This document is the Accepted Manuscript version of a Published Work that appeared in final form in **Journal of Language Contact** 15(3/4) : 512–561(2022), copyright © 2023 Brill. To access the final edited and published work see https://doi.org/10.1163/19552629-15030003

# A typological and diachronic analysis of replication: Body-part reflexives in Romance-lexifier pidgins and creoles

## Abstract

The fact that body-part reflexives (BPRs) are widespread in Romance-lexifier pidgin, creole and mixed (PCM) languages of the Atlantic area has usually been accounted for in terms of substratum influence from West African languages, in which such reflexives are common. However, this approach does not explain why BPRs are also frequently found in Romance-lexifier PCM languages like Zamboanga Chavacano and Malacca Creole, which lack a demonstrable African substrate, are spoken outside the Atlantic area and are in contact with languages that lack BPRs. Drawing on cross-linguistic as well as historical corpus data, this paper argues that the source of BPRs in these languages should be traced back to the late-medieval and early-Renaissance lexifiers. More specifically, it is proposed that speakers of Romance-lexifier PCM languages identified, recapitulated and replicated reflexive-like uses of words such as 'body' and 'head' in the lexifiers. A number of bridging contexts is argued to have fostered these processes.

## Keywords

body-part reflexives – Romance-lexifier pidgins, creoles and mixed languages – replication – grammaticalization

## 1 Introduction: Background, Aims and Scope of the Study

Studies on language typology, contact and change have frequently addressed the presence of reflexive constructions with terms denoting body parts ('body', 'head', 'skin' etc.) in African and creole languages (Muysken and Smith, 1994: 273; Heine, 2005: 247–248; Schladt, 2000: 109–110; Evseeva and Salaberri, 2018: 397 among many others). Accordingly, it has been observed that such body-part reflexives (BPRs) are widespread in pidgin, creole and mixed (PCM) languages with Romance lexifiers that

are spoken in the Atlantic area such as Haitian Creole (1a), Cape Verdean Creole of São Vicente (1b) and Papiamentu (1c), either as the only or as an alternative reflexive strategy:

| (1) | a. | *li* | *blese* | *kò* | *li* |
|---|---|---|---|---|---|
| | | 3SG.M | hurt.PST | body | 3SG.M |
| | | 'He hurt himself' (Lefebvre, 1998: 159) | | | |

(Haitian Creole)

| | b. | *el* | *matá* | *se* | *kabésa* |
|---|---|---|---|---|---|
| | | 3SG.M | kill.PST | GEN | head |
| | | 'He killed himself' (Michaelis et al., 2013) | | | |

(Cape Verdean Creole of Sâo Vicente)

| | c. | *e* | *ta* | *kita* | *nan* | *for* | *di* | *su* | *kurpa* |
|---|---|---|---|---|---|---|---|---|---|
| | | 3SG | be | take.off | 3PL | from | off | GEN | body |
| | | 'He takes them off himself/his body' (Muysken and Smith, 1994: 278) | | | | | | | |

(Papiamentu)

A number of reasons have been adduced to account for this state of affairs. The most recurrent one is that BPRs entered Atlantic pidgins and creoles through West African substrate influences, including Edo/Bini, Igbo, Yoruba, Gbe and other Kwa languages (Muysken and Smith, 1994: 279–280; Lefebvre, 1998: 167–170). Alternative explanations involve contact-induced grammaticalization (Heine, 2005: 247–248), independent development of BPRs based on universal cross-linguistic tendencies (Carden, 1993: 106) and transfer of BPRs from medieval and Renaissance-period lexifier Romance languages, notably Old French and Early Modern French, by means of superstrate effect (Chaudenson, 1973: 368).

All of these proposals are questionable to some extent. The BPRs of Haitian Creole (i.e., *tèt* 'head' + pronoun, *kò* 'body' + pronoun), on the one hand, have been argued not to be grammatically equivalent to comparable constructions either in French (pronoun + *même*) or in Gbe (verb + *wù* 'body'), which in this case speaks against both the 'substrate' and the 'superstrate' hypotheses (Lefebvre, 1998: 165, 169). On the other

hand, the ‘superstrate hypothesis’ has been argued to be unsustainable because of an alleged absence of BPRs in 16th-to-19th-century Romance languages, particularly French (Muysken and Smith, 1994: 273).

A problem with these critiques is that they rarely adopt a broader cross-linguistic perspective. Thus if one takes a look at the world-wide prevalence of BPRs in PCM languages with Indo-European lexifiers, a few facts suggest that there must be some specific motivation for the predominance of such reflexive constructions in these languages: (i) Schladt (2000: 110) counts 89/148 (60.1%) world-wide languages whose reflexive constructions originate in BPTs; (ii) 35/76 (46.1%) of all PCM languages across the world have BPRs (Heine, 2005: 228; Michaelis et al., 2013); (iii) 32/35 (91.4%) of these PCM languages with BPRs have Indo-European (i.e., Germanic and Romance) lexifiers (ibid.); (iv) 32/58 (55.2%) of PCM languages with Indo-European lexifiers have BPRs (ibid.); (v) 13/77 (16.9%, or one in six) of all languages known to have reflexive constructions based on the BPT ‘head’ are Indo-European-lexifier PCM languages (Evseeva and Salaberri, 2018: 397). These data lead to the generalization that BPRs in contact varieties with Indo-European lexifiers are slightly more frequent than on average (55.2% vs. 46.1%), whereas the proportion of pidgins and creoles with Indo-European lexifiers is extremely high (91.4%) among contact languages with BPRs, and considerable (16.9%) among all languages with ‘head’-reflexives.

In summary, then, PCM languages with Indo-European superstrate languages have a higher-than-average tendency of developing BPRs, independently of their substrate and of the linguistic area in which they are spoken. This fact has been pointed out in previous literature (Carden, 1993), but has not been accounted for. Moreover, the aforementioned data ((i)–(v)) render less plausible the view that BPRs in PCM languages with Indo-European lexifiers should be explained in universalist terms, that is to say, by arguing that the BPRs in question are due to inherent properties of creole languages in general (McWhorter 2002: 15–19) or the result of cross-linguistically recurrent grammaticalization tendencies (Carden, 1993: 106; Schladt, 2000: 108–110).

The ‘substrate’ and ‘areal influence’ hypotheses are likewise problematic because representative sampling and intralingual variation is often not taken into account by research on the topic (see Michaelis 2020 for general concerns and newer thoughts on representative sampling in creole contexts). More importantly, though, studies

advocating these hypotheses cannot account for the presence of BPRs in Romance-lexifier PCM languages which fulfill three distinctive conditions: (i) they do not have a demonstrable African substrate; (ii) they are spoken outside the Atlantic area (i.e., in the Indian Ocean and Pacific Ocean areas); (iii) they are in contact with languages that lack BPRs.[1] A handful of languages share these properties, including Réunion Creole (2a), Zamboanga Chavacano (2b) and Malacca Creole (2c):

| (2) | a. | *largg* | *lë* | *kor* |
|---|---|---|---|---|
| | | let | DEF | body |
| | | 'To let oneself go' (Staudacher-Valliamée, 2004: 125) | | |

(Réunion Creole)

| | b. | *ya* | *kulgá* | *éle* | *desuyo* | *kwérpo* |
|---|---|---|---|---|---|---|
| | | PFV | hang | 3SG | GEN | body |
| | | '(S)he hanged herself/himself' (Lipski and Santoro, 2007: 394) | | | | |

(Zamboanga Chavacano)

| | c. | *teng* | *ńgua* | *omi* | *ki* | *ja* | *matá* | *korpu* | *na* | *Muar* |
|---|---|---|---|---|---|---|---|---|---|---|
| | | be | one | man | REL | PFV | kill | body | LOC | Muar |
| | | 'There was a man who killed himself in Muar' (Baxter, 1988: 206) | | | | | | | | |

(Malacca Creole)

In view of this data situation, the following hypothesis is put forward in this paper: BPRs in Romance-superstrate PCM languages are to be traced back to the late-medieval and early-Renaissance lexifiers, and the mechanisms involved in this development include identification, recapitulation and replication of reflexive structures in contact grammaticalization; see Section 2.2 for details.[2] The approach adopted here is grounded on comparative data and corpus analyses, which will be laid out in Section 3.

