23-44.

Hancock, Ian. 1975. Malacca Creole Portuguese: Asian, African or European? *Anthropological Linguistics* 17(50), 211-236.

Hao, Zhidong. 2011. *Macau: History and society*. Hong Kong: Hong Kong University Press.

Harris, Alice C. & Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.

Hartmann, Katharina & Tonjes Veenstra. 2013. Introduction. In Katharina Hartmann & Tonjes Veenstra (eds.). *Cleft structures*, 1-34. Amsterdam: John Benjamins.

Haspelmath, Martin and the APiCS Consortium. 2013a. Functions of reduplication. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath and Magnus Huber (eds.). *The atlas of pidgin and creole language structures*. Oxford: Oxford University Press.

Haspelmath, Martin and the APiCS Consortium. 2013b. Order of recipient and theme in ditransitive constructions. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath and Magnus Huber (eds.). *The atlas of pidgin and creole language structures*. Oxford: Oxford University Press.

Haspelmath, Martin and the APiCS Consortium. 2013c. Passive constructions. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath and Magnus Huber (eds.). *The atlas of pidgin and creole language structures*. Oxford: Oxford University Press.

Heng, Derek. 2009. *Sino-Malay trade and diplomacy from the Tenth through the Fourteenth Century*. Athens, OH: Ohio University Press.

Hirschman, Charles. 1986. The making of race in colonial Malaya: Political economy and racial ideology. *Cornell University* 1(2), 330-361.

Hirschman, Charles. 1987. The Meaning and Measurement of Ethnicity in Malaysia: An Analysis of Census Classifications. *The Journal of Asian Studies* 46(3), 555-582.

Holm, John. 1988-9. *Pidgins and creoles*. Cambridge: Cambridge University Press.

Holm, John. 2000. *An introduction to pidgins and creoles.* Cambridge: Cambridge University Press.

Holm, John & Peter Patrick. 2007. *Comparative creole syntax: Parallel outlines of 18 creole grammars*. London: Battlebridge.

Hopper, Paul J. & Elizabeth Closs Traugott. 2003. *Grammaticalization*. Cambridge: Cambridge University Press.

Huber, Magnus. 2012. Existential sentences with RCs in 18th and 19th Century English. A diachronic study based on the Old Bailey corpus. Paper presented at the 33rd Conference of the International Computer Archive for Modern and Medieval English (ICAME), Leuven, Belgium, 30 May-3 June 2012.

Huber, Magnus and the APiCS Consortium. 2013. Order of possessor and possessum. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath and Magnus Huber (eds.). *The atlas of pidgin and creole language structures*. Oxford: Oxford University Press.

Huet, G. 1989. La communauté portugaise de Batavia. *Revista Lusitana* 12, 149-70.

Hymes, Dell (ed). 1971. *Pidginization and creolization of languages*. Cambridge: Cambridge University Press.

Inkelas, Sharon, & Cheryl Zoll. 2005. *Reduplication: Doubling in morphology*. Cambridge: Cambridge University.

Jarkey, Nerida. 2015. *Serial verbs in White Hmong*. Leiden: Brill.

Kihm, Alain. 2011. Pidgin-creoles as a scattered sprachbund: Comparing Kriyol and Nubi. *Journal of Pidgin and Creole Languages*, 26(1), 43-88.

Koh, Kelly. 2017, February 27. Intrudu water festival makes a splash at Malacca's Portuguese Settlement. *New Straits Times*.

Koopman, Hilda. 1986. The genesis of Haitian: Implications of a comparison of some features of the syntax of Haitian, French, and West African Languages. In Pieter Muysken & Norval Smith (eds.). *Substrata versus universals in creole genesis*. Amsterdam: John Benjamins.

Kouwenberg, Silvia. 2008. The problem of multiple substrates: The case of Jamaican Creole. In Susanne Michaelis (ed.). *Roots of creole structures: Weighing the contribution of substrates and superstrates*, 1-27. Amsterdam: John Benjamins.

