Susana Afonso* and Francesco Goglia

Linguistic innovations in the immigration context as initial stages of a partially restructured variety: Evidence from SE constructions in the Portuguese of the East Timorese diaspora in Portugal

Abstract: Portuguese became one of the official languages of independent East Timor after ca. 25 years of Indonesian rule; this prevented the partial restructuring of an East Timorese variety of Portuguese in a similar way to that undergone by other Portuguese varieties (e.g., Mozambican, Angolan and Vernacular Brazilian Portuguese). We will discuss the idiosyncratic use of SE constructions in the speech of literate Portuguese-speaking East Timorese immigrants in Portugal, who will go back to East Timor and will be likely to lead language change. Given this particular link between East Timor and its diasporas, linguistic innovations in the immigration context can shed light on the initial stages of a future partially-restructured East Timorese Portuguese variety. SE constructions are highly polysemous and marked and the data show that innovative patterns are emerging, comprising deletion and generalization of the clitic as well as creative uses of these constructions, mainly observed in impersonal and spontaneous situation types. These innovative patterns can be attributed to L2 acquisition and to the interference of Tetum Dili.

Keywords: Portuguese, East Timor, partial restructuring, SE constructions

1 Introduction

In the literature on language contact, there is general consensus on acknowledging the role of second language (L2) acquisition and use in contact-induced

*Corresponding author: Susana Afonso, The University of Exeter, Exeter, UK,
E-mail: s.p.c.afonso@exeter.ac.uk
Francesco Goglia, The University of Exeter, Exeter, UK

change and the creation of new contact varieties such as ethnolects in the immigration context, partially restructured varieties,[1] and creolization (Clements 2009; Clyne 2003; Holm 2004; Matras 2009; Thomason 2001; Thomason and Kaufman 1988; Winford 2003). The immigration context is an ideal setting for the study of language contact phenomena involving naturalistically acquired L2 and the formation of new varieties of both the language of the host country and the immigrant languages. Due to its abrupt and unstable nature, the immigration context can provide us with an ideal locus to observe the initial stages of language change, otherwise unobservable in other contact situations in which only the final stages of language change are available. Backus et al. (2011) and Heine and Kuteva (2005) use the term 'nonce replication' to refer to contact-induced innovations resulting from short-term contact, produced only in the idiolects of some speakers and which are context-dependent. The majority of innovations in the immigration context may not lead to a situation of emergent change or a new ethnic variety. Immigrant groups in Europe are under extreme pressure to be assimilated and achieve native-like proficiency in the language of the host country by the second generation. However, this may not be the case if the immigrant group speak a variety of a pluricentric language, in which case differences may be retained. In the last decades, there have been many studies in Europe focusing on language use and contact in immigrant communities having a postcolonial link with the host country, including those by Meeuwis (1997) on Zairians in Belgium, by Sebba and Wootton (1998) on the African-Caribbean community in England and by Märzhäuser (2011) on Cape Verdeans in Portugal.

In Portugal, the majority of immigrants came from former colonies and share a previous knowledge of Portuguese, in some cases of Portuguese L2, as in the case of Cape Verdeans, Mozambicans, Angolans and East Timorese. In Cape Verde, Mozambique and Angola, Portuguese is well established as a second language, stabilizing, undergoing nativization in the big cities and on its way to becoming a partially restructured variety. The case of East Timor is more peculiar. The presence of Portuguese is residual and the language became official again (sharing this status with Tetum Dili) only twelve years ago as part of an ongoing process of language planning. East Timorese immigrants in Portugal are mainly part of a highly educated elite, many of which are teachers,

---

1 We use the term 'partially-restructured languages' to refer to those varieties which have grown out of other varieties with which they share the majority of linguistic features. Examples of partially-restructured languages are Brazilian Vernacular Portuguese or Afrikaans. The term is used often in opposition to fully-restructured languages, or creoles, whose linguistic structure is radically different from that of their lexifier language (Holm 2004).

and are likely to go back to East Timor. Because of this unique sociolinguistic situation and link between the Portuguese diaspora and the new official reintroduction of Portuguese in East Timor, linguistic innovations in the idiolects of East Timorese in Portugal are likely to reveal interesting clues to the initial stages of a future partially restructured variety of East Timorese Portuguese in the making. This hypothesis is supported by the fact that in other partially-restructured varieties of Portuguese such as Mozambican Portuguese (Gonçalves 2004) and Vernacular Angolan Portuguese (VAP) (Inverno 2009), restructured linguistic features are observable in the informal speech of speakers who have more contact with Portuguese.

The present article results from the research conducted in a project entitled "Patterns of multilingualism among different generations of the East Timorese diasporic community in Portugal." It offers a qualitative analysis of occurrences of reflexive constructions and related types, here referred to as SE constructions, in the Portuguese L2 idiolects of East Timorese immigrants in Portugal. The choice of these constructions is justified by the fact, discussed in more detail below, that they constitute a marked linguistic construction, that is, a construction which is overtly coded by a morphological feature (Haspelmath 2006), and are hence less likely to be acquired and used in L2 communication; unmarked features are those which "are easiest to interpret and to learn from both the perceptual and productive standpoint" (Thomason and Kaufman 1988: 49). Overtly coded features are thus also less likely to form part of a contact variety such as a partially restructured one (Clements 1996 and Clements 2009; Thomason and Kaufman 1988). Clements (1996: 47) adds that the sociolinguistic characterization of the speakers, namely the type of bilingual situation – in addition, we would assert, to the attitudes towards the different languages, as well as the typological proximity of the languages in contact – are crucial to understanding the emergence of a new variety.

The innovative uses of the SE constructions may remain idiosyncratic or cease to exist if the speakers aim at European Portuguese (EP) as their target language (TL), but some may be retained and propagated in the speech of L2 learners or even future monolingual descendants, becoming part of an emergent variety. Studies of partially restructured varieties usually focus on the final stages of language change, i.e., the actual new variety,[2] but we believe that a discourse-level qualitative analysis of individual idiolects can shed light on mechanisms of partial restructuring which are already in advanced stages, for example in other Lusophone countries. In Brazil, the spread of Portuguese and its contact with African languages took place in the past and Vernacular

---

2 See Holm (2004) for an in-depth discussion on partially-restructured languages.

Brazilian Portuguese (VBP) has already undergone the process of partial restructuring (Holm 2004). On the other hand, in Angola and Mozambique, the varieties of Portuguese are still undergoing the process of partial restructuring (Gonçalves 2004; Inverno 2004), especially accelerated in the major cities, where Portuguese is increasingly being used as a lingua franca (Gadelii 2002: 27). The case of Portuguese in East Timor is similar to that in Angola and Mozambique, in the sense that they all are multilingual countries, colonized by the Portuguese, which became independent in 1974 and adopted Portuguese as their official language for use in administration and education, where Portuguese is acquired as L2 and EP is the TL (Petter 2009). However, the process of partial restructuring of Portuguese in East Timor was slower than elsewhere due to the very reduced number of native speakers of Portuguese (ca. 100 in the second half of the 19th century [Boxer 1947]) and the limited reach of Portuguese to the social elite as well as the very high illiteracy in the country (ca. 90% in 1970 [Thomaz 2002]). Furthermore, the process was interrupted by the Indonesian invasion and therefore remained at the initial stage, as can be observed today in the diaspora in Portugal.[3]

The paper is organized as follows: we begin by providing background information on the East Timorese community in Portugal and the linguistic situation in East Timor (Section 2). In Section 3, we describe the methodology used in the study. Sections 4 and 5 deal respectively with SE constructions in Portuguese and the equivalent constructions in Tetum Dili. Section 6 presents the results of our analysis. Finally, we provide some concluding remarks and reflections.

---

3 A similar idea is also expressed by Carvalho (2001: 21) but from a very different perspective, i.e., learner's errors perspective, concluding that there are deviant tendencies which remained in a set of former speakers of Portuguese whose competence in Portuguese declined during the Indonesian regime when the already reduced network of speakers who used Portuguese practically disappeared. Albuquerque (2012), on the other hand, regards some phonological and lexical features from 16th century Portuguese, as well as lack of number agreement as evidence for the existence of a national East Timorese variety of Portuguese on par with other national varieties, such as European and Brazilian Portuguese. However, there is contradictory sociolinguistic and historical evidence for the existence of a stable national variety. In our corpus we found a very high variability of agreement patterns at the idiolectal level, for example. Furthermore, speakers do not seem to perceive that the Portuguese they speak is a separate variety of Portuguese yet (Batoréo 2009; Goglia and Afonso 2012), which is a very important factor for the formation of a new variety (Muhr 2012) We, therefore, disagree with the existence of a national variety of Portuguese and we prefer to speak of Portuguese in East Timor rather than East Timorese Portuguese.