This study is thus structured as follows: Sections 2.1–2.2 define the most relevant

1 These include Tagalog, cf. Schachter and Otanes (1972: 138–141) for an overview of its reflexive markers, and a number of Dravidian languages, cf. Andronov (2003: 169–170) for reflexivity in this language family.

2 This proposal therefore rejects, in line with Fernández (2012), the so-called monogenetic hypothesis, i.e., the view that all Romance PCM languages derive from an early form of Portuguese-based pidgin spoken in various places in Africa and Asia during the 15th and 16th centuries. For arguments in support of this view, see Naro (1978), among others.

concepts and discuss the paper's hypothesis in detail. Section 3.1 lays out the sources and methodology. Sections 3.2-3.3 provide an analysis of specific examples of BPRs in the languages under discussion, and the textual, historic and sociolinguistic dimensions of the advocated hypothesis are discussed in Section 4. Finally, Section 5 concludes.

## 2 The Domains of Inquiry

### 2.1 *Reflexivity: Definition, Types and Related Categories*

A broad definition will be adopted here for all reflexivity-related grammatical phenomena: Faltz (1977: 3–4) defines an "archetypal reflexive context" as one in which a simple clause, consisting minimally of one verb, expresses a two-argument predication, one of which is a human agent or experiencer and the other a patient. Crucially, both arguments share the referent, and they may be overtly or covertly realized. In line with widespread use, one of these arguments will be called 'reflexive marker' and the other 'antecedent' (Faltz, 1977: 21; Kemmer, 1993: 44).

A number of typologies of reflexive markers have been proposed in recent decades; see Haspelmath (forthcoming) for an overview. A basic distinction differentiates between verbal reflexives, i.e., reflexives encoded by means of affixes and adverbs, and nominal reflexives, i.e., reflexives expressed through nouns and pronouns. From a diachronic perspective, these strategies are presumably arranged along a grammaticalization cline: reflexive markers are believed to arise as nominal reflexives and then become progressively grammaticalized until they merge morphologically with the verb, thus becoming verbal reflexives (Heine, 1999: 3–4, Schladt, 2000: 113). More details are discussed in Section 2.2.

In languages with at least two different reflexive means, the choice of strategy often correlates with predicate type: some reflexive markers are used with so-called 'introverted' verbs, i.e., verbs of body grooming, physical movement and positioning of the body such as 'to wash', 'to dress' and 'to shave', which "refer to actions which one generally performs upon one's self", whereas other reflexive markers are used with so-called 'extroverted' predicates such as 'to hate', 'to kill' and 'to see', which "describe actions which the subject usually performs towards others" (Haiman, 1983: 303).

Accordingly, it has been argued that, whenever a language contrasts two reflexive strategies, the more complex one tends to be used with extroverted verbs (König and Siemund, 2000: 58). Therefore, an analysis of BPRs in Romance-lexifier PCM languages should take into account this distinction between introverted and extroverted predicates (cf. Section 3).

Concerning categories related to reflexivity, one should first mention so-called ‘emphatic pronouns’ (Heine, 1999; Schladt, 2000) or ‘secondary reflexives’ (Faltz, 1977). The main difference between emphatic pronouns and reflexives lies in their syntactic behavior: while the former function as peripheral arguments of noun and verb phrases, the latter occur as core arguments (König and Siemund, 2000: 43, 50). In many languages emphatic and reflexive markers also have different diachronic origins and a different syntactic distribution, which motivates distinguishing between the two (König and Siemund, 2000: 41).

Another category associated with reflexivity is middle voice. According to Zúñiga and Kittilä (2019: 175–176), the middle should best be characterized, on the one hand, as a kind of form that can correspond in some languages to a number of diatheses including reflexive, causative and reciprocal, and, on the other hand, as a kind of syncretism caused by the semantic relatedness of different event types such as reflexive, reciprocal and anticausative. From a semantic point of view, middle markers tend to indicate actions in which the initiator is also an endpoint, and in which some aspects of the internal structure of the denoted event are considered less important from the speaker’s point of view (Kemmer, 1993: 243). Markers of middle voice are thus used to denote events in which the subject is the site of the action such as grooming (‘to comb’, ‘to wash’), change of body posture (‘to stand up’, ‘to sit down’), translational motion (‘to place’) and emotional reaction (‘to become angry’), among others (Kemmer, 1993: 18; Zúñiga and Kittilä, 2019: 170–171).

Despite the fact that reflexivity and middle voice sometimes overlap semantically and formally, the former should not be regarded as a subtype of the latter, or vice versa. Evidence that these are two close yet distinct grammatical categories stems from the number of semantic roles invoked by each, i.e., two by reflexivity and one by middle voice (Heine, 1999: 4). Accordingly, in some languages reflexive and middle are expressed by different formal means, and may therefore be argued to constitute different

categories. This is the case of Tz'utujil, where an infix *-j-* or *-ʔ-* indicates middle voice, whereas the reflexive/reciprocal marker is a noun *-iiʔ* (Dayley, 1981: 465, 476). In short, then, emphatic pronouns and middle voice are known to strongly interact and sometimes overlap with reflexive markers cross-linguistically.

## 2.2 *Grammaticalization of BPTs as Reflexive Markers under Contact*

Grammaticalization is understood here as a "subset of linguistic changes whereby a lexical item or construction in certain uses takes on grammatical characteristics, or through which a grammatical item becomes more grammatical" (Hopper and Traugott, 2003: 2). Regarding the specific stages of grammaticalization, four interrelated mechanisms are relevant: desemanticization or semantic bleaching (the form loses its meaning content), extension or context generalization (the form is used in new contexts), decategorialization (the form loses its morphosyntactic properties) and erosion or phonetic reduction (the form loses its phonetic substance) (Kuteva et al., 2019: 3). As far as the grammaticalization of BPTs into reflexive markers is concerned, several stages can be distinguished both at the semantic (desemanticization, extension) (3a–c) and formal (decategorialization, erosion) levels (4a–d, Schladt, 2000: 113–116):[3]

**Semantic change**

(3) a. Stage 1: the BPT is the object of the clause and has only its source meaning.

b. Stage 2: the BPT is reinterpreted by means of synecdoche and starts to stand for the subject referent, thus acquiring a reflexive function. The expression is, however, still ambiguous in the sense that it can have both the source and the target meaning.

c. Stage 3: the BPT functions only as a reflexive and can develop new uses such as the reciprocal one.

**Formal change**

(4) a. Stage 1: the BPT behaves as a full noun phrase both morphosyntactically (case marking, agreement) and syntactically (word order permutations, presence of a possessive pronoun).

3 The various stages of grammaticalization (both at the semantic and formal levels) are illustrated in Sections 3.2–3.3.

b. Stage 2: the BPT may optionally display reduced behavior morphosyntactically (loss of agreement) or syntactically (constraints on word order, elision of the possessive pronoun).

c. Stage 3: the BPT shows constrained syntactic behavior, in the sense that it must be co-referential with the subject, is confined to one particular function within the clause and may not undergo word order permutations such as topicalization.

d. Stage 4: the BPT does not behave morphosyntactically as a noun phrase anymore, but rather has the properties of a pronoun.