Kupperman, Karen Ordahl. 2012. *The Atlantic in world history*. Oxford: Oxford University Press.

Kusters, Wouter. 2003. The fate of complex languages: Classical Arabic and Old Norse in the age of globalisation. *Nordlyd* 31, 275-289.

La Batu, Safrin. 2016, March 28. Kampung Tugu tour traces Portuguese slaves, Mardijkers, history. *The Jakarta Post*.

Lalla, Barbara & Jean D'Costa. 2009. *Language in exile: Three hundred years of Jamaican Creole.* Tuscaloosa: University of Alabama Press.

Larrea, Elisabela. 2019. Bela Maquista. Retrieved on 2019/12/04 from <https://belamaquista.wordpress.com/>

Laub, Robert. 2014. Substrate influence and twentieth century decreolization in Macanese Creole. Master's thesis, SOAS University of London.

Lebel, Alexandre. 2018. Seven ways of saying *no*: The negation system in Macau Creole Portuguese. *Papia* 28(2), 157-200.

Lebel, Alexandre. 2021. Grammar of Macao Creole Portuguese with typological and semiotic considerations. PhD thesis, University of Saint Joseph, Macao.

Lefebvre, Claire. 1998. *Creole genesis and the acquisition of grammar. The case of Haitian Creole*. Cambridge: Cambridge University Press.

Lefebvre, Claire. 2008. On the principled nature of the respective contribution of substrate and superstrate languages to a creole's lexicon. In Susanne Michaelis (ed.). *Roots of creole structures: Weighing the contribution of superstrates and substrates*, 197-224. Amsterdam: John Benjamins.

Lefebvre, Claire. 2015. *Functional categories in the Atlantic creoles: Saramaccan, Haitian and Papiamentu*. Amsterdam: John Benjamins.

Lefebvre, Claire & John Lumsden. 1989. Les langues créoles et la théorie linguistique. *Revue Canadienne de Linguistique* 34, 319-377.

Li, Charles & Sandra Thompson. 1973. Serial verb constructions in Mandarin Chinese: subordination or coordination? CLS Comparative Syntax Festival.

Lim, Lisa, and Umberto Ansaldo. 2013. Singlish structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds). Atlas of Pidgin and Creole Language Structures Online. Leipzig: Max Planck Institue for Evolutionary Anthropology. (Available online at http://apics-online.info/contributions/21, Accessed on 2021-12-16.)

Luke, K.K., Adams Bodomo & Owen Nancarrow. 2001. The subject condition in Cantonese. Workshop on LFG approaches to Chinese, LFG2001.

Lumsden, John. 1999. Language acquisition and creolization. In Michel DeGraff (ed.). *Language creation and language change: Creolization, diachrony and development,* 129-158. Cambridge, MA: The MIT Press.

Lüpke, Friederike. 2005. Small is beautiful—the contribution of small field-based corpora to different linguistics disciplines, illustrated by Jalonke. *Language Documentation and Description* 3, 75-105.

Maiden, Martin. 2010. Morphophonological innovation. In Martin Maiden, John Charles Smith, & Adam Ledgeway (eds.). *The Romance languages volume 1: Structures*. Cambridge: Cambridge University Press.

Marbeck, Joan. 1995. *Ungua Adanza*. Lisbon: J. M. Marbeck.

Marbeck, Joan. 2004. *Linggu Mai*. Lisbon: Calouste Gulbenkian Foundation.

Matras, Yaron. 2010. Contact, convergence, and typology. In Raymond Hickey (ed.). *The handbook of language contact*, 66-85. Oxford: Wiley-Blackwell.

Matthews, Stephen. 2006. On serial verbs in Cantonese. In Alexandra Aikhenvald & R. M. W. Dixon (eds.). *Serial verb constructions: A cross-linguistic typology*, 69-87. Oxford: Oxford University Press.

