Immigrant speech, creoles, and the 'basic variety': a usage-based account of some traits in the Portuguese-based creoles

J. CLANCY CLEMENTS

Abstract

In order to account for some key structures in immigrant Spanish, 15th-16th-century African Portuguese, pidgin Portuguese, and the Portuguese-based creoles, we appeal to a model of Emergent Grammar in which linguistic structure emerges from frequency and perceptual saliency. We argue that key traits of the Portuguese-based creoles evolved from L2 varieties of Portuguese initially spoken by Africans in Portugal in the 15th and 16th centuries. First, we show how certain verb forms make their way into the speech of immigrants due to frequency of occurrence of verb forms based on their respective verb class. We also note that some forms become part of immigrant speech because of their more perceptually salient form (e.g. son *over* es, *and* são *over* é). *Second, we point out that the many traits that characterize Chinese immigrant Spanish also define African Portuguese, and are those that also have found their way into the Portuguese-based creoles. We claim that regardless of whether pidgin Portuguese had its beginnings in Portugal or emerged in the interaction between the Portuguese and Africans in Africa, the same relative frequency of occurrence of forms would apply, as would the same processes used in naturalistic L2 acquisition, including sensitivity on the part of the speakers/learners to perceptual saliency of items. Thus, this study contributes to the understanding of the role that naturalistic L2 acquisition played in the formation of the pidgin Portuguese and, subsequently, the Portuguese-based creoles.*

Introduction

Henry the Navigator's first voyage to Africa in 1417 initiated a Portuguese colonial expansion that by 1530 had reached India, Indonesia, China, and Japan. And although the Portuguese domination in the colonization of Africa and Asia began to wane in the 17th century, the contact with peoples of these

parts of the world from 1417 to 1620 changed the face of Portuguese society significantly.

In the first half of the 15th century, word of the Portuguese maritime exploits traveled quickly through the rural areas of Portugal, and it was not long before thousands of the landless farmers found their way to Lisbon and on to ships heading abroad, attracted by the possibilities of wealth and adventure. Boxer reports that during the 16th century alone, from 2000 to 4000 Portuguese men traveled overseas to the Portuguese colonies, the majority 'being able-bodied and unmarried young men, bound for Golden Goal and further east, relatively few of whom ever returned to Europe' (1975:67).

At the same time that Portuguese men were lining up to venture out to the newly-founded Portuguese colonies, Africans were being brought to Portugal as slaves. Based on data collected by Godinho (1981-1983), Ramos Tinhorão (1997:101, 113) notes that from around 1450 to 1505, sub-Saharan Africans came to make up more than 10% of the rural population in Portugal south of the Douro river. And by 1538, Lisbon had a population of around 100,000 inhabitants, of which sub-Saharan Africans made up around 15% (15,350) of that total. These African slaves filled the shortage of labor created by the exodus of young Portuguese men to the colonies. The Africans not only toiled in agriculture in the rural areas, but also worked as interpreters in the courts and on ships, as domestics, and as workers in small businesses (water sellers, goods transporters, street cleaners, etc.). And over time, some Africans also came to be performers, crafts people, etc.

The elevated numbers of Africans living in urban areas such as Lisbon and Porto naturally had a significant impact on Portuguese culture and language. In this study, we are interested in the variety of Portuguese that the Africans may have spoken during that period. Fortunately, there are portrayals of the Africans' speech in plays of the time, by playwrights such as Anrique da Mota and Gil Vicente. We examine the traits displayed in such portrayals, comparing them to what we know about immigrant speech today. Klein and Perdue (1992) have shown that, independently of the languages involved, it is possible to speak of a 'basic variety' of immigrant speech, which features are found in modern-day immigrant Spanish, as well as in the African variety of Portuguese found in plays. We account for this by appealing to a usage-based model, the description of which we turn to now.

Aspects of a usage-based model

The varieties of immigrant speech to be discussed in this study are examples of language change, more particularly, contact-induced language change. We assume that language change is detected and studied most accurately at its different stages by studying language use (accessible in language data bases). This is done through a number of approaches, such as discourse analysis, variationist   analysis,   computational   linguistics,   corpus   linguistics,

construction-based models, and models of language acquisition. All these approaches are largely compatible with one another and share key features: they make use of corpora (i.e. examples of natural language use) and they afford a key role to frequency of occurrence of items or sets of items for the creation of linguistic structure. That is, repetition is important. Part of this study involves the use of a data bases to test usage-based certain hypotheses of second language acquisition, to be discussed below.

This study assumes a type of Emergentist Grammar (EG) model. We subscribe to a somewhat weaker version of the EG model than that advanced in Hopper (1987, 1988, 1998). Hopper assumes that language is indeterminate, constantly under construction, and structured by the emergent patterns that come and go as the forms that carry them are found useful for their speakers, and that it consists of different kinds of repetition, some of which concern the lexical, some idiomatic, and some grammatical (Hopper 1998: 158, 172). While we agree that language is constantly evolving, we assume, with Goldberg (1999: 200), that grammar emerges primarily during acquisition (either L1 or L2), from a combination of linguistic input, the functional demands of communication, and general cognitive abilities and constraints. We also assume that once grammar is acquired, it is conventionalized, that is, it is fairly stable as a system.