2 Background information

The linguistic situation of the East Timorese diaspora in Portugal is a very interesting one and mirrors the language situation in East Timor. Language policies in the territory have been dictated variously by the Portuguese colonial enterprise until 1975, by the subsequent Indonesian regime until 1998 and by East Timorese independence since then. During the Portuguese presence in East Timor, Portuguese was the official language, but knowledge of it was not widespread among the population: only 0.25% of the pre-1975 East Timorese population had been educated in Portuguese (Weatherbee 1966: 684). Following its takeover in 1975, Indonesia banned Portuguese, made Indonesian the official language and implemented a wide education program aiming to Indonesianize the population, which was seen as fundamental to the national unity of the Republic of Indonesia (Dardjowidjojo 1998). Finally, after independence, East Timor adopted as official languages both Portuguese and Tetum Dili, one of the most widely spoken native languages in the territory. Indonesian and English are working languages.

In addition to the official languages, 22 native languages of East Timor continued to be spoken, the majority being oral, non-standardized languages (Lewis 2009). The East Timorese 2002 constitution grants the local languages the status of national languages and promotes their development. However, the *de facto* development of these languages has not been a priority of the government. East Timor has therefore always known a situation of multilingualism, and more or less aggressive official language policies have not greatly disrupted its stability.

Due to the political changes in East Timor, and more notably after the Indonesian invasion in 1975, many East Timorese left the country, seeking asylum abroad, especially in Portugal and Australia. The four distinct waves of immigration between 1976 and 1998 (Viegas 1998) meant that the diaspora in Portugal grew in linguistic diversity, reflecting the different linguistic repertoires of the successive generations, as Table 1 shows.

The sociolinguistic questionnaires which were distributed among the East Timorese diaspora in Portugal as part of the project reveal that respondents reproduce their multiple linguistic repertoire in the immigration context. However, this repertoire is reshaped by two main factors: 1) the age group, as different generations were exposed to changes in language policy and planning, still ongoing in East Timor, and 2) the linguistic situation in the host country in which Portuguese is the only official language. In terms of attitudes towards the languages, Portuguese and Tetum Dili are clearly perceived as the most useful

Table 1: Linguistic repertoire of East Timorese immigrant groups (Goglia and Afonso 2012).

Immigration waves	Characteristics of the immigration groups	Languages
until 1976	Literate elite of East Timor, including civil servants and former Portuguese colonial administrators	Portuguese, national languages and Tetum
1977–1981	East Timorese of Chinese origin, economically affluent	Hakka, Cantonese, Mandarin, Portuguese, Tetum and perhaps national languages (Hajek 2002: 182)
1982–1993	Reunion of East Timorese families	Different levels of command of Portuguese, national languages, Tetum and Bahasa Indonesia
1993–1998	Educated male asylum-seekers (18–30 years old) as a consequence of the Santa Cruz massacre.	Very little if any Portuguese, Bahasa Indonesia, national languages and Tetum
2002–2011	Educated East Timorese	Portuguese, Bahasa Indonesia, English, national languages and Tetum

and preferred languages. Bahasa Indonesia is still used in the community. National languages, which occupy the lowest level in the multiple linguistic repertoire in East Timor, appear to be readily abandoned in the immigrant context, where only Tetum Dili functions as an East Timorese identity marker (Goglia and Afonso 2012).[4]

3 Data collection and informants

The aim of this study is to investigate linguistic structures in the Portuguese of East Timorese immigrants in Portugal. In order to do so, spontaneous conversations in Portuguese among speakers of the East Timorese community were collected. The conversations were conducted and recorded by two research assistants (RAs) who are members of the community, as otherwise it would have been difficult to collect spontaneous data. The role of the RAs, who were fluent in the main languages of the community, Portuguese, Tetum Dili and Indonesian, was to trigger informal conversations with one or more members of

---

4 For these reasons, and also because we did not have access to grammars of the other national languages, this study will take into consideration the structures in Tetum Dili which may be transferred to the Portuguese SE constructions.

the community. They themselves became informants, since they were taking part in the conversations, the topics of which focused on aspects of the informants' lives in Portugal, including hobbies, study, work, friends and challenges, on memories from East Timor and on future perspectives in East Timor. The conversations were recorded and transcribed and relevant information on the data collection process was provided.

The objective of recording semi-structured conversations was to collect a sample corpus of the Portuguese spoken by different generations of the East Timorese diaspora, aiming at maximum diversity of structures spontaneously produced, rather than focusing on particular structures through elicitation. A total of 12 conversations were collected for a total of six hours between July 2010 and March 2011 in the cities of Porto, Braga, Coimbra and Évora. All necessary ethical steps were taken in order to gain the consent of the participants and preserve their anonymity.

The research also entailed the distribution of sociolinguistic questionnaires to gather information on language use and variation, language and identity, and language attitudes. The results of the questionnaire survey have already been presented (Goglia and Afonso 2012). In the present paper, our discussion is based on the analysis of the conversations and we will make use of information from the questionnaires only when needed for our discussion.

Each informant is referred to by a letter, while the RAs are identified as RA1 and RA2. Table 2 lists information regarding the birth dates of the informants and their dates of arrival in Portugal, while Figure 1 shows two time lines, referring respectively to the cut-off dates until which informants completed their primary education fully in Portuguese (1965) and Indonesian (1992) and to the changes in language policy in East Timor. In the upper timeline, the shadowed intervals correspond to the periods in which informants experienced the transitions between the two official languages in education. Figure 1 should be read in conjunction with Table 2 in order to perceive the relation between the informants' personal data, exposure to formal Portuguese and language policy in East Timor.

The informants were contacted either via the East Timorese university student associations in Braga, Coimbra and Évora, or through Tane, an East Timorese community association in Porto, or from among the RAs' personal contacts. All the informants belong to the educated elite. In fact, the diasporic community in Portugal is composed of highly educated or economic affluent individuals, who will potentially lead changes in the linguistic ecology, both in the immigration context and in East Timor (see Taylor-Leech 2007). All informants are multilingual (see Goglia and Afonso 2012 for a discussion on their repertoire), sharing and using Tetum Dili as a lingua franca in the immigrant context.

Table 2: Information on informants.

Conversation (Conv)	Informants (Inf) and birth date	Date of arrival in Portugal
1	A: 1963	2009
	RA2: 1981	2003
2	B: 1988	2009
	C: 1988	2009
	RA2	
3	D: 1942	1976
	E: 1975	1976
	F: 2002	Born in Portugal
	RA1: 1970	1976
4	G: 1944	1969
	RA1	
5	H: 1961	1994
	RA2	
6	I: 1985	2007
	RA2	
7	J: 1975	2004
	RA2	
8	K: 1961	1976
	RA2	
9	L: 1981	2010
	M: 1979	2010
	RA1	
10	N: 1980	2000
	RA1	

Figure 1: Time lines of changes in language policy and exposure to language of education.

From the sample corpus which was spontaneously produced by members of the East Timorese diaspora in Portugal with Portuguese as L2, the production of SE constructions will be described and analyzed. Given that the RAs facilitated semi-structured conversations and that SE constructions were not elicited, this study is necessarily qualitative, as the incidence of the different construction types varies in each conversation.

4 The Portuguese SE constructions

In contact-induced change, two factors may explain the existence of particular patterns: markedness and typological proximity. Marked features – overtly coded and which are harder to process and/or produce – are generally not transferred in contact situations, and the most probable outcome is simplification. However, this may not always be the case; not only may marked features be preserved if the substrate language displays similar features, but the overt features may be also observed in a wider range of contexts. In this section we will describe the Portuguese SE constructions and consider how they vary across varieties of Portuguese.

SE constructions in Portuguese, similarly to those in Romance, are highly polysemous constructions, forming a network organized by a similarity in form and, to a certain extent, in meaning. Kemmer (1993) identifies semantic domains or situation types in Romance which are typologically associated with constructions with an identical marking – the use of *se* and its personal inflected forms (Table 3).

Table 3: Situation types associated with reflexive and middle constructions (Afonso 2008).