Matthews, Stephen & Virginia Yip. 1994. *Cantonese: A comprehensive grammar*. London: Routledge.

Maurer, Philippe. 2011. *The former Portuguese creole of Batavia and Tugu (Indonesia)*. London: Battlebridge.

Maurer, Philippe and the APiCS Consotrium. 2013. 'Give' serial verb constructions. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath and Magnus Huber (eds.). *The atlas of pidgin and creole language structures*. Oxford: Oxford University Press.

Maurer, Philippe and the APiCS Consortium. 2013. 'Take' serial verb constructions. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath and Magnus Huber (eds.). *The atlas of pidgin and creole language structures*. Oxford: Oxford University Press.

Maurer, Philippe, Susanne Maria Michaelis and the APiCS Consortium. 2013. Directional serial verb constructions with 'come' and 'go'. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath and Magnus Huber (eds.). *The atlas of pidgin and creole language structures*. Oxford: Oxford University Press.

MacGregor, Ian A. 1955. Notes on the Portuguese in Malaya. *Journal of the Malaysian Branch of the Royal Asiatic Society*, 28(2), 4-47.

McNair, Major J.F.A. 1878. *Perak and the Malays: Sarong and Kris*. London.

McPherson, Kenneth. 1993. *The Indian Ocean: A history of people and the sea*. Oxford: Oxford University Press.

McWhorter, John. 1999. The Afrogenesis of plantation creole origin. In Magnus Huber & Mikael Parkvall (eds.). *Spreading the word: The issue of diffusion among the Atlantic Creoles*. London: University of Westminster Press.

McWhorter, John. 2001a. Defining "creole" as a synchronic term. In Ingrid Neumann-Holzschuh & Edgar W. Schneider (eds.). *Degrees of restructuring in creole languages*, 85-124. Amsterdam: John Benjamins.

McWhorter, John. 2001b. The world's simplest grammars are creole grammars. *Linguistic Typology*, 5(2/3), 125-166.

McWhorter, John. 2005. *Defining Creole*. New York: Oxford University Press.

McWhorter, John. 2018. *The Creole debate.* Cambridge: Cambridge University Press.

Merewether, E. M. 1892. *Report on the Census of the Straits Settlements, Taken on the 5ᵗʰ April 1891.* Singapore: Government Printing Office.

Meyerhoff, Miriam. 2013. Bislama structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath, & Magnus Huber (eds.). Atlas of Pidgin and Creole Language Structures Online. Leipzig: Max Planck Institute for Egolutionary Anthropology. (Available online at http://apics-online.info/contributions/23, Accessed on 2021-12-16.)

Michaelis, Susanne (ed.). 2008. *Roots of Creole structures: weighing the contribution of substrates and superstrates*. Amsterdam: John Benjamins.

Michaelis, Susanne Maria, Philippe Maurer, Martin Haspelmath and Magnus Huber (eds.). 2013. Atlas of Pidgin and Creole Language Structures Online. Leipzig: Max Planck Institute for Evolutionary Anthropology. <https://apics-online.info/>

Miestamo, Matti. 2008. Grammatical complexity in a cross-linguistic perspective. In Matti Miestamo, Kaius Sinnemäki, & Fred Karlsson (eds.). *Language complexity: Typology, contact, change*, 23-41. Amsterdam: John Benjamins.

Migge, Bettina & Laurence Goury. 2008. Between contact and internal development: Towards a multi-layered explanation for the development of the TMA system in the creoles of Suriname. In Susanne Michaelis (ed.). *Roots of creole structures: Weighing the contribution of substrates and superstrates*, 301-332. Amsterdam: John Benjamins.

Miller, Tom. 2017. *China's Asian Dream*. London: Zed Books.

Mintz, Malcolm. 1994. *A student's grammar of Malay & Indonesian.* Singapore: EPB Publishers.

Moseley, Christopher (ed). 2010. Atlas of the World's Languages in Danger, 3rd edn. Paris, UNESCO Publishing.