Two points stand out concerning Schladt's (2000: 113–116) grammaticalization path for reflexive markers: on the one hand, this model does not consider anaphoric uses to be part of the process. However, if one considers the claim that reflexive markers arise as noun phrases and become pronouns in their path towards grammaticalizing into affixes (Heine, 1999: 3–4; Schladt, 2000: 113), it may be assumed that, at least in some cases, noun phrases acquire anaphoric properties without necessarily becoming reflexive. There seems to be some cross-linguistic evidence for this path (Evseeva and Salaberri, 2018: 424), which may be illustrated as follows (5):

(5) NOMINAL > ANAPHORIC > REFLEXIVE > OTHER

On the other hand, the synecdoche/metonymy mechanism (3b) seems to play a major role in the use of BPTs as markers of spatial and grammatical relations. This may be related to the following matter: apparently, the human body is perceived as being divided into a few prototypical parts such as 'leg', 'back', 'head' and 'eye'. This corresponds to the principle of prototypicality of Natural Semantics, conceived of as a reflection of the properties of human cognition (Geeraerts, 1985: 127; though see Schladt, 2000: 112). Accordingly, body parts that are perceived to be prototypical will be more prone than non-prototypical parts to encode, via synecdoche, the body as a whole. This explains why words denoting non-prototypical body parts such as 'knee' and 'ear', are less prone to grammaticalize into reflexive markers.

The hypothesis presented here draws on the stages and mechanisms of

grammaticalization of reflexive markers discussed so far. In that respect, however, a few properties of this process of language change should be pointed out that have not been the object of much research in terms of contact-induced grammaticalization. First and foremost, it should be pointed out that grammaticalization often ensues in such a way that new stages emerge continually, but this does not necessarily imply that older stages are completely discarded. Indeed, older stages may coexist and even interact with newer ones; this phenomenon has been referred to as 'layering' (Hopper, 1991: 22). Moreover, when a lexical form shifts into a grammatical one in grammaticalization, some traces of the original lexical meaning may still be carried on by the form undergoing this process. Consequently, details on the diachrony of the grammatical form may be reflected in constraints on its use and distribution; this is the so-called persistence principle of grammaticalization (Hopper, 1991: 22).

The fact that grammaticalization involves layering and persistence of older uses suggests that, in contact situations, speakers of contact varieties are exposed to grammaticalized forms coexisting with and reflecting properties of earlier ((more) lexical) stages of similar forms in the substrate or superstrate languages (Matras and Sakel, 2007: 852). This indicates that speakers of PCM languages might be able to identify older stages and properties of grammaticalized forms in both the substrate and the superstrate (Ziegeler 2014: 133–134).[4] Identification of less grammaticalized forms can then be exploited for the more functionally transparent communicative needs of contact situations (Ziegeler 2014: 107), which are characterized by a higher communicative pressure that often leads to accelerated grammaticalization (Hagège, 1993: 130).

Indeed, grammaticalization often ensues at a faster pace in contact varieties than in non-PCM languages (Heine and Reh, 1984: 87–90; Hagège, 1993: 128–130; Ziegeler, 2000: 12, 2014: 108–109; Bruyn, 2009: 321; Michaelis and Haspelmath, 2020; though see also Ziegeler and Lee, 2019: 745). The faster grammaticalization rate of contact

4 In creolistic terms, the language that provides the model for replication (the M(odel)-language) could arguably be associated with the substrate or superstrate, whereas the language that makes use of that model (the R(eplica)-language) may be associated, in this case, with a contact variety. This assumption seems implicit to the definition of ordinary contact-induced grammaticalization: Heine and Kuteva (2003: 534–535, 2005: 83) illustrate this process with Tayo (a French-based creole of New Caledonia), which replicated the semantic category of dual number from Drubéa and Cèmuhî, two of the main Melanesian substrate languages that exerted an influence on Tayo. Later work on grammaticalization in contact languages (Mufwene, 2008: 176; Bruyn, 2009: 325; Ziegeler, 2014: 111, 2017: 312; Kuteva, 2017: 164 among others) seems to follow the same line, whether implicitly or explicitly.

languages should be ascribed to the fact that communicative urgency is greater in cases of linguistic contact, which leads to certain forms being used in an over-extended sense (Hagège, 1993: 130). Moreover, the size and ecology of the individual speaker community should be borne in mind, since these factors also have an impact on the speed at which innovations are selected, replicated and propagated (Ansaldo, 2009: 111). In any case, grammaticalization can also ensue at a slower rate in contact than in non-contact situations due to restraining pressure by the substrate, i.e., the substrate may cause a colonial lag, or because of a need for greater transparency of communicative intent in the contact situation (Ziegeler, 2000: 12, 2014: 136, 2017: 315).

Identification and exploitation of older stages and properties of grammaticalized forms in the substrate or superstrate language by speakers of a contact variety for functional purposes of the contact situation is labeled 'recapitulation' by Ziegeler (2014, 2017) and Ziegeler and Lee (2019). This is not to say that speakers of the contact language are aware of the diachronic stages the grammaticalized form in question went through. In fact, it is necessary for the older and newer stages of the grammaticalizing form to coexist synchronically in order for recapitulation to be possible at all (Pietsch 2009: 532–533). The process of recapitulation thus involves replication of structures of the substrate or superstrate language into the contact variety. Replication is particularly likely to occur in so-called bridging contexts, whereby "meaning B often comes into existence because a regularly occurring context supports an inference-driven contextual enrichment of A to B" (Evans and Wilkins, 1985: 5).

In light of the former, it is argued in Section 3 that speakers of Romance-lexifier PCM languages of the Indian and Pacific Ocean areas were able to recapitulate older stages and properties of reflexive-like uses of BPTs in the Renaissance-era Romance lexifiers. Recapitulation then favored replication of BPRs into the contact languages in question. Moreover, replication was fostered by a number of bridging contexts in which a contextual enrichment of BPTs into reflexive markers was particularly favorable. This contextual enrichment was probably assisted by reanalysis of BPTs as reflexive markers.

Reanalysis is commonly regarded as a process which is distinct from, but closely related to grammaticalization (Hopper and Traugott, 2003: 58–59). This closeness is illustrated by the fact that reanalysis may participate in grammaticalization, as shown by

the above-mentioned relevance of synecdoche or metonymy in the grammaticalization of BPTs into reflexive markers. Metonymy is, in fact, among the most common triggers for reanalysis (Detges and Waltereit, 2002: 154). Therefore, and by extension, reanalysis may also be claimed to play a role at least in some cases of recapitulation and subsequent replication. Bearing this in mind, one may argue that ambiguous expressions involving BPTs and which are subject to metonymy and subsequent reanalysis may constitute a potential source, via recapitulation and replication, of reflexive markers.

As a matter of fact, ambiguous constructions with BPTs can be found on many occasions in all of the analyzed Romance lexifier varieties. (6a) illustrates a fixed expression that was still in use in 18th-century French. The fixed expression comprises the reflexive clitic *se* and *le cœur et le corps* 'the heart and the body', which could be analyzed as serving as an intensified SELF-expression. (6b) is an example from Réunion Creole where the BPR can still be specified by a possessive pronoun. (6c) stems from early 17th-century Portuguese and shows a very similar structure as the examples in (6b), from Réunion Creole, and (6d) from Papiamentu:

| (6) | a. | *se* | *tuer* | *le* | *cœur* | *et* | *le* | *corps* |
|---|---|---|---|---|---|---|---|---|
| | | REFL | kill.INF | DEF | heart | and | DEF | body |
| | | 'To work hard at an ungrateful task (lit. to kill one's own heart and body)' (*Dictionnaire de l'Académie française*, 1694; *Dictionnaire de Trévoux*, 1704–1771; *apud* Wartburg, 1929: 1212) | | | | | | |

(French)

| | b. | *li* | *la* | *twe* | *sō* | *kor* |
|---|---|---|---|---|---|---|
| | | 3SG.M | PFV | kill | GEN | body |
| | | 'He has killed himself' (Bollée, 2000: 306) | | | | |

(Réunion Creole)

| | c. | *furtar* | *o* | *corpo* |
|---|---|---|---|---|
| | | steal.INF | DEF | body |
| | | 'To free oneself of all responsibility (lit. to steal one's own body)' (Francisco Rodrigues Lobo, 1619; Galves et al., 2017) | | |

(Portuguese)

| | d. | *skonde* | *su* | *kurpa* |
|---|---|---|---|---|
| | | hide.INF | GEN | body |
| | | 'To hide oneself' (Muysken, 1993: 300) | | |

(Papiamentu)

The hypothesis laid out in this paper is not meant to imply that all features of contact languages which are unlikely to have originated in the substrate or as a result of areal or cross-linguistic tendencies of language change should necessarily be explained away in terms of lexifier influence. Admittedly, contact languages can diverge sharply from their lexifiers in a number of traits (McWhorter, 2002: 10–12; Mufwene, 2008: 162).