Bybee addressed the role of frequency in the formation of grammar as early as 1985, in her book *Morphology*. There she made the distinction between frequency proper of a lexical item in discourse, which she called its 'lexical strength', and the frequency of a common form in a paradigm, which she called 'lexical connections' (cf. Bybee, 1985: 123-27). Here we include an example of lexical connections that will be relevant for our later discussion. The present-tense copula forms in Portuguese and Spanish, shown in Table 1, are irregular.

Table 1: Lexical Connections of Portuguese and Spanish *ser* 'be'

Portuguese		*Spanish*	
C V(G/C)	V(C)	CV(C)	V(C)
s o u 'I am'	é s 'you.SG are'	s o y 'I am'	e r e (s) 'you.SG are'
| |	|	| |	| |
s o m o s 'we are'	é 's/he/it is'	s o m o s 'we are'	e (s) 's/he/it is'
| |		| |	
| |		(s o i s) 'you.PL are'[1]	
| |		| |	
s ã o 'they are'		s o n 'they are'	

---

[1] This form is found only in the variety of Spanish spoken in Spain.

Important in Table 1 is that the most frequent forms are not the most perceptually salient ones from the perspective of processing. That is, although the 3SG form (Portuguese *é* and Spanish *e(s)*) is the most frequently occurring form, the more perceptually salient ones are those that have CV(C) structure. That is, given their V(C) structure, the vocalic realizations [e], [eh], and [es] are not as perceptually salient as those forms with CV(C) structure.[2]

The usage-based EG model we assume encompasses certain hypotheses about acquisition. For our purposes, we will concentrate here on two hypotheses regarding the L2 acquisition of verb forms: The Primacy of Aspect Hypothesis (POA) and the Distribution Bias Hypothesis (DBH). Based on lexical semantics and verb form meaning, Andersen and Shirai (1996: 529) formulated the POA to predict for L2 acquisition which forms of which verbs are learned first. They state that:

> '[I]t has been consistently observed that L1 and L2 learners, in the early stages of acquiring verbal morphology, use tense-aspect selectively according to the inherent lexical aspect of the verb to which the tense-aspect marker is attached or with which it is associated.

They propose the POA to account for these facts (Andersen and Shirai, 1996: 559):

1. Learners will initially restrict past or perfective marking to telic verbs (achievements and accomplishments), and later gradually extend the marking to atelic predicates (activity and then stative predicates), with states being the last category to be marked consistently;

2. In languages with an imperfective marker, imperfective past appears much later than perfective past and then is initially restricted to atelic predicates (stative and activity verbs), then extended to telic predicates (to accomplishments and then achievements);

3. Progressive marking is initially restricted to dynamic atelic predicates (activity verbs), then extended to telic predicates (to accomplishments and then achievements);

4. Progressive marking is not incorrectly overextended to states.

---

[2] As an example, the Spanish forms in the contemporary Spanish database maintained by the Real Academia Española (cf. www.rae.es) are as follows. Apart from the 5% differential between the 3SG and 3PL forms, the relative distribution overall and in oral speech is comparable.

Form	Overall distribution	Distribution in oral speech
soy	23756 (2%)	2951 (2.5%)
ere(s)	9767 (.9%)	1051 (.95%)
e(s)	978813 (78%)	98840 (82%)
somos	12563 (1%)	1660 (1.5%)
sois	1085 (.1%)	117 (.05%)
son	223417 (18%)	15638 (13%)

The POA is related to another hypothesis advanced by Andersen (1993: 320), the DBH, which his based on frequency of occurrence of verb forms:

> Native speakers will tend to use past or perfective inflections more with telic and punctual events than with states or activities, progressive inflections primarily with activities and imperfective inflections more with states and activities than with telic and punctual events.

To our knowledge, this second hypothesis has not been tested by analyzing key verbs in a database. The hypothesis makes the prediction that certain verb forms will appear with certain verbs more frequently than with others because of the semantic nature of the verb and the aspect of the form. Before we test the hypothesis on immigrant speech, we can verify whether the DBH makes the correct predictions for key verbs in oral speech of a language. For reasons that will become clear below, we have analyzed oral speech of Spanish from Madrid for the following verbs, grouped into their verb classes: States: *saber* 'know', *querer* want (love)', *poder* 'be able'; Activities: *trabajar* 'work', *buscar* 'look for'; Achievements: *llegar* 'arrive', *nacer* 'be born', *morir* 'die'. We have purposely excluded data for Accomplishments because they often involve complex (i.e. verb + object/adjunct) rather than just simple (i.e. only verb) predicates. In Table 3, we show the distribution of forms for each of the verb based on whether the form involves imperfective or perfective aspect.