Mufwene, Salikoko. 2000. Creolization is a social, not a structural process. In Ingrid Neuman-Holzschuh & Edgar Schneider (eds.). *Degrees of restructuring in creole languages*, 65-84. Amsterdam: John Benjamins.

Mufwene, Salikoko. 2002. Competition and selection in language evolution. *Selection* 3, 45-56.

Mühlhäusler, Peter. 1986. *Pidgin and creole linguistics*. Oxford: Basil Blackwell.

Mühlhäusler, Peter. 2012. The complexity of the personal and possessive pronoun system. In Bernd Kortmann & Benedikt Szmrecsanyi (eds.). *Linguistic complexity: Second language acquisition, indigenization, contact*, 101-126. Berlin: Mouton De Gruyter.

Muysken, Pieter. 1988. Parameters for serial verbs. In Katherine Demuth & Victor Manfredi (eds.). *Niger-Congo syntax and semantics 1*, 65-75. MS, Boston University.

Muysken, Pieter & Norval Smith. 1986. *Substrata versus Universals in Creole Genesis: Papers from the Amsterdam Creole Workshop, April 1985*. Amsterdam: John Benjamins.

Muysken, Pieter & Tonjes Veenstra. 1994. Haitian. In Jacques Arens, Pieter Muysken, & Norval Smith (eds.). *Pidgins and creoles: An introduction*, 153-164. Amsterdam: John Benjamins.

Newitt, Malyn. 2005. *A history of Portuguese overseas expansion 1400-1668*. London: Routledge.

Nordhoff, Sebastian (ed.). 2012. *The genesis of Sri Lanka Malay*. Leiden: Brill.

Nomoto, Hiroki. 2006. Voice in colloquial Malay relatives. In Yoichiro Tsuruga, Nobuyuki Miyake, & Yuji Kawaguchi (eds.). *Gengojouhougaku Kenkyuuhoukobu 12*, 97-116. Tokyo: Tokyo University of Foreign Studies.

Nieuwe Woordenschat, uyt het Nederduitsch in het gemeene Maleidsch en Portugeesch, zeer gemakkelyk voor die eerst op Batavia komen. 1780. Te Batavia, By Lodewyk Dominicus, Stats-Drukker, op de Tygers-Gragr, aan de West-Zyde.

Niew, Shong Tong. 1969. The population geography of the Chinese communities in Malaysia, Singapore and Brunei. PhD thesis, SOAS University of London.

Ostler, Nicholas. 2005. *Empires of the World: A language history of the World*. New York: Harper Collins.

Pakir, Anne. 1986. A linguistic investigation of Baba Malay. PhD thesis, University of Hawai'i.

Parkvall, Mikael. 2000. Reassessing the role of demographics in language restructuring. In Ingrid Neumann-Holzschuh & Edgar W. Schneider (eds.). *Degrees of restructuring in creole languages*, 185-213.

Parkvall, Mikael. 2008. The simplicity of creoles in a cross-linguistic perspective. In Matti Mestamo, Kaius Sinnemäki & Fred Karlsson (eds.). *Language complexity: typology, contact, change*, 265-285. Amsterdam: John Benjamins.

Parthesius, Robert. 2010. *Dutch ships in tropical waters: The development of the Dutch East India Company (VOC) shipping network in Asia 1595-1660*. Amsterdam: Amsterdam University Press.

Pawley, Andrew & Jonathan Lane. 1998. From event sequence to grammar: serial verb constructions in Kalam. In Anna Siewierska & Jung Song Jae (eds.). *Case, typology, and grammar*, 201-227. Amsterdam: John Benjamins.

Pereira, Alexius. 2006. No longer "other": The emergence of the Eurasian Community in Singapore. In Kwen Fee Lian (ed.). *Race, ethnicity, and the state in Malaysia and Singapore*. Leiden: Brill.