Another term often used in contact studies is 'convergence', defined by Bollée (2007b [1982]: 392) as a situation whereby similar structures in two languages under contact (typically, a lexifier and a substrate language) are associated with and reinforce each other. This is to say that, if a structure in language A is more grammaticalized than a similar structure in language B, then the structure in language A triggers grammaticalization of the similar structure in language B (Kriegel et al., 2019: 321–322). This reinforcing relationship does not seem to obtain for the languages under analysis here because the substrate languages in question (Dravidian, Philippine Austronesian etc.) do not exhibit BPRs. For a slightly different definition of this term, see Matras and Sakel (2007: 835).[5]

## 3 Reflexives in Latinate Languages and Romance-Lexifier Contact Varieties

### 3.1 *Sources and Methodology*

This study draws on the historical corpora of six Latinate languages: Catalan, French, Late and Medieval Latin (LML), Occitan, Portuguese and Spanish. The time span under analysis comprises, with respect to LML, the years 500–1400, as opposed to 1000–1700 concerning all others. The aim of this choice of time is to look for a possible diachronic continuity (i.e., involving the Middle Ages, the Renaissance and the century beyond)

5 Alternative models of contact-induced language change, whether focusing specifically on grammaticalization (Heine and Kuteva, 2003: 539, 555–559, 2005: 81) or on more general mechanisms (Matras and Sakel, 2007: 841–857; Mufwene, 2008: 115–132; Kuteva, 2017: 172–181) will not be discussed here for want of space.

regarding reflexive-like uses of BPTs in the languages under scrutiny. In addition to analyzing Romance languages, we choose to scrutinize LML sources since varieties of spoken Latin are not only the basis for the evolution of the Romance languages but in early medieval periods, there is often a complex multilingual situation found in the geographical areas of interest with Latin being one of the major languages involved. Moreover, written Latin also continues to be a major influence throughout the Middle Ages and in certain genres it constitutes *the* model language (especially in religious contexts), which also implies that certain discourse traditions are maintained in Romance texts (Lodge, 1993: 34ff., 163–164; Wright, 2017).

The data sources comprise digital historical corpora and dictionaries, including Torruella et al. (2013) and Alcover and Moll (2002) for Catalan, Guillot-Barbance et al. (2017) and Wartburg (1929) for French, Ricketts et al. (2003) and Stempel et al. (2013) for Occitan, Galves et al. (2017) for Portuguese and Real Academia Española (1994) for Spanish. LML data, on the other hand, have been collected from dictionaries (Blaise, 1975; Niermeyer, 1976; Migne, 1977), etymological dictionaries (Ernout and Meillet, 2001; De Vaan, 2008), text editions (Ramos, 2000), text collections (Le Blant, 1856; Lemay, 2017), digital corpora (Stotz et al., 1959–2015; Guillot-Barbance et al., 2018; Quetglas and Gómez, 2019) and others (Pérez, 2007).

One main focus concerning the analyzed material was on texts with a religious outlook and thematic links to the Bible. These involve epic poems, early translations of (parts of) the Bible into vernacular languages as well as other kinds of religious writings, including doctrines and lives of saints. Two reasons motivate this choice: (i) because of their formal similarities, these texts constitute a suitable basis for comparison between Romance languages and contact varieties; (ii) many of these kinds of writings were presumably involved in the evangelization processes of PCM language-speaking populations in colonized countries; therefore, in accordance with the hypothesis laid out in Sections 1 and 2.2 they may indeed be seen as a global source of ambiguous structures in creolization and as the main model for recapitulation and replication by speakers of contact varieties.

In addition, it should be borne in mind that many of the traits found in Romance-lexifier PCM languages are also believed to emanate from colloquial and dialectal Renaissance varieties (Chaudenson, 1973: 345; Bollée, 2000: 306–307; Kriegel et al.

2019).[6] Consequently, texts which reflect colloquial and dialectal Romance languages have also been considered. Most of these data have been extracted from dictionaries and collections of phrases, the latter of which reflect popular use. Phrases and proverbs moreover have the advantage of preserving syntactic relics longer than other parts of grammar (Harris and Campbell, 1995: 354–355). Details regarding the textual, historic and sociolinguistic factors potentially involved in the emergence of BPRs in Romance-lexifier PCM languages are discussed in Sections 3.2–3.3 and 4 below.

Concerning contact languages, compilations of early (18th–20th century) texts, grammatical outlines and etymological dictionaries of four main PCM language groups have been considered:[7] Indian Ocean Creole French (Mauritius, Réunion, Rodrigues and Seychelles), Indian Creole Portuguese (Diu, Kannur, Kochi, Korlai, Mangalore and Sri Lanka), Southeast Asian Creole Portuguese (Batavia (Tugu), Bidau (East Timor) and Malacca) and Philippine Creole Spanish (Cavite, Cotabato, Ternate and Zamboanga). Examples of BPRs in these languages have been classified into a typology of reflexives and compared to occurrences of reflexive-like uses of BPTs in the lexifier languages. Details regarding all textual sources are exhibited in the Appendix.[8]

Further Romance-lexifier PCM languages of the Southeast Asian-Pacific areas such as Tayo and Vietnamese Pidgin have not been taken into consideration, since these varieties have a different past (they are spoken in areas that were colonized later) which may have influenced the data situation regarding reflexive markers or are only scarcely documented (e.g., for Davao and Ermita Chavacano, Nagapattinam, Macau, Bengal and Burma Creole, see Arana-Ward (1977), Lipski (1987), Syea (2017) and Cardoso (2020)).

## 3.2 *Reflexives in Late-Medieval and Renaissance Latinate Varieties*

### 3.2.1 Late and Medieval Latin

In Classic and Vulgar Latin personal pronouns such as *mē*(*d*) '1SG.ACC', *tē*(*d*) '2SG.ACC' and *nōs* '1PL.ACC' serve, in object position, as reflexive markers of first and second

6 This is probably true only for some of the contact varieties under analysis. Admittedly, the formation process of each PCM language is unique and can hardly be extrapolated to others (Sippola, 2011: 7).
7 The use of these specific labels is meant to reflect the predominant lexifier and to allow areal grouping.
8 Sources cited only in examples or in Section 3.1 are provided exclusively in the Appendix. Accordingly, for all sources not listed in the bibliography the reader is referred to the Appendix.

person (7a). In the third person, on the other hand, reflexivity is usually encoded by the etymologically obscure reflexive pronoun *sē* (7b, Weiss, 2009: 325–333):

| (7) | a. | *si* | *ego* | *me* | *ad* | *iudicem* | *sic* | *defend-am* |
|---|---|---|---|---|---|---|---|---|
| | | if | 1SG.NOM | 1SG.ACC | before | judge.ACC | thus | defend-1SG.FUT |
| | | 'If I shall thus defend myself before the judge' (Cicero's *Tusculanae Quaestiones* 29.5, 1st century BCE; *apud* Viti, 2009: 149) | | | | | | |