Construction types		Situation types	Marker
REFLEXIVE		Direct reflexive	Inflected marker
		Indirect reflexive	for person
MIDDLE	Central middle domains	Reciprocal	
		Translational and non-translational motion	Inflected marker for person
		Grooming	
		Body posture	
		Cognitive and emotional	
	Peripheral middle domains	Passive and impersonal	*se*
		Event spontaneity	

For each situation type there is one marker, the reflexive clitic pronoun *se*, although with variations in form. Apart from the passive, impersonal and spontaneous SE constructions, which exclusively use *se* as a marker, the other situation types are marked by the inflected forms of the pronoun (i.e., 1sg *me*, 2sg *te*, 3sg/pl *se*, 1pl *nos*). The complexity of these constructions is very high, as the same marker expresses a range of situations which, synchronically, are not always clearly semantically related. The reflexive and reciprocal SE constructions can also be marked syntactically through a prepositional phrase: in the former, the preposition is followed by a strong, stressed pronoun *si* followed by *mesmo*, while in the latter, a "complex non-reflexive construction consisting of a quantifier and the word for 'other'" (*outro*) can be used (Gast and Haas 2008: 315). The syntactic expression of reflexivity seems to increase semantic transparency, but it may or may not replace the use of clitic, depending on the type of verb.

The semantic relationship between the constructions is related to transitivity and affectedness. The SE constructions are mapped onto the continuum of transitivity between 1-participant and 2-participant events (Hopper and Thompson 1980), being detransitivizing or valency-reduction devices (Klaiman 1991; Haspelmath and Müller-Bardey 2004). Affectedness is also embedded in the continuum of transitivity: the more transitive an event is, the more control a participant has over the event (Cennamo 1993). Less transitive events seem to be the more patient-oriented ones (Comrie 1981: 70–71), in which the patient's change of state is highlighted. To establish a cline of transitivity and affectedness in relation to SE constructions in particular is more difficult, however. The relative transitivity and affectedness of each particular SE construction will depend, according to Kemmer (1993, 1994), on the relative elaboration of events and consequently on the distinguishability of the participants, i.e., the extent to which they can be conceptualized as two distinct entities.

All varieties of Portuguese, standard or non-standard, exhibit similar situation types, ranging from reflexive to impersonal types in a transitivity continuum. The following examples show the SE construction types in standard EP and Brazilian Portuguese (BP).[5]

---

5 The examples are from the CETEMPúblico (Rocha and Santos 2000) and CETEMFolha, a Brazilian Portuguese corpus, part of the NILC project (*Núcleo Interinstitucional de Lingüística Computacional*), available from the AC/DC project (Santos and Bick 2000).

(1) Reflexive

    a. *O presidente* (…) **fecha- se** *na Casa Branca* (EP)
       the presidente shut.PRES.[6]3sg REFL in.the White House
       'The president shuts himself in the White House.'

    b. *Você* **se acha** *louca?* (BP)
       you REFL think.PRES.3sg crazy
       'Do you think you're crazy?'

(2) Reciprocal

    a. *Apertaram as mãos e* **abraçaram- se** *uns aos*
       squeeze.PAST.3pl the hands and hug.PAST.3pl RECIP one to.the
       outros (EP)
       others
       'They shook hands and hugged each other.'

    b. *A família é muito unida e sempre* **nos**
       the family be.PRES.3sg very united and always RECIP
       apoiamos *uns aos outros* (BP)
       support.PRES.1pl one to.the others
       'Our family is very close and we always support each other.'

(3) Translational motion

    a. *O realizador egípcio inspirou- se na história*
       the director Egyptian inspire.PAST.3sg MM in.the story
       de S. José mas
       of St. Joseph but
       não **se afastou** *do Corão* (EP)
       not MM move.away from.the Quran
       'The Egyptian director took inspiration from
       the story of St. Joseph but remained faithful
       to the Quran.'

    b. *O Charade* **se aproxima** *do Honda Civic em*
       the Charade MM come.closer.PRES.3sg to.the Honda Civic in

---

6 The abbreviations used in this article are: PRES – Present; PAST – Simple Past; IMPF – Imperfect; FUT – Future; INF – Infinitive; PP – Past Participle; SUBJ – Subjunctive; PRT – Verbal particle; REFL – Reflexive marker, MM – Middle marker; PASS/IMP – Passive/impersonal marker; INTR – Intransitive marker; ADJ – Adjective; sg – singular; pl – plural; fem – feminine; masc – masculine; POS – Possessive marker; RECI – Reciprocal marker; FOC – focus marker.

> *tamanho   e*
> size   and
> *custa   bem   menos* (BP)
> cost.PRES.3sg   much   less
> 'The Charade comes closer to the Honda Civic
> in size but costs far less.'

(4)   Grooming[7]

    a.   *Dois   fugitivos   descem   as   escarpas   e   vão*
           two   fugitives   go.down.PRES.3pl the   cliffs   and   go.PRES.3pl
           ao   mar
           to.the   sea
           lavar-   se (EP)
           wash.INF.   MM
           'Two fugitives go down the cliffs to the sea to wash themselves.'

    b.   *Quando   **me   barbeava**   na frente   do   espelho,*
           when   MM   shave.IMPF.1sg in.the.front of.the   mirror
           percebia   meus   traços
           see.IMPF.1sg my   lines
           envelhecidos (BP)
           aged
           'When I shaved myself in front of the mirror I could see my wrinkles.'

(5)   Body posture

    a.   *Sem   pensar,   **levantei-   me** para o   ir*
           without   think.INF,   get.up.PAST.1sg   MM   to   it   go.INF
           buscar (EP)
           get.INF
           'Without thinking, I got up to get it.'

    b.   *Sem   alternativas, voltou   a   **se   sentar***
           without   alternatives, go.back.PAST.3sg PRT MM   sit.INF

---

7 Grooming is a situation type of middle reflexives but not of direct reflexives. The events denoted in this category are different from events in the direct reflexive constructions in that the former are not expected to be performed to another participant, e.g., *castigar-se* 'punish oneself' vs. *pentear-se* 'to comb one's hair.' This is pointed out by Kemmer (1993) and also by Haiman (1983: 803) who termed verbs depicting this kind of events as introverted ("actions which the subject generally performs upon one's self"). The latter are called extroverted verbs. Hence the level of elaboration of the event is higher with extroverted verbs than with introverted ones.

em torno de   uma   mesa (BP)
around   a   table
'Having no alternative, he sat again at a table.'

(6)   Cognitive and emotional

    a.   *Tu   podes   **sentir**- **te**   satisfeito* (EP)
      you   can.PRES.2sg   feel.INF MM contente
      'You should feel content.'

    b.   *57%   não   **se   lembram**   em   quem   votaram* (BP)
      57%   not   MM   remember.PRES.3pl   in   whom   vote.PAST.3pl
      '57% don't remember whom they voted for.'

(7)   Passive and impersonal

    a.   *E   aí   **colocam**-   **se**   novas   dúvidas* (EP)
      and   there pose.PRES.3pl PASS new   doubts
      'And in this case new questions are put forward.'

    b.   *O   mesmo não **se   pode   dizer**   da   pintura*
      the   same   not PASS/IMP can.PRES.3sg say.INF of.the painting
      de   Cy Twombly (BP)
      of   Cy Twombly
      'The same cannot be said of Cy Twombly's art.'

    c.   *Até   a   Primeira Guerra Mundial, não   **se   admitiam***
      until the   First   War   World   not   PASS admit.IMPF.3pl
      mulheres
      women
      em   escritórios (BP)
      in   offices
      'Until the First World War women were not allowed in offices.'

(8)   Event spontaneity

    a.   *Nos   últimos dias o   ritmo   **tornou**-   **se***
      in.the last   days the rhythm become.PAST.3sg MM
      frenético (EP)
      frantic
      'During the last days the rhythm became frantic.'

    b.   *O   litoral algarvio   corre   o   risco de **se afundar*** (EP)
      the   coast   algarve.ADJ run.PRES.3sg the risk   of MM sink.INF
      'The coast of the Algarve runs the risk of sinking.'

   c. *A   relação   entre   os   dois* **se   desenvolveu**   *em*
      the   relationship between the   two   MM develop.PAST.3sg in
      especial   em
      particular   in
      outubro   de   1962 (BP)
      October   of   1962
      'The relationship between the two developed particularly in
      October, 1962.'

   d. *De   alguma   forma*, *o   preconceito   começa   a* **se**
      in   some   way,   the   prejudice   start.PRES.3sg PRT MM
      quebrar, *com*
      break.INF   with
      auxílio   de   coisas   como   o   lançamento   do   CD latino (BP)
      help   of   things   like   the   release   of.the   CD Latin
      'In some way, the prejudice begins to break down with the help of
      events such as the release of the Latin CD.'