Pereira, João F. M. 1899-1900. *Ta-ssi-yang-kuo: Archivos e annaes do Extremo-Oriente português.* 1(1-2). Lisbon: Antiga Licraria Bertrand.

Pillai, Patrick. 2015. *Yearning to belong: Malaysia's Indian Muslims, Chitties, Portuguese Eurasians, Peranakan Chinese and Baweanese*. Singapore: ISEAS-Yusof Ishak Institute.

Pillai, Stefanie. 2011. Malaccan Portuguese Creole: A Portuguese-Based Creole. Endangered Languages Archive at SOAS University of London.
<https://elar.soas.ac.uk/Collection/MPI130545>

Pinharanda Nunes, Mário. 2008. Os demonstrativos em maquista: uma análise morfossintáctica contrastiva. In *Papia, Revista Brasileira de Estudos Crioulos e Similares* 18, 8-22.

Pinharanda Nunes, Mário. 2012. Traces of superstrate verb inflection in Makista and other Asian-Portuguese creoles. In Hugo Cardoso, Alan Baxter, & Mário Pinharanda Nunes (eds.). *Ibero-Asian creoles: Comparative Perspectives*, 291-326. Amsterdam: John Benjamins.

Pinharanda Nunes, Mário, & Alan Baxter. 2004. Os marcadores pré-verbais no crioulo de base lexical portuguesa de Macau. *Papia, Revista Brasileira de Estudos Crioulos e Similares* 13, 31-46.

Population and Housing Census. Malaysia, 2010.

Sadock, Jerrold M. & Zwicky, Arnold M. 1985. Speech acts distinctions in syntax. In Timothy Shopen (ed.). *Language typology and syntactic description*, 155-196. Cambridge: Cambridge University Press.

Santos, Ana Lúcia. 2002. The acquisition of answers to yes-no questions in European Portuguese: Syntactic, discourse and pragmatic factors. *Journal of Portuguese Linguistics* 2, 61-91.

Santos Ferreira, José. 1967. *Macau Sã Assi*. Macau: Tipografia da Missão Padroado.

Santos Ferreira, José. 1973a. *Qui-Nova Chencho*. Macau: Tipografia da Missão do Padroado.

Santos Ferreira, José. 1973b. *Poéma di Macau*. Macau: s.n.

Santos Ferreira, José. 1990. *Macau di Tempo Antigo: Poesia e Prosa, Dialecto Macaense.* Macau: s.n.

Schuchardt, Hugo. 1883. Über das Indoportugiesische von Diu. *Kreolische Studien* 3, 3-17.

Schuchardt, Hugo. 1889. Allgemeineres über das Indoportugiesische. *Zeitschrift für Romanische Philologie* 13, 476-516.

Schuchardt, Hugo. 1890. Über das Malaioportugiesische von Batavia und Tugu. *Kreolische Studien* 9, 255.

Sebba, Mark. 1987. *Syntax of serial verbs: An investigation into serialization in Sranan and other languages.* Amsterdan: John Benjamins.

Selbach, Rachel. 2008. The superstrate is not always the lexifier: Lingua Franca in the Barbary Coast 1530-1830. In Susanne Michaelis (ed.). *Roots of creole structures: Weighing the contribution of substrates and superstrates*, 29-58. Amsterdam: John Benjamins.

Shi, Yuzhi. 2002. *The establishment of modern Chinese grammar. The formation of the resultative construction and its effects.* Amsterdam: John Benjamins.

Siegel, Jeff. 1997 Mixing, leveling, and pidgin/creole development. In Arthur K. Spears & Donald Winford (eds.). *The structure and status of pidgins and creoles*, 111-149. Amsterdam: John Benjamins.

Siegel, Jeff. 2007. Sociohistorical contexts: Transmission and transfer. In Umberto Ansaldo, Stephen Matthews, & Lisa Lim (eds.). *Deconstructing creole*, 167-202. Amsterdam: John Benjamins.