(Classical Latin)

| | b. | *is* | *non* | *inprudentiā* | *se* | *defend-et* |
|---|---|---|---|---|---|---|
| | | 3SG.M | NEG | carelessness.ABL | REFL | defend-3SG.PRS |
| | | 'He does not defend himself with carelessness' (*Rhetorica ad Herennium* 2.24.5, 1st century BCE; *apud* Viti, 2009: 149) | | | | |

(Classical Latin)

In addition to the pronominal strategy, Classic and Vulgar Latin also encode reflexivity by means of the mediopassive voice marker -(*tu*)*r* in order to denote events in which the subject is the site of the action (8a–b, cf. Section 2.1):[9]

| (8) | a. | *lav-o-r* |
|---|---|---|
| | | wash-1SG-MID |
| | | 'I wash (myself)' (Weiss, 2009: 381) |

(Classical Latin)

| | b. | *ferr-ō* | *accing-o-r* |
|---|---|---|---|
| | | sword-ABL | gird-1SG-MID |
| | | 'I gird myself with a sword' (Weiss, 2009: 381) | |

(Classical Latin)

Already in Classic and Vulgar Latin, and even more so in LML, the BPTs *caput* 'head' and *corpus* 'body' are frequently used in a metonymic sense, i.e., in order to denote the whole self: "*Caput* [est] souvent employé dans de sens dérivés ou imagés pour désigner

9 For a comprehensive overview regarding the range of uses of the suffix -(*tu*)*r,* its interaction with reflexivity and evolution over time in Latin see Cennamo (2009).

la personne tout entière” (Ernout and Meillet, 2001: 98). As a result, both these terms increasingly function as anaphora, i.e., as devices that refer to antecedents mentioned in previous discourse (9a–c):

| (9) | a. | *corpus* | *reg-is* |
|---|---|---|---|
| | | body.NOM | king-GEN |
| | | ‘The king (lit. the king’s body)’ (Blaise, 1975: 256) | |

(Late and Medieval Latin)

| | b. | *corpus* | *castr-i* |
|---|---|---|---|
| | | body.NOM | castle-GEN |
| | | ‘The castle (lit. the castle’s body)’ (Migne, 1977 [1858]: 619) | |

(Late and Medieval Latin)

| | c. | *in* | *eodem* | *monte* | *XV* | *capita* | *amplius* |
|---|---|---|---|---|---|---|---|
| | | on | same | mountain.ABL | XV | head.PL | more.than |
| | | *noscuntur* | *esse* | *interfecta* | | | |
| | | know.3PL.PASS.PRS | be.INF | killed.F.PL | | | |
| | | ‘More than fifteen people (lit. heads) are known to have been killed on that same mountain’ (*Colección documental del monasterio de Santa María de Otero de las Dueñas*, I.20.37, 876; *apud* Pérez, 2007: 316) | | | | | |

(Late and Medieval Latin)

A direct outcome of this state of affairs is for anaphoric uses of *caput* and *corpus* to grammaticalize, via the grammaticalization path sketched in (3)–(4) above (Section 2.2), into reflexive markers. Accordingly, reflexive uses of the BPTs in question are not uncommon in LML (10a–d):

| (10) | a. | *caput* | *meum* | *in* | *illorum* | *protectionem* | *ad* |
|---|---|---|---|---|---|---|---|
| | | head | 1SG.POSS.ACC | in | 3PL.POSS.ACC | protection.ACC | at |
| | | *clausuram* | *collocavi* | | | | |
| | | closure.ACC | place.1SG.PST | | | | |
| | | ‘I placed myself (lit. my head) under their protection under lock and | | | | | |

| | | key' (*Traditiones Frisingenses*, 7th–12th centuries; Stotz et al., 1959–2015) |
|---|---|---|

(Late and Medieval Latin)

| | b. | *si* | *posseat* | *fidare* | *et* | *credere* | *in* | *illum* |
|---|---|---|---|---|---|---|---|---|
| | | if | be.able.3PL | entrust.INF | and | commend.INF | in | 3PL |
| | | *de* | *suo* | *corpus* | *vel* | *de* | *sua* | |
| | | of | 3SG.POSS.M | body | or | of | 3SG.POSS.F | |
| | | *honore* | | | | | | |
| | | honor | | | | | | |
| | | 'If they are able to entrust and commend themselves (lit. their bodies) or their honor to them' (*Liber Feudorum Maior*, 1118; Quetglas and Gómez, 2019) | | | | | | |

(Late and Medieval Latin)

| | c. | *nec* | *crederem* | *capiti* | *meo* | *de* | *exposicione* |
|---|---|---|---|---|---|---|---|
| | | neither | trust.1SG.PRS | head.DAT | 1SG.POSS.DAT | of | exposition |
| | | *sensus* | *litteralis* | | | | |
| | | sense.GEN | literal.GEN | | | | |
| | | 'Neither do I trust myself (lit. my head) to express a literal sense' (John Wycliffe's *Tractatus de benedicta incarnatione*, 65, ca. 1400) | | | | | |

(Late and Medieval Latin)

| | d. | *in* | *primis* | *concedo* | *corpus* | *meum* |
|---|---|---|---|---|---|---|
| | | in | first | confer.1SG.PRS | body | 1SG.POSS.ACC |
| | | *ad* | *domum* | *S.* | *Cucuphati* | |
| | | to | house.ACC | Saint.GEN | Cucuphas.GEN | |
| | | 'First of all, I confer myself (lit. my body) to the House of Saint Cucuphas' (*Cartulario de Sant Cugat del Vallés*, 1097; Quetglas and Gómez, 2019) | | | | |

(Late and Medieval Latin)

The BPTs in (10a–d) can each be interpreted literally ('head', 'body') or as standing for

the subject referent. Furthermore, they display the properties of a full noun phrase concerning case marking (accusative, dative) as well as the presence of possessive pronouns (*meum*, *suo*, *meo*). This suggests that examples (10a–d) illustrate Stage 2 of the semantic (3b) and Stage 1 of the formal (4a) level of the grammaticalization cline for reflexive markers. Nevertheless, it may be no coincidence that most reflexive-like uses of BPTs in this language tend to be direct objects, as also shown by (10a–d). This may be indicative of slightly constrained syntactic behavior, which is a feature of Stage 2 of the formal cline. In any case, BPRs of this kind can never have been widespread enough to represent the primary reflexive strategy of LML. In fact, they seem to co-occur with a specific set of verbs including *collocāre* 'to place', *fīdere* 'to (en)trust', *crēdere* 'to believe', *concēdere* 'to confer' and *pōnere* 'to place', among others.

### 3.2.2 Romance languages (Catalan, French, Occitan, Portuguese, Spanish)

Romance languages use the *me te se* paradigm as a primary option to express reflexivity. Besides being a reflexive marker, especially *se* has evolved to mark a variety of constructions that deal with argument alternations and promoting a theme or patient argument to the subject position. *Se* can be found to mark typical middle environments (see Kemmer, 1993) and anticausative constructions (cf. Haspelmath, 1993 among many others); it can function as a passive marker and as a marker of impersonal active constructions that are close to so-called *man*-impersonals in their meaning and functional range (cf. Cennamo, 1993; Wolfsgruber, 2017a–b among many others).