The types illustrated in (1–8) correspond to a continuum of elaboration based on the relative distinguishability of the participants (Kemmer 1993 and Kemmer 1994). The reflexive and reciprocal situation types in (1) and (2) correspond to events which are highly elaborated, as the participants are conceptualized as distinct, although in (1) they are co-referential and in (2) they are simultaneously agents and patients. The passive/impersonal situation types (7) are less elaborated than the reflexive and reciprocal types, since the agent, although implied, is unspecified. The remaining central middle types (3–6) show lower elaboration than the reflexive and reciprocal middle types; the distinguishability of participants is very low and conceptualized as different facets of the same unitary entity (physical, psychological, emotional). The lowest elaboration of events is associated with the spontaneous SE construction (8), where external agency is completely removed from the event, making it closer to prototypical one-participant events. The standard EP and BP SE constructions differ in the position of the marker.

    In non-standardized varieties there is far more variation in the way these constructions are marked. Kliffer (1982: 425) notes that in informal BP, the range of situation types typically marked by *se* is undergoing reduction regardless of the type of verb (stative, action, change of state) and the animacy of the subject, as (10) shows. Moreover, *se*-deletion is also observed in impersonal SE constructions such as (11). Indeed, impersonal SE constructions are "relatively infrequent" in spoken BP (Azevedo 2005: 238), with lexical strategies being preferred in informal speech to convey impersonalization, such as the use of

você (2sg) or *a gente* (1pl), both taking a third person verb inflection (Cyrino et al. 2000: 60).

(10) a. *Eu ø **preocupava**   tanto*
         I   worry.IMPF.1sg   much
         'I worried so much.'

      b. *Eu não estou   de acordo   que a   pessoa não ø*
         I   not be.PRES.1sg in agreement that the person   no
         esforça
         make.an.effort.PRES.3sg
         também
         also
         'I don't agree that the person didn't make any effort either.'

      c. *A   lâmpada ø quebrou*
         the light.bulb   break.PAST.3sg
         'The light bulb broke.'

(11) *No   Brasil ø circula   à   direita*
      in.the Brazil   circulate.PRES.3sg on.the right
      'In Brazil they drive on the right.'

One reason for the absence of marking of these constructions is suggested by Azevedo (2005: 237); the loss of reflexive clitics – "the vanishing reflexive" – may be due to the loss of awareness of the function of the marker. Kliffer (1981) puts forward the hypothesis that *se* marks some sort of limit which can be related to the number of participants (in the case of the reflexive, reciprocal and passive/impersonal constructions) or to characteristics of the subject such as motivation and responsibility. According to Kliffer, however, the unmarked construction, without the clitic *se*, does not imply that there are no limits imposed on the event; in certain cases, the limits in an unmarked construction are encoded by other elements present in the immediate linguistic context. For example, in (11) the reduction of participants (or at least the specificity of the only participant in the event) is not marked by *se* but is implied by the presence of the locative *no Brasil*. Another variation, also observed in some regional dialects and sociolects of EP, is the generalized use of the marker *se* with all grammatical persons, particularly in situation types where the marker inflects, as (12) illustrates.

(12) a. ***Vou-   se***   *embora*
         go.PRES.1sg MM.3sg away
         'I am going away.'

   b. *Não   se   entendemos*
      not   RECI.3sg   understand.PRES.1pl
      'We don't get along.'/'We don't understand each other.'
      (Naro and Sherre 2004: 187)

The explanation given by Azevedo and Kliffer for the generalization and deletion of the marker *se* is language internal, but this variation should also take into consideration the early contact phenomenon involving European Portuguese, displaced Africans in Brazil and the local substrate indigenous languages. It is possible that the displaced communities would have acquired imperfectly all the situation types associated with the Portuguese SE constructions and the respective markers. Taking into consideration the hypothesis that marked features do not tend to be transferred to the L2 variety, a level of simplification may have occurred. Reinforcing this hypothesis is the deletion and generalization of *se* to all grammatical persons observed in other contact varieties of Portuguese: VAP (13) and Mozambican African Portuguese (MAP, henceforth).

(13) a. *Cansada*, *sentou*   *ø*   *no*   *caixote*
         tired,   sit.PAST.3sg   on.the   box
         'Tired, she sat on the box.'
         (Mendes 1985 *cit. in* Inverno 2009: 102)

      b. *Ele* *chama*   *ø*   *MS*
         he   call.PRES.2s   MS
         'His name is MS.'

      c. *Nós* *conseguimos*   *se*   *entender*
         we   manage.PRES.1pl   RECI.3pl understand
         'We can understand each other.'

Inverno (2009: 102) pinpoints Bantu substrate influence as the explanation for the latter phenomenon, as in Bantu languages there is only one marker to convey reflexivity and reciprocity for all grammatical persons. In MAP the tendency is for the deletion of the marker in spontaneous SE constructions, especially when a verb occurring in this construction enters an unaccusative/ transitive alternation or is reanalyzed as such (Gonçalves 2004). This variation is one consequence of a more general phenomenon in MAP, the homogenization of the syntactic constructions in predicates with "agent-oriented meaning components" (Haspelmath 1993), implying that transitive and unaccusative verbs are treated in a similar way. In this sense, the spontaneous SE construction whose function is to mark spontaneity morphosyntactically therefore becomes obsolete.

The variation in the marking of the SE constructions in varieties of Portuguese which emerged through contact show, in line with Thomason and Kaufman (1998: 51), that "shifting speakers are likely to fail to acquire marked features from the TL [...], which may result in a decrease of the overall markedness of a (sub) system." It is also possible in some cases, although not in all, that the decrease of markedness through simplification may be due to substrate influence, although this will depend on the typological similarity between the TL and the substrate language.

Taking into consideration the variation of marking in SE constructions in the varieties of Portuguese which underwent partial restructuring, we will assess in the next sections the potential interference of Tetum Dili in the Portuguese L2 spoken by East Timorese.

5 Reflexives and related types in Tetum Dili

Tetum Dili is part of the linguistic repertoire of the East Timorese, being the native language and the lingua franca of a large part of the East Timorese diaspora and/or their everyday language. As Portuguese spoken by the East Timorese diaspora is an L2 variety, it is likely that some of the patterns observed in the speech of the informants are due to substrate influence.

In Tetum Dili, reflexives and reciprocals have the respective functions of co-referentiality and reciprocity. Reflexivity is encoded either by the noun *án* 'self' possessed by a pronoun or by the clitic *–an* attached to the verb (Williams-van Klinken et al. 2002: 62).[8] Reciprocity is marked by *malu* 'each other' and occurs in two different constructions: one with plural referents and the other with a singular referent, the instigator. The other referent(s) are introduced by a preposition *ho* 'with,' as in example (14) (Williams-van Klinken et al. 2002: 60).

---

8 An example is 'He killed himself:'

i) *Nia   oho   nia   **án**   rasik*
    He   kill   3sg.pos   self   own

ii) *Nia   oho-**an**   rasik*
    He   kill-self   own

(14) a. *Sira   baku   **malu***
        They   hit   RECI
        'They hit each other.'

     b. *João   istori   **malu**   ho   Maria*
        João   quarrel   RECI   with   Maria
        'João quarrelled with Maria.'

The encoding of reflexivity by cliticization is similar to Portuguese. However, the reciprocal construction includes no cliticization strategy similar to the use of *se*. Instead, the syntactic marker *malu* is used, corresponding to *um ao/com o outro* ('each other'/'with each other'). Although reciprocity is typically marked by *se*, in some cases, particularly when the construction may not be semantically reciprocal, the heavy marker alone is used, as in (15).

(15) a. *Simpatizaram   **um com o**   outro   e   decidiram*
        sympathize.PAST.3pl   one   with   the   other   and   decide.PAST.3pl
        passear   juntos
        stroll.INF   together
        'They sympathized with each other and decided to take a stroll.'

     b. ***Simpatizaram-***   **se**   e   decidiram   passear   juntos*
        sympathize.PRES.3pl   RECI   and   decide.PAST.3pl   stroll.INF   together

The other situation types, which in Portuguese are encoded by the same marker as the reflexive marker, are marked in Tetum by other types of constructions. Motion, posture, cognitive and emotional events, impersonal/ passive and event spontaneity are marked in Tetum Dili by different constructions. Emotions are expressed by a compound structure consisting of a noun *laran* 'inside' followed by an adjective, which can be introduced by *senti*, a borrowing from the Portuguese verb *sentir* 'to feel' (16), in the non-pronominal form (Williams-van Klinken et al. 2002: 57–58).