Table 2: The distribution of forms for representative State, Activity, and Achievement verbs

	States			Activities		Achievements		
	saber	poder	querer	trabajar	buscar	morir	nacer	romper
PRESENT								
1sg	5288	1275	1618	46	24	15	1	7
2sg	1308	814	635	61	16	14	1	5
3sg	1130	4550	1392	222	103	88	56	42
1pl	444	861	338	40	29	3	10	2
3pl	512	1129	495	84	38	42	29	19
Present Subtotal	*8682*	*8629*	*4478*	*453*	*210*	*162*	*97*	*75*
PAST								
1/3sg-pret	12/38	73/157	41/67	31/25	5/13	5/159	48/148	5/36
2sg-imp/pret	43/4	76/2	49/1	9/1	2/1	2/2	1/10	0/0
1/3sg-imp	394	667	875	86	17	17	9	13
1pl-imp/pret	53/10	104/38	59/7	145/1	4/0	2/0	0/4	1/1
3pl-imp/pret	52/14	137/40	111/21	20/8	5/5	10/30	1/16	4/9
Past Subtotal	*620*	*1294*	*1231*	*326*	*52*	*227*	*237*	*69*
NON-FIN.								
gerund	51	19	25	449	211	29	4	10
infinitive	813	439	109	620	349	186	58	87
Non-fin Subtotal	*864*	*458*	*134*	*1069*	*560*	*215*	*62*	*97*
TOTAL	10,166	10,381	5843	1848	822	604	396	241

Table 3: Distribution of verb forms (raw numbers and percentages) according to imperfective (present, imperfect) and perfective (preterit) verb forms.

	States			Activities		Achievements		
	saber	poder	querer	trabajar	buscar	morir	nacer	romper
Imperfective	9224 (99)	9613 (96.9)	5572 (97.6)	713 (91.5)	238 (90.8)	193 (49.6)	108 (32)	93 (64.6)
Perfective	78 (1)	310 (3.1)	137 (2.4)	66 (8.5)	24 (9.2)	196 (50.4)	226 (68)	51 (35.4)

In Table 4, we show the distribution of non-finite forms for each of the chosen verbs (i.e. infinitives and gerunds). Here we take the infinitive form to be neutral with respect to imperfective or perfective grammatical aspect. The gerund form, however, is imperfective.

Table 4: The distribution (raw numbers and percentages) of non-finite forms for the verbs chosen.

Verb Form	saber	poder	querer	trabajar	buscar	morir	nacer	romper
Gerund	51 (6%)	19 (4%)	25 (19%)	449 (42%)	211 (38%)	29 (13%)	4 (6%)	10 (10%)
Infinitive	813 (94%)	439 (96%)	109 (81%)	620 (58%)	349 (62%)	186 (87%)	58 (94%)	87 (90%)

The data presented in Tables 2-4 bear out the predictions made by the DBH: States and Activities, which are atelic, appear overwhelmingly in imperfective forms (over 96% for States, over 90% for Activities). Achievements, which are not only telic but also punctual, appear much more frequently than Activities or States in perfective forms (51.8% for Achievements vs. 4.8% for States and Activities).

As for the distribution of non-finite forms, the POA predicts that in L2 acquisition the progressive form will not be overextended to States. This is not surprising given that the input is not here: the representative State verbs chosen for this analysis rarely appear in the gerund form (6%). Achievement verbs, being punctual, though not incompatible with the gerund form, appear in it infrequently. Activity verbs, on the other hand, show up the most frequently of all classes in the gerund form. This is due to the facts that Activities are dynamic situations (as opposed to States) and atelic (as opposed to Achievements, which are telic). We see, then, that the predictions made by the

DBH are borne out by the distributional data of verbs representing the verb classes of States, Activities, and Achievements. Moreover, the predictions made by the POA follow from the data in Tables 2-4. This suggests that the DBH is possibly the basis for the POA.

Having seen what the distribution of forms looks like in oral speech of a language, let us now turn to the distribution of these forms, as well as other traits, in modern-day immigrant speech. Once we have seen what modern-day immigrant speech looks like, we can then compare it to the portrayals of speech we find of Africans in 15th-16th century Portugal to check for similarities.

Chinese Immigrant Spanish and its similarities to African Portuguese of the 15th-16th centuries

It may seem odd that we have chosen for this study a comparison of 15th-16th--century African Portuguese with 20th-21st-century Chinese immigrant Spanish. We justify this comparison by noting that the verb forms of Portuguese and Spanish we are dealing with are largely identical in terms of their respective grammatical aspect. Moreover, in other ways Portuguese and Spanish are largely comparable for what we propose to do. As for the relation between African languages and Portuguese on the one hand, and Chinese and Spanish on the other, two points are noteworthy. Both the West African languages and Chinese are substantially different from Portuguese and Spanish, respectively. Second, we pointed out earlier the observation made by Klein and Perdue (1992) that immigrant language varieties display enough properties in common that one can speak of a basic immigrant variety of a language.