Beside this highly grammaticalized reflexive marker, we also find BPTs in contexts where they express anaphoric relations and in which it can be argued that these body parts function as a way of expressing reflexivity. The contexts that seem to be most crucial for our cause can roughly be grouped into three main patterns, which will be exemplified with the help of French examples:

(i) constructions in which an opposition between 'body' and 'soul' is created explicitly. These contexts may indicate a crucial point of departure for nouns derived from Latin *corpus* to be used as reflexive markers. They are often found in biblical texts and contexts that have to do with the salvation of the soul. An example of this is found in (11). Medieval literature is often heavily influenced by Christian moral concepts and, in

fact, a majority of texts is linked to Christian ideas (and the Bible) or to political issues. The duality of flesh (body) and the soul is one of the most prominent concerns in the religiously motivated texts of that time.[10]

| (11) | *qui* | *croit* | *et* | *aimme* | *fole* | *famme* | *il* | *gaste* | *avoir* |
|---|---|---|---|---|---|---|---|---|---|
| | who | believe.3SG | and | love.3SG | crazy | woman | 3SG.M | waste.3SG | wealth |
| | *et* | *cors* | *et* | *ame* | | | | | |
| | and | body | and | soul | | | | | |
| | ‘Whoever believes and loves a foolish woman, loses wealth and both body as well as soul’ (*Anciens proverbes*, ca. 13th–15th centuries; Guillot-Barbance et al., 2017) | | | | | | | | |

(French)

(ii) constructions that involve one’s body in battle activities, which often include a possessive pronoun + body part. These instances commonly feature the offering of one’s body to the battle or fighting head-to-head. This type of construction is illustrated by the following example:

| (12) | *l’=une* | *fois* | *quant* | *il* | *abandonna* | *le* | *prince* |
|---|---|---|---|---|---|---|---|
| | DEF=one | time | when | 3SG.M | abandon.3SG.PST | DEF | prince |
| | *son* | *lige* | *et* | *ala* | *avec* | *le* | *doc* |
| | 3SG.POSS | liege | and | go.3SG.PST | with | DEF | duke |
| | *d’=Atthenes* | | *qui* | *estoit* | *revellez* | *contre* | *le* |
| | of=Athens | | who | AUX | rebel.PTCP | against | DEF |
| | *prince,* | *et* | *porta* | | *armes* | *et* | *se* |
| | prince | and | carry.3SG.PST | | weapons | and | REFL |
| | *combati* | | *cors* | *a* | *cors* | *a* | *son* |
| | fight.3SG.PST | | body | to | body | to | 3SG.POSS |
| | *seignor* | *lige* | | | | | |
| | lord | liege | | | | | |
| | ‘Then one time when he abandoned his allegiance to the prince and went with | | | | | | |

10 The picture is, of course, more complex: there are nuanced constellations of this dual representation that originate in much more ancient times. Delving deeper into details would lead us too far astray; for a concise overview, see e.g. Baschet (2000).

| | the Duke of Athens, who rebelled against the prince, and he carried weapons and fought head-to-head with his liege lord's army' (*Chronique de Morée*, 223, ca. 1320; Guillot-Barbance et al. 2017) |
|---|---|

(French)

(iii) The third and seemingly more advanced means of expressing reflexivity are instances with extroverted verbs or verbs that express (a change in) body postures. An example of this construction type is illustrated in the following sentence, in which a body is put somewhere, e.g. in the proximity of another person or body or into a new/different location, etc.

| (13) | *et* | *fut* | *recrust* | *jusques* | *a* | *chest* | *premiers* |
|---|---|---|---|---|---|---|---|
| | and | AUX | recruit.PTCP | until | to | this | first |
| | *ples* | *apres* | *chest* | *prochain* | *eschequier,* | *a* | *rendre* |
| | plea | after | this | next | parliament. meeting | to | surrender.INF |
| | *son* | *cors* | *en* | *prison* | | | |
| | his | body | to | prison | | | |
| | 'And by the time of this first plea, after the next parliament meeting, he was required to surrender himself (lit. his body) to prison' (*Plaids de Mortemer*, 60, ca. 1320; Guillot-Barbance et al. 2017) | | | | | | |

(French)

The *body* vs. *soul* opposition reaches into the semantic field of *offering one's own body to fight*. This is the case in battle situations, where there is often a notion of fate, perishing, etc. In medieval times, warfare was not a purely political issue but frequently intertwined with ecclesiastical interests, as can be observed in medieval texts such as the *Chanson de Roland* or the *Cantar de mío Cid*. Therefore, the *body* vs. *soul* dichotomy is often prominent in these contexts. Notice, as well, that the semantic field of *offering one's own body to fight* is the first one mentioned so far in this section that allows an anaphoric reading of the BPT 'body', which, depending on context, may also be understood as standing for the subject referent. Therefore, examples such as (12) —just like (15) further below— may be regarded as enabling reanalysis of the BPT 'body' as a

reflexive marker and thus opening the path for a shift to Stage 2 of the semantic level of grammaticalization.

The third kind of construction mentioned above —i.e., instances involving the positioning of a body in a new context in which the body is more on the foreground— is different and may present an important evolution, as the *body* vs. *soul* dichotomy is not evoked in these examples anymore. It is relevant to note that changes in body posture are closely connected with the typical middle contexts described in Kemmer (1993). The following is an overview of the new contexts hitherto discussed (14a–c):

(14) a. new context #1: *body* vs. *soul* dichotomy
b. new context #2: possessive pronoun + *body* involved in battle situations
c. new context #3: possessive pronoun + *body* expressing middle contexts

The three main contexts presented above with the help of medieval French material (11–13) are extant in all the Romance languages we have examined, i.e., in French, Occitan, Catalan, Spanish and Portuguese. This is not to say, however, that extension in the use of BPTs into these three new contexts (14a–b) occurs in a chronological sequence. On the basis of the available data, all three seem to have surfaced by about the first half the 14th century.

In the following, a trident of Catalan examples is presented: (15a) is again a representation for the *body* vs. *soul* dichotomy. In (15b), the hero offers his body for combat, i.e., he offers himself to the battle. (15c) is a typical example of indicating a change in body posture or location:

| (15) | a. | *ells* | *saben* | *certament* | *que* | *ell* | *és* | *perdut* |
|---|---|---|---|---|---|---|---|---|
| | | 3PL | know.3PL | surely | that | 3SG.M | be.3SG | lost.PTCP |
| | | *en* | *cos* | *e* | *en* | *ànima* | | |
| | | in | body | and | in | soul | | |
| | | ‘They surely know that he is entirely lost (lit. in body and soul)’ (*Viatge d’en Ramon de Perellós al purgatori de Sant Patrici*, 138, second half of the 15th century; Torruella et al. 2013) | | | | | | |

(Catalan)

| | b. | *yo* | *offir* | *lo* | *meu* | *cos* | *a* | *combatre* |
|---|---|---|---|---|---|---|---|---|
| | | 1SG | offer.1SG.PRS | DEF | 1SG.POSS | body | to | fight.INF |
| | | ‘I offer myself/my body to fight’ (Bagà Pere Tomic, 1438; Torruella et al. 2013) | | | | | | |

(Catalan)

| | c. | *lanç-ant* | *lo* | *cos* | *ab* | *esforç* | *de* | *amor* | *cruel* |
|---|---|---|---|---|---|---|---|---|---|
| | | throw-PTCP | DEF | body | with | strain | of | love | cruel |
| | | *pesada* | *sobre=l* | *de* | *Leànder* | | | | |
| | | heavy | on=DEF | of | Leànder | | | | |
| | | ‘Throwing her heavy self/her body, with strain of cruel love, on top of Leànder’s’ (Roís de Corella, *Proses mitològiques*, 173, ca. 1460; Torruella et al. 2013) | | | | | | | |

(Catalan)

While Occitan also displays BPT constructions that stress the opposition between *body* and *soul* as well as body positioning, it has another very intriguing feature. According to Jensen (1986), *mos cors* ‘my body’ can likewise function as an anaphor, and it is not exclusively found with the first person singular, but it can also be used in the third person or second person plural, as in *sos cors* and *vostre cors* respectively. Jensen (1986: 88) moreover links these Occitan expressions with the French phrase *à son corps défendant* ‘reluctantly’.