Siegel, Jeff. 2008a. *The emergence of pidgin and creole languages*. Oxford: Oxford University Press.

Siegel, Jeff. 2008b. In praise of the cafeteria principle: Language mixing in Hawai'i Creole. In Susanne Michaelis (ed.). *Roots of creole structures: Weighing the contribution of substrates and superstrates*, 59-82. Amsterdam: John Benjamins.

Singho, Michael. et al. 2017. *Beng prende Portugues Malaká (Papiá Cristang): Come, let's learn Portugues Malaká (Papiá Cristang).* Kuala Lumpur: University of Malaya Press.

Sippola, Eeva. 2012. Indefinite terms in Ibero-Asian Creoles. In Hugo Cardoso, Alan Baxter, & Mário Pinharanda Nunes (eds.). *Ibero-Asian creoles: Comparative perspectives,* 149-180. Amsterdam: John Benjamins.

Sit, Victor F. S., Vivian Hui, & Gladys Leung. 2012. *Macau through 500 years*. Silkroad Press.

Smith, Ian. Unpublished field notes from Batticaloa 1973-4.

Smith, Ian. 2012. Measuring substrate influence: Word order features in Ibero-Asian Creoles. In Hugo Cardoso, Alan Baxter, & Mário Pinharanda Nunes (eds.). *Ibero-Asian creoles: Comparative perspectives,* 125-148. Amsterdam: John Benjamins.

Smith, Ian. 2013. Sri Lanka Portuguese structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds). Atlas of Pidgin and Creole Language Structures Online. Leipzig: Max Planck Institue for Evolutionary Anthropology. (Available online at http://apics-online.info/contributions/41, Accessed on 2021-12-16.)

Stewart, William. 1965. Urban Negro speech: Sociolinguistic factors affecting English teaching. In R. Shuy, A. Davis, & R. Hogan (eds.). *Social dialects and language learning*, 10-18. National Council of Teachers of English.

Sun, Chaofen. 2006. *Chinese: A Linguistic Introduction*. Cambridge: Cambridge University Press.

Szmrecsanyi, Benedikt & Bernd Kortmann. 2012. Introduction: Linguistic complexity – Second language acquisition, indigenization, contact. In Bernd Kortmann & Benedikt Szmrecsanyi (eds.). *Linguistic complexity: Second language acquisition, indigenization, contact*, 6-34. Berlin: Mouton De Gruyter.

Thomason, Sarah. 2001. *Language contact*. Edinburgh: Edinburgh University Press.

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

Ulung, A. Kurniawan & Gyanindra Ali. 2019, January 18. Kampung Tugu: Hidden Portuguese gem in North Jakarta. *The Jakarta Post*.

UNESCO. Network of the Silk Road Cities. Retrieved on 2019/12/04 from
<https://en.unesco.org/silkroad/network-silk-road-cities-map-app/en>

University of Potsdam. Questionnaire for Information Structure.

Valdman, Albert. 1978. *Le créole: Structure, statut et origine*. Paris: Klincksieck.

Vinson, Julien. 1889. Créoles. In Adolphe Bertillion et al, eds. *Dictionnaire des sciences anthropologiques*, 345-347. Paris: Doin.

Wallace, Stephen. 1978. What is a Creole? The example of the Portuguese language of Tugu, Jakarta, Indonesia. Suñer, M. (ed). *Contemporary studies in Romance linguistics*, 340-377. Washington, DC: Georgetown University Press.

Weinreich, Uriel. 1953. *Languages in contact: Findings and problems.* The Hague: Mouton.

Whinnom, Keith. 1971. Linguistic hybridization and the 'special case' of pidgin and creoles. In Dell Hymes (ed). *Pidginization and creolization of languages*, 91-115. Cambridge: Cambridge University Press.

Williams, Edwin. 1962. *From Latin to Portuguese: Historical phonology and morphology of the Portuguese language.* Philadelphia: University of Pennsylvania Press.