After studying many cases of immigrant language involving different European languages, Klein and Perdue (1992: 315) conclude that most immigrants acquire a basic, naturalistically learned second-language (L2) variety. They found that, independently of the linguistic background and the target language of the learners, the varieties of language, though not identical, were found to be remarkably similar. They consider this a stage of development and note that it can be transcended if there is ongoing influx of language, i.e., the acquisition of an increasing amount of vocabulary. This in turn would arguably support the acquisition of more grammar (i.e., tense--aspect markers, grammatical relation markers and other corresponding syntactic structures). Some of the traits common to the basic variety of immigrant speech they studied were the following: in the development from a nominal utterance organization (NUO) to an infinitive utterance organization (IUO) and finally to a finite-verb utterance organization (FUO) of sentences, the basic variety is located at the IUO stage, in which largely an invariant verb form is used. At this stage, other inflectional markers, such as pluralization, etc. are most likely also inconsistent. The comment by Klein and Perdue

(1992: 315) that '[i]t is perhaps not accidental that we find essentially the same patterning in pidgin languages,' is revealing, as we also have noted the same connection elsewhere (Clements, 2003) and in this chapter hope to show that connection between what we consider L2 varieties of 15th-16th century Portuguese portrayed in plays of the time and pidgin Portuguese, which we argue served as the basis of the Portuguese-based creoles.

We have analyzed the speech of two Chinese immigrants living in Madrid who have learned Spanish naturalistically. Jenny came to Madrid in 1985, in her late twenties. Born in Nanking, Jenny is a speaker of Mandarin Chinese. She knew she wanted to emigrate and thus learned Chinese massage and acupuncture before leaving her homeland. She knew no Spanish upon arriving in Madrid, and spent the first nine years working in the restaurant industry where she had little contact with Spanish speakers. After that, she found a job as a professional manicurist, masseuse and acupuncturist, and now works and interacts mainly with Spanish speakers. She maintains little contact with the Chinese community in Madrid, preferring instead to spend her time with Spaniards.

During this formative stage of language development, she acquired neither articles nor clitic pronouns *(lo, los, la las, le, les, se,* nor features such as gender and number agreement. Since 1994, she has increasingly become integrated into Spanish society, and as a consequence her vocabulary is substantial. Even though the feature pool she has access to contains largely Spanish lexical material and structure, her language development is constrained by the system she created during the first nine years of her residence in Madrid.

The other informant, Luis, came to Madrid in 1989 from Taiwan, where he and his wife were hair stylists. He made an initial trip to Spain to determine in which Spanish city he wanted to settle with his family, deciding on Madrid, because he had a sister living there who while in Spain had married a man of Chinese descent who himself had grown up in Spain. For the first several years in Madrid, they leased a hair cutting and styling business. From the beginning, then, he has had contact with Spanish-speaking clients six days a week, 48 weeks a year. In his interview, Luis reported that he learned a lot of his Spanish from his clients. Initially, he used textbooks brought from Taiwan to master the Spanish alphabet and language, and he has been able to practice his Spanish with his clients. Apart from work, Luis and his wife have virtually no contact with Spanish speakers. They socialize with people from their own town in Taiwan who also reside in Madrid.

In Tables 5 and 6, we show the distribution of perfective and imperfective forms classified by verb class in Jenny and Luis' speech. The data were collected in interviews with the subjects.

Table 5: Distribution of perfective and imperfective forms in Jenny's speech
(p < 0.001, Chi-square = 116.7 [df = 2])

Verb Form	States	Activities	Achievements
Imperfective	218 (99.1%)	83 (96.5%)	48 (58%)
Perfective	2 (.9%)	3 (3.5%)	35 (42%)

Table 6: Distribution of perfective and imperfective forms in Luis' speech
(p < 0.001, Chi-square = 27.15 [df = 2])

Verb Form	States	Activities	Achievements
Imperfective	85 (96.6%)	36 (77%)	5 (45%)
Perfective	3 (3.4%)	11 (23%)	6 (55%)

There are three striking points about the distribution of forms in their speech. First, it coincides very closely with the percentages found in the database sampling of oral speech from Spain (see comparison in Table 7). Since the informants reside in Spain and have received the entirety of their Spanish input from there, we take the oral sampling to be representative of what the informants have received as their input. It is apparent that both Jenny and Luis have mapped the verb classes onto the aspectually appropriate forms: States and Activities onto imperfective forms over 90% of the time; Achievements onto perfective forms over 50% of the time.

Table 7: Comparison of the distributions of verb forms (raw numbers and percentages) according to imperfective (present, imperfect) and perfective (preterit) verb forms in oral speech in Spain, Jenny's speech, and Luis' speech.