(16) a. *Hau   **laran susar**   tanba   hau   nia   inan   moras*
        1sg   inside difficult   because   1sg   POS   mother   sick
        'I am distressed because my mother is ill.'

     b. *Nia   kolega sira   **sentí laran susar**   tebes*
        3sg.POS   friend   PL   feel   inside difficult   truly
        'His friends felt truly distressed.'

Events denoting motion are expressed by serial constructions in which two verbs follow one another: the first verb indicates motion and the following verb indicates direction (Williams-van Klinken et al. 2002: 93).[9]

With regard to spontaneous events, Tetum Dili uses the prefix *nak-* as a detransitivization strategy. The derived intransitive verb does not imply an external agent and the patient is in the subject position, as example (17) shows (Williams van-Klinken et al. 2002: 10–20, 63). This strategy is very similar to the unaccusative SE construction in Portuguese, in which the detransitivization of a typically transitive event is marked by *se*.

(17) a. *Ita*   **fakar**   *bé*
          1pl   spill   water
          'We spill the water.' (on purpose)

   b. *Ita*   *halo*   *bé*   **nak-fakar**
       1pl   make   water   INTR-spill
       'We cause the water to spill.' (accidentally)

Constructions with the detransitivizing prefix *nak-* can also be used to convey passive-like functions, in the absence of a voice system in Tetum (Williams-van Klinken et al. 2002: 51). In *nak-* constructions, the patient is the subject, gaining prominence, while the agent is not expressed and is consequently removed from the event. Other agent-demoting constructions are the impersonal construction with generic *ema* 'person, people, someone' or *ita*, the generic first person plural pronoun (Williams-van Klinken et al. 2002: 51–53).

Although emotions, motion and spontaneous events are not expressed by using the reflexive marker, there are some cases where a transitive verb is followed by the reflexive clitic *-an* to convey these situation types. According to Williams-van Klinken et al. (2002: 63), this pattern is a Portuguese calque where the function of the Portuguese *se* in the motion, emotion and spontaneous event types were transferred to the Tetum *-an*. Hence, the reflexive *-an* attached to transitive verbs *bók* 'move' and *hairak* 'lower' will produce a motion and emotion event type: *bók-an* 'move,' *haraik-an* 'be humble/humble oneself.'

---

9 An example is:

   (i) *Ami*   **halai**   **sae**   *tó*   *Dare*   *nebá*
        We   run   ascend   until   Dare   there
        'We ran up there to Dare.'

In case of event spontaneity, the clitic follows a derived intransitive verb (18) (Williams-van Klinken et al. 2002).

(18) *Odamatan   né   **nak- loke- an***
      Door   this   INTR-open-self
      'The door opened (by itself).'

While Portuguese and Tetum Dili are not typologically related, they both encode events placed in the transitive continuum between 1-participant and 2-participant event types, using similar strategies in some cases. The reflexive, spontaneous and passive types are marked respectively by a clitic (-*an*) and a prefix (-*nak*) which are, like the Portuguese *se*, detransitivizers. In addition, some situation types (motion, emotion and spontaneity) are expressed in Tetum Dili by the reflexive marker, calqued from the Portuguese reflexive marker.

6 The SE constructions in the corpus

In this section we will analyze the interference effects emerging from our corpus, taking into consideration the markedness of the SE constructions in Portuguese (the TL) and the similarities between Tetum Dili (the substrate language) and Portuguese in relation to the encoding of reflexive and related situation types.

Our data show in general, similarly to other contact varieties of Portuguese, that there is both elision and generalization of the marker, but we also observe innovation in the use of the marker. Examples (19) and (20) are cases of marker deletion and (21) is a case of generalization of the marker to other grammatical persons.

(19) *para   sentir   que   não   estamos   cá   sós   e   para   sentir*
      to   feel.INF that   not   be.PRES.1pl   here   alone   and   to   feel.INF
      que
      that
      *temos   família   cá   eu   **aproximo**   aos*
      have.PRES.1pl   family   here   I   come.closer.PRES.1sg   to.the
      timorenses (Conv.2, Inf B)
      Timorese
      'In order to feel that we are not here alone and in order to feel that we have family here I come closer to the other East Timorese.'

In example (19), *aproximar* is a motion verb which can occur in either a transitive or a middle construction. Given that the event bears no object, it refers to the motion of the subject of the event: the speaker. The event is thus a middle event and the construction is therefore marked in standard Portuguese by the clitic *se*. However, motion events, especially motion towards, include the target of the motion marked by the prepositional phrase: *aos timorenses* in (19). In such a construction the use of *se* is redundant from a communicative point of view, as it loses its iconicity (clear form/function correspondence). This is a similar case to the deletion of *se* in VBP, in that the other elements of the construction allow the agent to be implied. Substrate influence should also be taken into consideration. The fact that in Tetum Dili, motion events are not encoded by a clitic, but by a serial verb structure (see (i) in footnote 9), may also explain the absence of the marker in Portuguese. A similar situation is illustrated in example (20).

(20) *...a vida de estudante, a conhecer amigos nós aqui*
      the life of student, to know.INF friends we here
      somos...   nossa
      be.PRES.1pl...our
      terra, tentei sempre **adaptar** com os meus colegas,
      country, try.PAST.1sg always adapt.INF with the my colleagues,
      tentei
      try.PAST,1ag
      adaptar com eles (Conv.2, Inf B)
      adapt.INF with them
      'The student life, knowing friends..here we are....our country, I always tried to adapt with my colleagues, I tried to adapt with them.'

The verb *adaptar*, just like *aproximar*, enters a transitive/middle alternation. In the absence of the object of the event and in the presence of the prepositional phrase *com eles*, the event cannot be conceptualized as transitive, but rather as a middle event.

The utterance in (21) is an example of generalization of the marker to all grammatical persons. Informant K is talking about the attitudes towards the native languages in East Timor and their place in the nation-building project. He uses the marker *se* in an emotional middle situation type with reference to the first person plural.

(21)   *agora*   *eu*   *acho*   *é que*   **temos**   **que**   *em*   *primeiro*
        now   I   think.PRES.1sg   FOC   have.PRES.1pl   to   in   first
        lugar   **sentir-se**
        place   feel.INF MM.3sg
        orgulhosos   *e*   *estimularmos*   *a nós próprios,*   *não é,*   *e*
        proud   and   stimulate.INF.1pl   ourselves,   as it were,   and
        para   *depois*   *ver*
        to   then   see.INF
        a nível de,   *de*   *línguas*   *que*   *existem*   *em*   *Timor,*   *olharmos*
        at level of,   of   languages   that   exist.PRES.3pl   in   Timor   look.INF.1pl
        para   *as*
        at   the
        línguas   *como*   *se*   *fossem*   *uma*   *chave*   *para*   *abrirmos*
        languages   as   if   be.IMPF.SUBJ.3pl   a   key   to   open.INF.1pl
        as   *portas*
        the   door
        para   *o*   *nosso*   *futuro*   *profissional*   *e*   *não*   *como*   *um*   *obstáculo*
        to   the   our   future   *professional*   *and*   not   as   an   obstacle
        para   *nós (Conv. 8, Inf. K)*
        for   we
        'Now, what I think is that we need first of all to feel proud and encourage ourselves, as it were, so that afterwards, in relation to the languages that exist in Timor, we can look at the languages as if they were a key to open the doors to our professional future, rather than an obstacle for us.'

As mentioned in Section 4, this pattern is attested not only in other varieties of Portuguese such as VAP (Inverno 2009: 102) and VBP (Naro and Scherre 2004: 187), but also in dialects of EP (Marques 1968, as cited in Naro and Scherre 2004: 187). We could therefore question the extent to which these simplifications are in fact contact-induced rather than simply cases of internal language drift.[10] It is worth noting, however, that in Tetum Dili, there is only one reflexive marker for all grammatical persons, *-an*, which is used only as a clitic and can also encode situation types very similar to the middle types in Portuguese, for example *bók-an* and *haraik-an* (Williams-van Klinken et al. 2002: 63). Similarly, in the case of VAP, Inverno (2009: 102) considers that the extension of *se* to all grammatical

---

10 Simplification by generalization of the marker to all grammatical persons is also attested in other Romance languages, for instance in some Lombard dialects (Nigel Vincent, p.c.), which bears the question if simplification could be due to natural syntactic change. For a discussion on natural vs. contact-induced simplification see Thomason and Kaufman (1988: Ch. 2).

persons is due to contact with one of the main Bantu language spoken as L1, Cokwe, in which there is only one reflexive marker for all grammatical persons.