Yip, Virginia & Stephen Matthews. 1999. *Basic Cantonese: A grammar and workbook*. London: Routledge.

Yip, Virginia & Stephen Matthews. 2000. *Intermediate Cantonese: A Grammar and Workbook.* Milton Park: Taylor & Francis Group.

Youssef, Valerie. 1988. The language bioprogram hypothesis revisited. *Journal of Child Language*, 15: 451-458.

Youssef, Valerie. 1990. On the confirmation of bioprograms. *Journal of Child Language*, 17: 233-235.

Yunus, Maris M. 1980. *The Malay sound system*. Kuala Lumpur: Penerbit Fajar Bakti Sdn. Bhd.

Zimmerman, Malte & Caroline Féry. 2002. *Information structure: Theoretical, typological, and experimental perspectives*.

Appendix 1: Consent form

Participant Consent Form

Robert W. Laub
Doctoral Student
SOAS, University of London

Title of Research Project: A Comparison of the Portuguese Creoles of Malacca and Macau
Introduction
The purpose of this form is to provide you with information so you can decide whether to participate in this study. Any questions you may have will be answered by the researcher or by the other contact persons provided below. Once you are familiar with the information on the form and have asked any questions you may have, you can decide whether or not to participate. If you agree, please either sign this form or else provide verbal consent if you do not wish your name to be registered on the form. Please also indicate whether or not you are willing for your contribution to be audio recorded. Please note that this recording, a transcript of it and any relevant notes may be kept to be re-used in a future project.

Please note your participation is voluntary and you may decide to leave the study at any time. You may also refuse to answer specific questions you are uncomfortable with. You may withdraw permission for your data to be used, at any time up to September 2019 in which case notes, transcriptions and recordings will be destroyed.

Purpose of the Study
You have been asked to participate in a research study about Malacca Portuguese Creole or Cristang. The purpose of the research is to examine the grammatical features of the language, and how it relates to the Portuguese Creole spoken in Macau, as well as Malay and other varieties of the Portuguese language.

Use of the data
The findings will be used to form part of my thesis and will potentially additionally be published in relevant journals or archives.
If you wish to receive a copy of the final thesis once completed I will happy to provide you with an electronic or physical copy.

Procedures to be followed

To assist my research I am asking you to agree to participating in a recorded interview or elicitation. We can arrange a time and date which is convenient to you once you have confirmed your consent. The recording will take no longer than one hour to complete. If necessary, additional recordings may be requested.

Compensation

You will receive a token as a favour for your participation. Please see the Honorarium Reception Form for details.

Statement of Privacy and Confidentiality

In any publication based on the findings of this study, the data presented will contain no identifying information that could associate it with you unless you specifically request to have your real name associated with your responses.

Contact Information

My telephone number is: +601124081355 (SMS)
                                        +447946176765 (WhatsApp)

My email address is: 602735@soas.ac.uk

Alternatively, you may wish to contact my supervisor, Professor Peter Austin at pa2@soas.ac.uk

Confirmation and consent

I confirm that I have freely agreed to participate in the research project of Robert W. Laub. I have been briefed on what this involves and I agree to the use of the findings as described above. I give/ do not give permission for the interview to be recorded.

I understand the recording, transcript and any notes may be archived and put on line for public use. In this case I give my permission for this including data that might identify me/provided all possible identification is removed.

Copyright Statement
By completing this form, you permit Robert W. Laub to edit, copy, disseminate, publish (by whatever means) and archive your contribution to this research project in the manner and for the purposes described above. You waive any copyright and other intellectual property rights in your contribution to the project, and grant SOAS, the project coordinator and other researchers a non-exclusive, free, irrevocable, worldwide license to use your contribution for the purposes of this project and similar future research projects.

**Participant
signature:**_____

Name:_____
__

Date:_____
__

I confirm that I agree to keep the undertakings in this contract.

Researcher signature:_____

Name:_____ ____

Date:_____ ____

Please keep this form for future reference.