Oral Speech in Spain	**States**	**Activities**	**Achievements**
Imperfective	24409 (98%)	951 (91%)	394 (45%)
Perfective	525 (2%)	90 (9%)	473 (55%)
Jenny			
Imperfective	218 (99%)	83 (96.5%)	48 (58%)
Perfective	2 (1%)	3 (3.5%)	35 (42%)
Luis			
Imperfective	85 (96.6%)	36 (77%)	5 (45%)
Perfective	3 (3.4%)	11 (23%)	6 (55%)

The reason for this could be that the informants are sensitive to lexical and grammatical aspectual distinctions. This would favor the POA hypothesis. However, the fact that the distribution of forms in oral speech from Spain is so close to the distribution found in the informants' speech suggests that the DBH is strongly supported. Second, this is all the more remarkable because neither informant uses present or past forms in a target-like way. In speaking largely about past events, Jenny only uses past forms 7.5% (41/544) of the

time, largely with telic predicates. In his speech, Luis uses past tense forms 11% (22/222) of the time to refer to past events. Again, these are largely with telic predicates. In other words, neither Jenny nor Luis displays a command of Spanish verb morphology.

From the discussion so far, we obtain an idea of what type of verb forms we can expect to find in contact varieties of Portuguese or Spanish, such as immigrant speech. Before discussing African Portuguese, we need to mention some of the other traits found in Jenny and Luis' speech.

Regarding temporal markers, in Jenny's speech, we find the use of *ya* 'already' (Portuguese *já*) and the past participle suffix *-do*. Examples are shown in (1) and (2).

(1) a. Yo luego a   la cuatro, cuatro a╱o **ya**   volvé   Sanghay.
      I   after at   the four   four year already return-INF Shanghai
      'I returned to Shanghai after four years.'
   b. Mil novesiento ochenta cuatro, de junio o julio.
      1984   of June or July
      ya coge pasaporte
      already get-3SG-PRES passport
      'I got my passport in June or July of 1984.'

(2) Mi tio   macha-do   Estados Unido... Tu sabe   Yuan
    my uncle leave-PPART United States   2SG know   Yuan
    Taiwan   veni-do?
    Taiwan   come-PPART
    Yuan más   detlás viene.   Yo más   temprano llega-do
    Yuan more behind come-3SG-PRES 1sg more   early arrive-PPART
    'My uncle went to the United States. You know Yuan, who came from Taiwan? Yuan came afterwards. I had arrived before her.'

Of note is that the verb forms that appear with *ya* in (1) are either in a non-finite (*volvé* < *volver* 'return') or 3sg present-tense form (*coge* 's/he gets'). Also of note is that in (2) *-do* either is used as a preterit, as in *mi tio macha-do Estados Unidos* or as a pluperfect, as in *Yuan más detlás viene. Yo más temprano llega-do*, where Jenny had arrived before Yuan came.

Another marker that appears often is *luego* 'then, afterwards'. In Jenny's speech we find it often, as shown in the passage in (3).

(3) Yo nació   Nanking, y   **luego** a la dos año, dos año con mi
    I   born-3SG Nanking and then   to the two year, two year with my
    papá
    dad
    Fue al Peking,   y   **luego** a cuatro año, volvé   Sanhay,
    Went-3sg   to-the Peking and the to four   year return-INF Sanhay
    mi mamá
    my mom

    Sanghay, mi abuela   Sanghay, y Sanghay, poque   mi papá trabajo
    Sanghay my grandma Sanghay, and Sanghay because my dad work
    Nanking, Peking, y   <u>luego,</u> Sanghay, yo <u>luego</u> a la cuatro,
    Nanking, Peking, and then   Sanghay I then   to the four
    cuatro año ya,   volve Sanghay
    four   year already return Sanghay
    'I was born in Shanghai and then I went with my father at two years
    old to Beijing, and then after four years I returned to Shanghai; my
    mom was in Shanghai, my grandma was in Shanghai, and I was in
    Shanghai because my dad was working in Nanking and Beijing, and
    then Shanghai and then four years after he returned to Shanghai.'

In Jenny and Luis' speech, zero copula are not uncommon. An example of this is given in (4).

    (4) Antes solisid◁ visado ‿ ma fásil.
        before apply.for visa ‿ more easy
        'Before it was easier to get a visa.'

In Luis' speech, we find overt copula, but interestingly it is the form *son* 'they are'. We take this to suggest that *son* is not only frequent (the second most frequent form after the 3sg, but it is more easily perceptible because of its CV(C) form.

    (5) S∡, tenemo   dos…dos niños, laa mayó ‿ catose   año,
        yes have-1pl two two kids the older   fourteen year
        la pequeño   son laa nueve, eh, ocho, ocho y media
        the small-masc   they.are nine, uh eight, eight a a half.
        'Yes, we have two children, the older is 14, the younger is 9, eh, 8 ½.'

Note also that his speech has zero copula and that he prefers the determiner *la* 'the.FEM' to *el* 'the.MASC'. We interpret this as a further indicator that CV structure is preferred to VC structure when there is a choice.

    The last trait to be discussed is the multi-purpose preposition. In Jenny's speech, *de* 'of' is used to mark possession, source, and cause, as in Spanish, but also location (6) and as a linker between adjectives and nouns (7), much like the Chinese particle *de* is used in Mandarin, to link a modifier with a modified element.