In relation to the distribution of construction types by informant and situation type, we observe the following; The informants who had had longer contact with Portuguese (because they had either lived in Portugal for longer and/or been educated in Portuguese) unsurprisingly showed a more target-like use of the construction, especially the marking of the impersonal situation type. This contrasts with the scarce use of impersonal/passive SE constructions by speakers who had been in Portugal for less time and who experienced the transition from Indonesian to Portuguese while at school. In most cases, it seems that the impersonal situation type was acquired as a semi-idiomatic structure (non-analyzed forms used as discourse markers): *como se diz, como se fala*. As we saw in Section 5, the category of voice is absent from Tetum Dili, with impersonalization/passive being conveyed by generic pronouns (1PL) or lexical strategies conveying non-specific referents (e.g., *ema* 'people'). It is also important to note that these situation types are peripheral to the reflexive and mental/psychological core situation types, making it harder to use the same marker by analogical extension.

The reflexive and middle situation types, particularly the psychological/emotional and mental types, are the ones which are more commonly marked, as they are in the TL. From a language-internal viewpoint, the fact that it is the reflexive and the mental/psychological situation types which exhibit less variation can be attributed to the centrality of some SE constructions in relation to others. The reflexive type is the core construction, the middle ones being closer to the core than more peripheral construction types such as the passive or impersonal types. These constructions are also less grammaticalized than the passive or the impersonal constructions,[11] which may indicate that they have a higher level of transparency between form and function. From a language acquisition perspective, this means that they are more easily acquired (DeKayser 2005). In addition, the relative similarity in marking these situation types in Tetum Dili and Portuguese may also contribute to the transfer of these marked patterns into the East Timorese version of the TL. However, even in the encoding of the core reflexive situation type, there are some cases of marker deletion, as examples (22–24) show.

---

11 Vitral (2006) observes that marker in impersonal se constructions displays characteristics which are more like an affix than a clitic, reinforcing the grammaticalization hypothesis by which the marker becomes less meaningful.

(22) *quando*   *davam*   *tiros*   *pronto*   *os*   *meus*   *pais*   *disseram*
      when   give.IMPF.3pl   shots   ok   the   my   parents   tell.PAST.3pl
      para   **esconder**
      to   hide.inf
      para   *ficarmos*   *dentro*   *da*   *casa* (Conv.2, Inf. B)
      to   stay.INF.1pl   inside   the   house
      'When shooting began, my parents told us to hide, to stay inside.'

As it happens with all reflexive events, the verb is transitive and the object is co-referential with the subject. In (22), the reflexive marker is absent, but so are any noun phrases which could potentially be considered objects. The agent of the event of *esconder* is implied in (22) but not marked by *se*. Instead, the presence of the first person singular of the possessive pronoun *meus* in *meus pais* allows this reading.

In (23), another strategy in addition to *se*-deletion is used to convey reflexivity. The speaker is talking about how her experience might contribute to East Timorese society and uses only the reflexive heavy marker *a si próprio* after the verb *conhecer*, without the clitic *se*. The use of the heavy marker is not unusual in varieties of Portuguese, but it is used more as an emphatic marker and does not often occur on its own. In vernacular varieties such as VBP, on the other hand, this is widely observed (Mello 1997: 153, as cited in Holm 2004: 105).

(23) *...nós*   *podemos*   *partilhar*   *a*   *nosso*   *conhecimento,*   *a*
      we   can.PRES.3pl   share.INF   the.fem   our.masc   knowledge.masc,   the.fem
      nossa...   *o*
      our.fem   the.masc
      nosso   *conhecimento,*   *estas*   *experiências*   *todas*   *para*   *ajudar-nos*
      our.masc   knowledge   these   experiences   all   to   help.INF us
      os   *nossos*
      the   our
      povos   *lá*   *em*   *timorense,*   *para*   *crescer*   *mais,*   *para*   **conhecer**
      people   there   in   Timorese   to   grow.INF   more,   to   know.INF
      melhor   **a**
      better   to
      si   **próprio** (Conv. 9, Inf. L)
      one   self
      'We can share our knowledge, our knowledge, all these experiences to help us our people there in Timor, in order to grow more, to know oneself better.'

A similar strategy is observed in (24). Informant K expresses the need to encourage the East Timorese, in which group he includes himself, to be proud of their roots. Reflexivity is marked syntactically (by the heavy marker *a nós próprios*) rather than morphologically (by the light marker *se*), thus simplifying the construction.

(24) *agora   eu   acho   é que   temos   que   em   primeiro*
      now   I   think.PRES.1sg FOC   have.PRES.1pl to   in   first
      lugar
      place
      *sentir-se   orgulhosos e   **estimularmos   a nós próprios**,*
      feel.INF MM.3sg   proud   and stimulate.INF.1pl ourselves,
      não é,   e   para
      as it were,   and to
      depois   ver   a   nível   de,   de   línguas   existem
      afterwards   see.INF at   level   of   of   languages   that exist.PRES.3pl
      em   Timor,
      in   Timor
      olharmos   para   as   línguas   como   se   fossem   uma
      look.INF.1pl   at   the   languages   as   if   be.IMPF.SUBJ.3pl a
      chave   para
      key   to
      abrirmos   as   portas para   o   nosso   futuro   profissional   e
      open.INF.1pl the   doors   to   the   our   future   professional   and
      não   como   um
      not   as   an
      obstáculo   para   nós (Conv. 8, Inf. K)
      obstacle   for   we
      'Now, what I think is that we need first of all to feel proud and encourage ourselves, as it were, so that afterwards, in relation to the languages that exist in Timor, we can look at the languages as if they were a key to open the doors to our professional future, rather than an obstacle for us.'

In (24) there are two situation types which in Standard Portuguese would be marked by the clitic *se*: the emotional middle (*sentir-se*) and the reflexive type (*estimular-se*). Speaker K only marks the first situation type. One of the reasons for this may be the fact that the verb *sentir*, being a verb denoting an emotional event, is more frequently encoded by *se* in the TL. In contrast, *estimular* is just a transitive verb which can be used reflexively (i.e. in a direct reflexive construction).

Despite the expected reduction of markedness, it is worth noting as well that overt marking is observed in unmarked situations. In contact situations, the social conditions of contact define the direction and degree of interference

(Thomason and Kaufman 1988: 19). Moderate or heavy interference is likely to complicate the grammar of the borrowing language or the TL in shifting situations (Thomason and Kaufman 1988: 29). Overt marking of typically unmarked situation types are innovative patterns observed in our data across informants and across situation types. It shows that our informants are marking situation types associated with SE constructions in a creative way, using already acquired linguistic structures and knowledge from previously acquired languages.

One of these constructions is the use of *se* with overt subjects. In impersonal SE constructions, the agent/experiencer of the event is demoted, either because it is not topical or salient or because it is unknown. In examples (25) and (26), the agents, *uma criança* and *grupos, ativistas, diplomatas*, are present in otherwise impersonal constructions. However, if we look at the semantic properties of the agents, they are non-specific and therefore fall into the realm of non-salient, generic participants.

(25) *A   minha   infância   foi   muito   feliz [laughs]   e*
     the   my   childhood   be.PAST.3sg   very   happy   and
     assim   sei
     like,   know.PRES.1sg
     lá   brincar   com   os   meus   amigos,   como   é   óbvio,
     not,   play.INF with   the   my   friends,   as   be.PRES.3sg   obvious,
     como   se
     as   IMP
     faz   uma   criança   e   a   estudar (Conv. 2, Inf. B)
     DO.PRES.3sg   a   child   and   PRT   study.INF
     'My childhood was very happy [laughs] and like, I don't know, playing with my friends, obviously, as a child normally does, and studying.'

(26) *Foi   com   isso   que   saí,   como   se   diz,   a*
     FOC   with   that   FOC   leave.PAST.1sg,   how   IMP   say.PRES.3sg,   PRT
     testemunhar,
     testify.INF
     chamar   a   atenção   internacional   e   pedir   à   comunidade
     call.INF   the   attention   international   and   ask.INF   to.the   community
     internacional
     international
     para   que   se   enviar   os   grupos   ou   os   ativistas   ou   os
     so   that   IMP   send.INF   the   groups   or   the   activits   or   the

diplomatas   etc.,   para que
diplomats,   etc.,   in order to
se   possam   *fazer   alguma   investigação   em*
IMP   can.PRES.SUBJ.3pl   do.INF   some   investigation   in
Timor (Conv. 5, Inf H)
Timor

'That was the reason why I testified, to attract international attention and to ask the international community to send groups or activists or diplomats, etc, so that they could investigate the matter in Timor.'