These different developments are proof of how present the concept of body (and soul) was in the Gallo-Romance area during the period in question.[11] The following is an example of *vostre cors* acting as an anaphor (16). Here it is particularly noteworthy that *mos cors* co-occurs with reflexive *se*-verbs as in *vostre cors s'orgouilla*, ‘your body/you take(s) pride’:

| (16) | *meravill* | *me* | *cum* | *vostre* | *cors* | *s'=orgouilla* |
|---|---|---|---|---|---|---|
| | amaze.3SG.PRS | 1SG.OBL | how | 2PL.POSS | body | REFL=take.pride.3SG.PRS |

11 Jensen (1990) mentions that a similar situation is also attested in Old French varieties. Some of his examples, however, do not only testify to the use of *cors* as a subject pronoun but also include cases that are more suitably ascribed to reflexive uses as in *car des que ge serai armez …, porrai mon cors deffendre contre vos* (*Mort Artu* 147.61) ‘for as soon as I am armed …, I can defend myself against you’ (Jensen, 1990: 140).

| | ‘It amazes me how you (lit. your body) take pride’ (Beatriz de Dia, ca. 1225, Jensen 1986: 88) |
|---|---|

(Occitan)

While we find the same data situation in French, i.e., *body* vs. *soul*, extroverted verbs and verbs of body posture co-occurring with body parts, as has already been shown above, we also find an array of fixed expressions that are partially still in use today and partially also attested in other Romance varieties (17a–e, also 6a and 6c above). These fixed expressions are mentioned here because, apart from the Bible and biblical texts, they are likely to have constituted a major source for anaphoric uses of BPTs in creole settings. In line with our hypothesis, the exposure of speakers of contact varieties to this kind of constructions would have enabled recapitulation and replication of reflexive-like uses of BPTs:

| (17) | a. | *mettre* | *corps* | *et* | *cœur* | *à* | *faire* | *quelque* | *chose* |
|---|---|---|---|---|---|---|---|---|---|
| | | put | body | and | heart | to | do | some | thing |
| | | ‘To give oneself much trouble’ (Estienne, 1549; Stoer, 1625; *apud* Wartburg, 1929: 1212) | | | | | | | |

(French)

| | b. | *se* | *traiter* | *bien* | *le* | *corps* |
|---|---|---|---|---|---|---|
| | | REFL | treat.INF | well | DEF | body |
| | | ‘To care about one’s own health’ (Furetière, 1690; *apud* Wartburg, 1929: 1212) | | | | |

(French)

| | c. | *faire* | *corps* | *neuf* |
|---|---|---|---|---|
| | | make.INF | body | new |
| | | ‘To recover after a long illness’ (Furetière, 1690; Larousse, 1869; *apud* ibid.) | | |

(French)

| | d. | *porter* | *bien* | *son* | *corps* |
|---|---|---|---|---|---|

| | | carry.INF | well | 3SG.POSS | body |
|---|---|---|---|---|---|
| | | ‘To have good health’ (Furetière, 1690; *Dictionnaire de Trévoux*, 1704–1771; *apud* ibid.) | | | |

(French)

| | e. | *faire* | *litière* | *de* | *son* | *corps* |
|---|---|---|---|---|---|---|
| | | make.INF | litter | of | 3SG.POSS | body |
| | | ‘To boldly expose onself to danger, to work too hard’ (*Dictionnaire de Trévoux*, 1704–1771; Féraud, 1787; Larousse, 1869; *apud* ibid.) | | | | |

(French)

In the same vein as the other Romance languages discussed so far, examples of BPTs in 15th-to-17th-century Portuguese can be classified into the three major uses that have been described above. These include statements which create an opposition between the concepts *body* and *soul* (18a), those which comprise the expression ‘to offer one’s own body to fight’ (18b) and BPTs that co-occur with extroverted verbs such as *entregar* ‘to surrender, hand over’ (18c). BPTs with introverted predicates such as *defender* ‘to defend, protect’ are also attested (18d):

| (18) | a. | *e* | *fugir-am* | *todos* | *e* | *nam* | *curar-am* |
|---|---|---|---|---|---|---|---|
| | | and | flee-3PL.PST | all | and | NEG | succeed-3PL.PST |
| | | *senam* | *de* | *salvar-em* | *seu-s* | *corpo-s* | |
| | | but | of | save-3PL.PST | 3PL.POSS-PL | body-PL | |
| | | ‘And they all fled, and they did not succeed but in saving their own bodies (i.e., as opposed to their souls) (Anonymous, ca. 1450, *Livro das estorias da biblia*; Galves et al. 2017) | | | | | |

(Portuguese)

| | b. | *offereç-ia* | *seu* | *corpo* | *a* | *pelejar* | *hum* | *per* |
|---|---|---|---|---|---|---|---|---|
| | | offer-3SG.PST | 3SG.POSS | body | to | fight.INF | one | by |
| | | *hum* | *com* | *o-s* | *filho-s* | *de* | *Isrrael* | |
| | | one | with | DEF-PL | son-PL | of | Israel | |
| | | ‘He offered himself (lit. his body) to fight one-by-one against the sons | | | | | | |

| | | of Israel' (Anonymous, ca. 1450, *Livro das estorias da biblia*; ibid.) |
|---|---|---|

(Portuguese)

| | c. | *escolh-endo* | *antes* | *entregar* | *o-s* | *seu-s* | *corpo-s,* | |
|---|---|---|---|---|---|---|---|---|
| | | choose-GER | rather | surrender.INF | DEF-PL | 3PL.POSS-PL | body-PL | |
| | | *do* | *que* | *servir* | *ou* | *adorar* | *a* | *deus* |
| | | than | SUB | serve.INF | or | worship.INF | OBJ | god |
| | | *algum* | | | | | | |
| | | any | | | | | | |
| | | 'Choosing to surrender themselves (lit. their bodies) rather than to serve or worship any god' (Ferreira de Almeida, 1681, *O novo testamento*; ibid.) | | | | | | |

(Portuguese)

| | d. | *com* | *que* | *se* | *sustentassem,* | *e* | *defendassem* | *o* |
|---|---|---|---|---|---|---|---|---|
| | | with | which | REFL | feed.3PL.SBJ | and | defend.3PL.SBJ | DEF |
| | | *corpo* | *d-as* | *injúria-s* | | *d-o* | | *tempo* |
| | | body | of-DEF.F.PL | hardship-PL | | of-DEF.M.SG | | time |
| | | 'With which they could feed themselves and defend themselves (lit. their bodies) from the hardships of the weather' (Severim de Faria, 1624, *Discursos vários políticos*; ibid.) | | | | | | |

(Portuguese)

Much the same is true of Spanish, where there are plenty of attestations involving an opposition between *body* and *soul* (19a) as well as of BPTs co-occurring with extroverted verbs such as *poner* 'to place, give up' and *querer* 'to love' (19b–c):

| (19) | a. | *salu-o* | *el* | *cuerpo* | *e* | *el* | *anima* | *e* | *la* |
|---|---|---|---|---|---|---|---|---|---|
| | | save-3SG.PST | DEF | body | and | DEF | soul | and | DEF |
| | | *hazienda* | | | | | | | |
| | | estate | | | | | | | |
| | | 'He saved his body and his soul and his estate' (García de Santa María, 1485, *Evangelios e epístolas con sus exposiciones en romance*, 42; Real | | | | | | | |

| | | |
|---|---|---|
| | | Academia Española, 1994) |

(Spanish)

| | b. | *Iudas* | *Machabeo,* | *que* | *est-aua* | *apareiado* | *de* | *morir* |
|---|---|---|---|---|---|---|---|---|
| | | Judas | Maccabee | REL | be-3SG.PST | bound | to | die.INF |
| | | *e* | *de* | *poner* | | *el* | *cuerpo* | *por* |
| | | and | to | give.up.INF | | DEF | body | for |
| | | *los* | *ciudadan-os* | | | | | |
| | | DEF.PL | citizen-M.PL | | | | | |
| | | ‘Judas Macabee, who was bound to die and to give himself (lit. his body) up for the citizens’ (Anonymous, ca. 1260, *Los libros de los Macabeos*; Real Academia Española, 1994) | | | | | | |