    (6) Víve   de   España.
        Live-3SG of   Spain
        'She lives in Spain.'

    (7) China   de   general
        China   of   general
        'Chinese general'

In Luis' speech, the multi-purpose preposition is *porə*, which has no exact equivalent in Spanish, the nearest being *por*, which itself is a multifunctional preposition, with meanings as diverse as 'because of, through, around, by, for'. Examples of *porə* are given in (8). (Apart from the use of *porə*, we also find *ya* as a temporal marker alluded to above, and we note Luis' use of the pronominal clitic *me* 'me' instead of *yo* 'I'.)

(8) a. Ya   llegá   por?   españó   (Directional)
         already   arrive-inf por?   Spanish
         'He arrived to Spain.'
     b. Mi cuñado   estudio por? aquí a coregio.   (Locative)
         My brother-in-law   study   por? here to high.school
         'My brother-in-law went to secondary school here'
     c.   ...   no paraa… no por? caná (ganar) dinero (Purpose)
         (I moved here) NEG for … no por?   earn   money…'.
         'I moved here, not to make money …'
     d. Me pensaw   por? viene aquí. (Complement-clause marker)
         me thought-PASTP por?   comes   here
         'I was thinking to come here.'

In sum, the immigrant speech of both informants displays similar traits, both in the use of the verb forms, as well as in the use of adverbs and tense-like markers, zero copula, multi-purpose prepositions, the use of *me* for *yo*, and the preference of CV over VC structure where there is a choice. In the next section, we will see that these traits are also prevalent in the portrayals of African Portuguese of the 15th and 16 centuries.

African Portuguese in 15th-16th century Portugal and the Portuguese-based creoles

Various scholars, including Naro (1978), Goodman (1987, 1988), and Clements (1992, 1993) have discussed the formation of pidgin Portuguese. The difference between those studies and the present one is the focus here on immigrant speech and naturalistic second language acquisition as the basis of the traits found in pidgin Portuguese and, subsequently, in many of the Portuguese-based creoles.

Given that the population of southern Portugal and Lisbon in 15th and 16th centuries had 10-15% African population, as mentioned in the introduction above, it would be reasonable to assume a considerable amount of naturalistic second language acquisition taking place in the area because the Portuguese and the Africans had to communicate in some manner and the target language was undoubtedly Portuguese. In such a contact situation, the prediction would

be that the Africans in Portugal would speak something very similar to the variety that Jenny and Luis speak. In fact, the speech we have – even though it constitutes an approximation of African speech by playwrights of that time – displays many of the characteristics we find Jenny and Luis' speech.

In plays of Anrique da Mota (15th century) and Gil Vicente (16th century), there are examples of African speech in the mouths of some of the characters of their plays. If these portrayals approximate significantly the manner of speaking of the Africans in Portugal at that time, we would expect them to show traits found also in the immigrant speech examined above. Examples found in Anrique da Mota (from *A lamentação do clérigo* 'the lamentation of the priest', a poetic farce written in the last quarter of the 15th century) are shown in (9), and those from Gil Vicente in (10), taken from Naro (1978). The forms in bold are those forms that have exact or near correspondences in immigrant Spanish discussed above.

(9) a. **a mym** nunca **sar** ruim.
 to me never be bad
 'I'm not (or never) bad.'

 b. Vós **logo** todos chamar …
 you right.away all call
 'You call everyone right away.'

 c. Aqui estar juiz no fora,
 here be judge in outside
 'Here is the judge outside.'

 d. **a mim logo vai** 'té lá.
 to me right.away goes until there
 'I will go there.'

 e. **Mim** também falar mourinho …
 me also speak Moorish
 'I also speak Moorish.'

 f. Mim não ‿ medo no toucinho …
 me no fear in.the lard
 'I'm not afraid of lard.'

(10) a. **logo a mi bae** trazee.
 right.away to me go-3sg bring
 'I'll go bring it right away.' (Naro 1978: 329)

 b. Quando **já** paga a rinheiro, deytá a mi fero **na** pé.
 when already pay the money put to me iron on foot
 'When he paid the money, he put irons on my feet.' (Naro 1978: 329)

 c. bosso barba **já** cajaro.
 your beard already white
 'Your beard {has turned/is already} white.' (Naro 1978: 330)

    d. **Ja**   **mim**   disee   isso   **ja**
      already me   say   this   already
      'I already said that.'   (Naro 1978: 329)

    e. Porque tu   nam ‿   **burguntando**?
      why   you   NEG   asking
      'Why aren't you asking?'   (Naro 1978: 330)

    f. ... e   levare elle   **na**   bico.
        andcarry he   PREP beak
      '... and carry it off on the sly.'   (Naro 1978: 332)

First, we note that there is no verbal inflection apparent. For temporal marking, we find frequent use of the adverbials *já* 'already' in (10b, 10c, 10d) and *logo* 'right away' in (9b, 9d, 10a). This corresponds to use of *ya* 'already' and *luego* 'then' documented in (1a, 1b, 8a) and (3), respectively, in Spanish immigrant speech. In the examples in (9) and (10), we find no use of the past participle forms in *-do*, but we also acknowledge that the data are few.