This pattern seems to be semantically driven, since the syntactic constraint that impersonal constructions cannot exhibit an overt agent is disregarded. The semantic properties of the agent – generic and non-specific – determine that an impersonal SE construction is posited. This could also be a contact-induced pattern, since in Tetum, as well as in Indonesian (Antonia Soriente, p.c.), impersonal constructions are conveyed using an overt lexical item (*ema* 'people' in Tetum Dili and *orang* 'person' in Indonesian). In Portuguese, impersonalization can also be conveyed by lexical and pronominal strategies (Afonso 2008). This may mean that the speaker has acquired the knowledge that one of the situation types of the SE construction is impersonalization, making overt generic subjects seem good candidates to appear in such a construction.

Another case is the overtly coded constructions in which spontaneous verbs occur, such as *eclodir* in (27).

(27) *o   segredo foi   bem guardado que   foi*
      the   secret   be.PAST.3sg well keep.PP   which   be.PAST.3sg
      ultimamente,   que
      lately   which
      *foi   contra a   minha vontade que isso **se***
      be.PAST.3sg   against the my   will   that that MM
      eclodiu,   porque (...) as
      explode.PAST.3sg because   the
      pessoas estavam   a (...) a   escutarem   por meu telemóvel e
      people   be.IMPF.3pl PRT   PRT listen.INF.3pl by   my   mobile   and
      descobriram (Conv. 4, Inf. G)
      find.out.PAST.3pl
      'The secret was well kept which was recently, which was against my will that it came to light, because people tapped my mobile and found out.'

Informant G marks the spontaneity of an already spontaneous event. In contrast, in the TL, spontaneous events are unmarked and only transitive events such as *abrir, fechar, partir* are marked for spontaneity, being in this way coerced into expressing a spontaneous meaning. In (27), the speaker extends, by analogy, the spontaneous SE construction to an already unaccusative event type. This pattern is also observed in (28) and (29) where, similarly to *eclodir*, informant H marks *existir* and *acontecer* for spontaneity.

In (28), informant H is talking about his involvement in the 1994 clash between the Indonesian military and East Timorese demonstrators in Dili. In (29), the reaction of the East Timorese diaspora in Portugal to the results of the referendum in East Timor is described by the same speaker.

(28) *Essa   é   outro   massacre   que   **se   existe***
     that._{fem} be.PRES.3sg other._{masc} massacre._{masc} that   MM   exist.PRES.3sg
     em
     in
     Timor   que   ninguém   sabia (Conv. 5, Inf. H)
     Timor   that   nobody   know.IMPF.3sg
     'That is another massacre that happens in Timor that nobody knew about.'

(29) *Aqui   em   Portugal,   naquela   altura   havia   muitos*
     Here   in   Portugal   at.that   time   exist.IMPF.3sg many
     timorenses   uma
     Timorese   a
     grande   comunidade   de   timorenses   que   estavam   aqui   a
     big   community   of   Timorese   who   be.IMPF.3pl   here   PRT
     viver,   todos   eles,
     live.INF   all   they
     pronto,   a   maioria   das   pessoas   ficaram   muito   contente,
     well   the   majority   of.the   people   be.PAST.3pl   very   happy._{sg},
     ficaram   muito
     be.PAST.3pl   very
     alegria,   algumas   ficaram   surpresa,   porque   nunca
     happiness,   some   be.PAST.3pl   surprise   because   never
     tinha
     have.IMPF.3sg
     esperança   de   que   o   referendo   será   irá   ser
     hope   of   that   the   referendum   be.FUT.3sg   go.FUT.3sg   be.INF
     realizado   em
     do.PP   in

> *Timor   mas   pronto*, **aconteceu**-   **se** (Conv. 5, Inf. H)
> Timor but   well   happen.PAST.3sg   MM
> 'At the time in Portugal there were many East Timorese, a big community
> of East Timorese who were living here, all of them, or perhaps better, the
> majority of the people were very happy, were very happiness, some were
> surprise, because they never had the hope that the referendum will be, is
> going to be done in Timor, but, well, it happened.'

This pattern may be the result of L1 transfer. As we can see from example (18), in
addition to the detransitivizing clitic *nak-*, the derived intransitive verb is also
followed by the reflexive clitic *–an*, in the case of expression of spontaneity, a
calque from European Portuguese lexicalized forms with *se*. It is possible, there-
fore, that the presence of the clitic *se* in (28) and (29) is a calque of *-an*.

7 Conclusion

The present study has sought to describe and discuss the SE constructions in the
Portuguese L2 of the East Timorese diaspora in Portugal. We have considered
uses of these constructions as potential instances of the initial stages of lan-
guage change. The socio-historical features of the East Timorese community in
Portugal make language use and innovation in this community of utmost inter-
est for the study of language change. While such innovations may be ephemeral
and not lead to language change in the immigration context, they may con-
tribute to the formation of a new East Timorese variety of Portuguese in East
Timor. In this former colony, the use of Portuguese was eradicated by
Indonesian rule and the restructuring of an East Timorese variety of
Portuguese was impeded by political instability in the territory and the changing
language policies of the past 35 years or so. East Timorese immigrants in
Portugal are L2 speakers of Portuguese and members of the educated elite,
many of whom will return to East Timor and are likely to lead the process of
partial restructuring in their homeland.

Considering the effect of contact on the transfer of marked features, we
investigated the extent of the presence of marked SE constructions in the
Portuguese spoken by the East Timorese in Portugal. In line with the evidence
from studies in L2 acquisition on these constructions in contact varieties (e.g.,
Brazilian Vernacular Portuguese and Mozambican Portuguese), we would have
expected a decrease in markedness.

We have observed that speakers who have sustained longer contact with Portuguese produce SE constructions of more types and in greater numbers. Overall, some of the situation types, particularly the reflexive constructions, are produced in a more target-like fashion than others. Furthermore, the passive/impersonal type is not only less often produced but also exhibits a reduction in markedness. This may be due in part to the substrate influence of Tetum Dili, which has no voice system and encodes impersonal constructions by using lexical and pronominal strategies. The reduction of markedness is also observed in the generalization of the marker *se* to all grammatical persons.

Our study further reveals that innovations are present across speakers and types of SE constructions. New structures using acquired linguistic structures in Portuguese and knowledge from previously acquired languages were found in our data. Impersonal SE constructions and spontaneous constructions are the types which exhibit the highest degrees of innovation. This is an interesting phenomenon, as it corresponds to the use of the marker in new situation types typically not overtly coded not very often observed in the initial stages of contact situations, where a reduction in markedness is typical. The presence of innovations is in line with what happens to these construction types in fully and partially restructured languages/varieties, where innovations emerge after the initial deletion of the marker present in the constructions of the lexifier languages. Further study of the Portuguese spoken in East Timor will show which innovative L2 usage patterns observed in the Portuguese of East Timorese immigrants, even if highly idiosyncratic, will be selected in the emergent restructured variety of East Timorese Portuguese.

References

Afonso, Susana. 2008. *The family of impersonal* constructions in European Portuguese: An onomasiological constructional approach. Manchester: University of Manchester dissertation.

Albuquerque, David. 2012. Especificidades do léxio do português de timor-leste. *PAPIA* 22(1). 201–223.

Azevedo, Milton. 2005. *Portuguese: A linguistic introduction*. Cambridge: Cambridge University Press.

Backus, Ad, A Seza Dogruöz & Bernd Heine. 2011. Salient stages in contact-induces grammatical change: Evidence from synchronic vs. diachronic contact situations. *Language Sciences* 33. 738–752.

Batoréo, Hanna. 2009. O português – uma língua pluricêntrica: O caso de timor-leste. De que forma of timorenses perspectivam e avaliam uma das suas línguas oficiais falada apenas por cinco por centro da população? *Revista Portuguesa De Humanidades* 13(1). 75–92.

Boxer, Charles R. 1947. *The Topasses of Timor*. Amsterdam: Indisch Instituut.

Carvalho, Maria José Albarran. 2001. Timor lorosa'e: Características das línguas crioulas e do português conservado na zona. Contribuição para a língua official. *Studies in Languages and Cultures of East Timor* 4. 20–36.

Cennamo, Michela. 1993. *The reanalysis of reflexives: A diachronic perspective*. Naples: Linguori Editore.

Clements, Joseph Clancy. 1996. *The genesis of a language: The formation and development of Korlai Portuguese*. Amsterdam: John Benjamins.

Clements, Joseph Clancy. 2009. *The linguistic legacy of Spanish and Portuguese: Colonial expansion and language change*. Cambridge: Cambridge University Press.