In terms of verb forms, the 3sg-present or the infinitive seem to be the default forms (*vai* or *bae* (< Ptg. *vai* 's/he goes') for *vou* 'I go', *sar* (< Ptg. *ser* 'be') for *sou* 'I am', *chamar* 'call' for *chamades* 'you.pl call', *estar* 'be (located)' for *está* 's/he is (located)', *paga* 'pays' for *pagou* 's/he paid', *deytá* 'put-INF' for *deytou* 's/he put', *disee* (< Ptg. *dizer* 'say-INF') for *disse* 's/he said'. We take the forms of the dynamic verbs to be most likely accurate portrayals of African Portuguese at that time. There is a significant correspondence between these in normal oral speech, in Spanish immigrant speech, and African Portuguese. The use of the infinitive forms *sar* 'be' and *estar* 'be' is not what would be expected given that these are stative verbs. Rather we would expect the 3sg present *é* 's/he/it is' or possibly the corresponding plural *são* 'they are', and *está*. However, we find null copula in both varieties of Portuguese and Spanish. Regarding function words, we note the multi-purpose preposition *na* in (10b, 10f), which is used to mark location. Recall that in Jenny's speech it is *de* (in [6] and [7]) and in Luis', it is *por?* (in [8]).

As in Chinese immigrant Spanish, we also find in African Portuguese copula deletion, shown in (9f) and (10e). We note also that in the latter case of copula deletion it occurs with a gerund form, namely *burguntando* (< Ptg. *perguntando*). In Jenny's speech, we find gerund forms for dynamic, durative predicates, of which *burguntando* can also be considered one.

As for pronominals, we find in the portrayal of African Portuguese the use of *(a) mim* 'to me' for the subject pronoun 'I' (in [9a, 9d, 9e, 9f, 10a]), somewhat comparable to to Luis' use of the clitic *me* 'me' for 'I' in (8d).

We see, then, that the use of verb forms is comparable in all the varieties under consideration and that there is no verbal morphology to be found. Moreover, the adverbs selected as temporal markers are the same ones (*luego, ya* and *logo, já*) and both varieties exhibit null copula and exhibit no nominal

agreement. Finally, they both have a multi-purpose preposition. With this in mind, we raise the question about whether these traits would have found their way into pidgin Portuguese and from there into the Portuguese-based creoles of Africa and Asia. The answer to the question is affirmative to a large extent.

First, we note that although we found no trace of past participle *-do* in African Portuguese as portrays in the plays of the 15th and 16 centuries, in all Portuguese-based creoles *-do* is present in at least one function (cf. Clements forthcoming). In many of the creoles, the verb form for many of the stative verbs is the 3sg present-tense form (e.g. *sab* 'know', *pod* 'be able'). This corresponds directly, we argue, with the finding of the overwhelming presence of 3sg present-tense forms found in Jenny and Luis' speech. We note also the presence of *vai* or *bae* 'goes' in African Portuguese (in [9d, 10a]). In Jenny's speech, we find *vasa* (< Spanish *vas a* 'you-sg go to') for 'go'. Important here is that the form is not taken from the infinitive *ir*, but rather from a singular form. As Holm (1988:267-68) notes, 'the overall interrelatedness of the Portuguese-based creoles is also supported by other common features, e.g. form for 'to go' in African, Asian, and New World creoles derived not from the Portuguese infinitive, as is the case of most creole verbs, but rather from *vai*, the third person singular form of the present indicative and (probably a more likely model) the second person singular imperative.'

In Spanish immigrant speech, 15th-16th century African Portuguese, and the Portuguese-based creoles, the basis for the dynamic verbs is the infinitive, and the gerund, where present in the creoles, marks either imperfective aspect or present tense. That is, the form has been extended to all dynamic verbs, even though in Chinese immigrant Spanish and 15th-16th century African Portuguese it was only found in durative, dynamic verbs. This may be a case where the adults who created the immigrant speech and the pidgin used the gerund only for a certain verb class (this would be considered an innovation). The form later came to be generalized to apply to all dynamic verbs through creolization, a process of regularization by children who are learning the pidgin as a native language.

Noteworthy is also that forms such as Spanish *ya* and Portuguese *já* 'already' are found in both varieties of immigrant speech and also mark past tense in some Indo-Portuguese creoles, i.e. those spoken in Korlai, Bombay, Cochin, Sri Lanka, as well as in Philippine Creole Spanish (which has a Portuguese base; see Whinnom 1956), and Malaysian Creole Portuguese. Similarly, *luego* and *logo* are found in the immigrant varieties examined here, and reflexes of *logo* are found as future markers in the Asian Portuguese creoles. In Macau Creole Portuguese, the form *sã* (< Portuguese *são* 'they are'), which corresponds to *son* found in Luis' speech, is the invariant copular form. We find this highly interesting given that this creole is the product of Chinese-Portuguese contact. A form we find in immigrant speech and in the creoles is *está* or *ta*; in the former it is a copula, in the latter it is often used as

a marker of the present-tense. Finally, in all varieties examined here, Chinese immigrant Spanish, African Portuguese, and the Portuguese-based creoles, we find evidence of the null copula.

Conclusion

In this study, we have assumed a model of Emergent Grammar in which linguistic structure emerges from frequency and perceptual saliency. We have concentrated on the traits of the Portuguese-based creoles that can arguably have come from L2 varieties of Portuguese initially spoken by Africans in Portugal in the 15th and 16th centuries. To argue for this, we have showed how certain verb forms make their way into the speech of immigrants, due to frequency of occurrence of verb forms based on their respective verb class (the POA and the DBH). And we suggested that some forms become part of immigrant speech because of their more perceptually salient form (e.g. *son* over *es*, and *são* over *é*). We found that the many traits that characterize Chinese immigrant Spanish also define African Portuguese, and are those that also have found their way into the Portuguese-based creoles. Whether pidgin Portuguese had its beginnings in Portugal, incorporating these traits into a variety of Portuguese that then traveled with the Portuguese sailors, or whether pidgin Portuguese emerged in the interaction between the Portuguese and Africans in Africa, the same relative frequency of occurrence of forms would apply, as would the same processes used in naturalistic L2 acquisition, including sensitivity on the part of the speakers/learners to perceptual saliency of items. In this way, we hope to have contributed to the understanding of the role that naturalistic L2 acquisition played in the formation of the pidgin Portuguese and, subsequently, the Portuguese-based creoles.

References

Andersen, R.W. (1993) Four operating principles and input distribution as explanation for underdeveloped and mature morphological systems. In *Progression and regression in language* (Hyltenstam, K. & Viborg, A., editors), p. 309-339. New York: Cambridge University Press.

Andersen, R.W. & Shirai, Y. (1996). The primacy of aspect in first and second language acquisition: The pidgin-creole connection. In *Handbook of second language acquisition* (Ritchie, W. C. & Bhatia, T. K., editors), pp. 527-570. San Diego, CA: Academic Press.

Boxer, C. (1975) *Women in Iberian expansion overseas, 1415-1815.* Oxford: Oxford University Press.

Bybee, J. (1985) *Morphology.* Amsterdam: Benjamins.

Clements, J.C. (2003) The tense-aspect system in pidgins and naturalistically learned L2. *Studies in Second Language Acquisition*, **25**, 245-281.

Clements, J.C. (1993) Rejoinder to Naro=s >Arguing about Arguin=. *Journal of Pidgin and Creole Languages*, **8**, 119-124.

Clements, J.C. (1992) On the origins of pidgin Portuguese. *Journal of Pidgin and Creole Languages*, **7**, 75-92.

Clements, J.C. (Forthcoming-a) The Portuguese-based creoles in Africa and Asia. *Berkeley Linguistic Society 29.*

Clements, J.C. (Forthcoming-b) *The linguistic legacy of Spanish and Portuguese: Colonial expansion and language chang*e. Cambridge: Cambridge University Press.

Godinho, V.M. (1981-1983) *Os descobrimentos e a economia mundial.* 4 vols. (2nd edition). Lisbon: Arcádia.

Goldberg, A. (1999) The emergence of the semantics of argument structure constructions. In *The emergence of language* (MacWhinney, B., editor), pp. 197--212. Mahwah, NJ: Lawrence Erlbaum.

Goodman, M. (1988) Response to Naro. *Journal of Pidgin and Creole Languages*, **3**, 103-107.

Goodman, M. (1987) Pidgin origins reconsidered. *Journal of Pidgin and Creole Languages*, **2**, 149-162.

Holm, J. (1988-1989) *Pidgins and creoles.* 2 vols. Cambridge: Cambridge University Press.

Hopper, P.J. (1998) Emergent grammar. In *The new psychology of language* (Tomasello, M., editor), pp. 155-175. Mahwah, NJ: Lawrence Erlbaum.

Hopper, P.J. (1988) Emergent grammar and a a priori grammar postulate. In *Linguistics in context* (Tannen, D., editor), pp. 117-134. Norwood, NJ: Ablex.

Hopper, P.J. (1987) Emergent grammar, *Papers of the 13th annual meeting, Berkely Linguistics Society*. Berkeley, CA: Berkeley Linguistics Society, 138-157.

Klein, W. & Perdue, C. (1992) *Utterance structure: Developing grammars again.* Amsterdam: Benjamins.

Naro, A. (1978) A study on the origins of pidginization. *Language*, **54**, 314-347.

Tinhorão, J.R. (1988 [1997]) *Os negros em Portugal. Uma presença silenciosa*. Lisbon: Caminho.

Whinnom, K. (1956) *Spanish contact vernaculars in the Philippine Islands.* Hong Kong: Hong Kong University Press.

J. Clancy Clements
University of New Mexico