Clyne, Michael. 2003. *Dynamics of language contact*. Cambridge: Cambridge University Press.

Comrie, Bernard. 1981. Aspect and voice: Some reflections on perfect and passive in tense and aspect. In Philip Tedeschi & Annie Zaenen (eds.), *Syntax and semantics, vol. 14*: Tense and aspect, 65–78. New York: Academic Press.

Cyrino, Sonia, Maria Eugênia Duarte & Mary Kato. 2000. Visible subjects and invisible clitics in Brazilian Portuguese. In Mary Kato & Esmeralda Negrão (eds.), *Brazilian Portuguese and the null subject parameter*, 55–73. Frankfurt am Main/Madrid: Vervuert/Iberoamericana.

Dardjowidjojo, Soenjono. 1998. Strategies for a successful national language policy: The Indonesian case. *International Journal of the Sociology of Language* 130. 35–47.

DeKayser, Robert. M. 2005. What makes learning second-language grammar difficult? A review of issues. *Language Learning* 55(S1). 1–25.

Gadelii, Karl Erland. 2002. Pronominal syntax in Maputo Portuguese (Mozambique) from a comparative creole and Bantu perspective. *Africa and Asia* 2. 27–41.

Gast, Volker & Florian Haas. 2008. On reciprocals and reflexive of anaphors in German and other European languages. In Ekkehard König & Volker Gast (eds.), *Reciprocals and reflexives: Theoretical and typological explorations*, 307–346. Berlin: Mouton de Gruyter.

Goglia, Francesco & Susana Afonso. 2012. Multilingualism and language maintenance in the East Timorese diaspora in Portugal. *Ellipsis* 10. 73–99.

Gonçalves, Perpétua. 2004. Towards a unified vision of classes of language acquisition and change: Arguments from the genesis of Mozambican African Portuguese. *Journal of Pidgin and Creole Languages* 19(2). 225–259.

Haiman, John. 1983. Iconic and economic motivation. *Language* 59. 781–819.

Hajek, John. 2002. Language maintenance and survival in East Timor: All change now? Winners and losers. In David Bradley & Maya Bradley (eds.), *Language endangerment and language maintenance*, 182–202. London: Routledge.

Haspelmath, Martin. 1993. More in the typology of inchoative/causative verb alternations. In Bernard Comrie & Maria Polinsky (eds.), *Causatives and transitivity*, 87–111. Amsterdam: John Benjamins.

Haspelmath, Martin. 2006. Against markedness (and what to replace it with?). *Journal of Linguistics* 42. 25–70.

Haspelmath, Martin & Thomas Müller-Bardey. 2004. Valency change. In Geert Booij, Christian
　　Lehmann & Joachim Mugdan (eds.), *Morphology: A handbook on inflection and word
　　formation*, vol. 2, 1130–1145. Berlin: Mouton de Gruyter.
Heine, Bernd & Tania Kuteva. 2005. *Language contact and grammatical change*. Cambridge:
　　Cambridge University Press.
Holm, John. 2004. *Languages in contact: The partial restructuring of vernaculars*. Cambridge:
　　Cambridge University Press.
Hopper, Paul J. & Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language*
　　56(2). 251–299.
Inverno, Liliana. 2004. Português vernáculo do brasil e português vernáculo de angola:
　　Reestruturação parcial vs. Mudança linguística. In Mauro Fernández, Manuel Fernández-
　　Ferreiro & Nacy Vázquez Veiga (eds.), *Los criollos de base ibérica: ACBLPE 2003*, 201–213.
　　Madrid/Frankfurt am Main: Iberoamericana/Vervuert.
Inverno, Liliana. 2009. A transição de angola para o português vernáculo: Estudo
　　morfossintáctico do sintagma nominal. In Ana Maria Carvalho (ed.), *Português em contato*,
　　87–106. Frankfurt/Madrid: Vervuert/Iberoamericana.
Kemmer, Suzanne. 1993. *The middle voice*. Amsterdam: John Benjamins.
Kemmer, Suzanne. 1994. Middle voice, transitivity, and the elaboration of events. In Barbara
　　Fox & Paul Hopper (eds.), *Voice: Form and function*, 179–230. Amsterdam: John Benjamins.
Klaiman, M.H. 1991. *Grammatical voice*. Cambridge: Cambridge University Press.
Kliffer, Michael. 1981. Reflexive deletion in Brazilian Portuguese. In William Cressey & Donna
　　Napoli (eds.), *Proceedings of the 9th linguistic symposium on romance languages*,
　　318–330. Washington, DC: Georgetown University Press.
Kliffer, Michael. 1982. The case of the missing reflexive. *Hispania* 65(3). 424–427.
Lewis, M. Paul (ed.). 2009. *Ethnologue: Languages of the world*. 16th edn. Dallas: SIL
　　International. Online version: http://www.ethnologue.com/.
Märzhäuser, Christina. 2011. Cape Verdean creole in Lisbon: The young generation's perspec-
　　tive. In Eric A. Anchimbe & Stephen A. Mforteh (eds.), *Postcolonial linguistic voices:
　　Identity choices and representations*, 299–322. New York/Berlin: Mouton De Gruyter.
Matras, Yaron. 2009. *Language contact*. Cambridge: Cambridge University Press.
Meeuwis, Michael. 1997. *Constructing sociolinguistic consensus: A linguistic ethnography of
　　the Zairian community in Antwerp, Belgium*. Duisburg: LICCA papers.
Muhr, Rudolf. 2012. Linguistic dominance and non-dominance in pluricentric languages: A
　　typology. In Rudolf Murh (ed.), *Non-dominant varieties of pluricentric languages: Getting
　　the picture*, 23–48. Frankfurt: Peter Lang.
Naro, Anthony & Maria Marta Pereira Sherre. 2004. Sobre as origens do português brasileiro –
　　retrospectiva de um garimpo. In Inês Duarte & Isabel Leiria (eds.), *Actas do XX encontro
　　nacional da associação portuguesa de linguística*, 183–194. Lisboa: APL.
Petter, Margarida Maria Tadoni. 2009. Aspectos morfossintáticos comuns ao português ango-
　　lano, brasileiro e moçambicano. *PAPIA* 19. 201–220.
Rocha, Paulo & Diana Santos. 2000. CETEMPúblico: Um corpus de grandes dimensões de
　　linguagem jornalística portuguesa. In Maria das Graças Volpe Nunes (ed.), *V encontro para
　　o processamento computacional da língua portuguesa escrita e falada (PROPOR 2000)*,
　　131–140. São Paulo: ICMC/USP.
Santos, Diana & Eckhard Bick. 2000. Providing internet access to Portuguese corpora: The
　　AC/DC project. In Maria Gavrilidou, Elias G. Carayannis, Stella Markantonatou, Stelios

Piperidis & G. Stainhauer (eds.), *Proceedings of the second international conference on language resources and evaluation (LREC 2000)*, 205–210. ELRA

Sebba, Mark & Tony Wootton. 1998. We, they and identity: Sequential versus identity-related explanation in code-switching. In Peter Auer (ed.), *Code-switching in conversation*, 262–289. London: Routledge.

Taylor-Leech, Kerry. 2007. *The ecology of language planning in Timor-Leste: A study of language policy, planning and practices in identity construction*. Queensland: Griffith University dissertation.

Thomason, Sarah G. 2001. *Language contact: An introduction*. Washington, DC: Georgetown University Press.

Thomason, Sarah G. & Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley, CA: University of California Press.

Thomaz, Luís F. 2002. *Babel loro sa'e: O problema lingüístico de Timor Leste*. Lisboa: Instituto Camões.

Viegas, Telma. 1998. *Migrações e associativismo de migrantes: Estudo de caso timorense.* Lisboa: Fundação Oriente/Universidade Aberta.

Vitral, Lorenzo. 2006. A evolução do *se* reflexivo em português na perspectiva da gramaticalização. In Tânia Lobo, Ilza Ribeiro, Zenaide Carneiro & Norma Almeida (eds.), *Para a história do português brasileiro, vol. 6: Novos dados, novas análises*, 107–133. Salvador: Editora da UFBA.

Weatherbee, Donald. 1996. Portuguese Timor: An Indonesian dilemma. *Asian Survey* 6(12). 683–695.

Williams-van Klinken, Catharina, John Hajek & Rachel Nordlinger. 2002. *Tetun dili: A grammar of an East Timorese language*. Canberra: Pacific Linguistics.

Winford, Donald. 2003. *Introduction to contact linguistics*. Oxford: Blackwell.