(Spanish)

| | c. | *más* | *querría* | *que* | *me* | *quisiese* | *como* | *a* |
|---|---|---|---|---|---|---|---|---|
| | | rather | want.3SG.CND | SUB | 1SG | love.3SG.SBJ | as | OBJ |
| | | *su* | *cuerpo* | | | | | |
| | | 3SG.POSS | body | | | | | |
| | | ‘I had rather that he would love me like (he loves) himself (lit. his body)’ (Santa Cruz, 1574, *Floresta española*, 276; Real Academia Española, 1994) | | | | | | |

(Spanish)

To summarize so far, the data suggest that in medieval and Renaissance Latinate languages the opposition between inherited *se*-reflexives and innovated BPRs represents, to a certain degree, a division of labor: whereas the first strategy tends to be used with all sorts of verbs and gradually spreads to other functional realms (i.e., passives and impersonals), the second mostly co-occurs with extroverted verbs. BPRs —which involve the use of a full noun and possessive pronoun as well as number agreement— are syntactically more complex than the pronominal *se*-reflexives. This finding is therefore in alignment with cross-linguistic observations regarding the relationship between choice of reflexive strategy and predicate type (cf. Section 2.1).

### 3.3 *Reflexives in Romance-Lexifier Contact Varieties*

#### 3.3.1 French-lexifier PCM languages

Mauritius, Réunion, Rodrigues and Seychelles Creole have three BPRs, *kor* 'body' and *lekor* 'the body', both of which can occur with or without a possessive pronoun (cf. 2a, 6b above; Chaudenson, 1973: 360); the third BPR is *latet* 'head' (Syea, 2017: 118–119). According to the grammaticalization path laid out in Section 2.2, then, these reflexive markers have reached Stage 2 of both the semantic and formal clines. They are believed to have emerged because of the loss of French clitics in the process of creolization (Syea, 2017: 115; see Heine, 2005 for a more detailed discussion).

Another reflexivization strategy involves the intensive pronoun *mem* 'self' and a form of a personal pronoun, for instance, *zotmem* 'themselves' (and *momem* 'myself', *numem* 'ourselves', etc.) (Syea 2017: 115). This constitutes the most common reflexive strategy, i.e., *mem*-reflexives are used with most transitive verbs such as *koz* 'to talk' and *esplik* 'to explain', especially if a prepositional phrase is involved (20a) (Muysken and Smith, 1994: 300). Furthermore, *mem*-reflexives are often interchangeable with plain personal pronouns (20b). Another way of obtaining a coreferential interpretation is to use plain strong pronouns such as *mwa*, *twa*, *zot* etc. (20c). For the 3rd person, ambiguity may arise between a reflexive reading and a disjoint interpretation when the plain pronoun strategy is used. Thus, another plausible meaning for (20c) is that they brought someone else ('them') into a difficult situation (Syea, 2017: 115–116).

| (20) | a. | *mo* | *pa* | *finn* | *rann* | *mwa* | *kont* | *ki* | *li* | *pa* | *ti* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1SG | NEG | PFV | take | 1SG | notice | SUB | 3SG | NEG | PST |
| | | *la* | *e* | *ki* | *mo* | *ti* | *pe* | *koz* | *ar* | *momem!* | |
| | | there | and | SUB | 1SG | PST | PRS | speak | to | REFL | |
| | | 'I didn't even realize that he was not there and that I was talking to myself!' (Police-Michel et al., 2011: 72) | | | | | | | | | |

(Mauritian Creole)

| | b. | *mo* | *pa* | *kone* | *ki* | *manyer* | *mo* | *pu* | *esplik* | *mwa/momem* |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 1SG | NEG | know | which | manner | 1SG | FUT | explain | 1SG/REFL |

| | | ‘I don’t know how I shall explain myself’ (Corne, 1988: 72) |
|---|---|---|

(Mauritian Creole)

| | c. | *zot* | *fin* | *met* | *zot* | *dan* | *enn* | *problem* |
|---|---|---|---|---|---|---|---|---|
| | | 3PL | PFV | put | 3PL | in | a | problem |
| | | ‘They brought them/themselves into some difficulty’ (Syea, 2017: 116) | | | | | | |

(Mauritian Creole)

In (20c) a reflexive interpretation is at hand without the use of any overt reflexive marker. Such verbs mainly include change in body posture such as *alonze* ‘to lie down’, *repoze* ‘to rest’, *benyen* ‘to bathe’, *asize* ‘to sit down’ and *dibute* ‘to stand up’ (Syea, 2017: 116, also Corne, 1988). These verbs, i.e., body-grooming predicates and predicates of (change in) body posture, are often characterized as part of the middle domain (see Kemmer, 1993: 16ss.). Verbs that are not to be located within these domains normally need some kind of morphological marker in order to express reflexivity. Consider (21a) below, which illustrates that a verb like ‘to put’ cannot be interpreted reflexively without a morphological marker, as opposed to (21b), which denotes a change in body posture; here the sentence is also grammatical and yields a reflexive interpretation without a morphological marker (Syea, 2017: 116–117).

| (21) | a. | **li* | *met* | *dan* | *enn* | *problem* |
|---|---|---|---|---|---|---|
| | | 3SG | put | in | a | problem |
| | | ‘He’s brought himself into some difficulty’ (Syea, 2017: 117) | | | | |

(Mauritian Creole)

| | b. | *mo* | *pe* | *alonz(e)* | *(mwa)* |
|---|---|---|---|---|---|
| | | 1SG | PROG | lay | 1SG |
| | | ‘I’m having a lie down’ (Syea, 2017: 116) | | | |

(Mauritian Creole)

According to Muysken and Smith (1994: 286–287), *kor* and *lekor* have a more reduced scope, since they mostly occur with verbs that refer to a physical action or movement, including *ran* ‘to surrender’ and *twe*/*tye* ‘to kill’ (22a–b). Of these two, *lekor* is even

more limited than *kor*, as the former only occurs with a few predicates (22c). Syea (2017:118) mentions that the BPRs *kor*, *lekor* and *latet* are only used with "[…] a handful of verbs (verbs which denote actions adversely affecting the speaker)." Those verbs include in Mauritian Creole *zete* 'to throw', *pini* 'to punish', *fatige* 'to tire/to worry' and *kase* 'to break' (22c–d).

| (22) | a. | *li* | *la* | *ran* | *sõ* | *kor* |
|---|---|---|---|---|---|---|
| | | 3SG | PFV | surrender | 3SG.POSS | body |
| | | 'He surrendered (himself) (to the police, to justice)' (Bollée, 2000: 306) | | | | |

(Réunion Creole)

| | b. | *li* | *la* | *twe* | *sõ* | *kor* |
|---|---|---|---|---|---|---|
| | | 3SG | PFV | kill | 3SG.POSS | body |
| | | 'He killed himself' (Bollée, 2000: 306) | | | | |

(Réunion Creole)

| | c. | *li* | *finn* | *al* | *zet* | *so* | *lekor/*li/*li- mem* | *dan* | *larivier* |
|---|---|---|---|---|---|---|---|---|---|
| | | 3SG | PFV | go | throw | 3SG.POSS | body/3SG/3SG-self | in | river |
| | | 'He's gone to drown himself' (Syea, 2017: 118) | | | | | | | |

(Mauritian Creole)

| | d. | *pa* | *kas* | *u* | *latet/*u/*u-mem* | *ar* | *sa!* |
|---|---|---|---|---|---|---|---|
| | | NEG | break | 2SG | head/2SG/2SG-self | with | that |
| | | 'Don't trouble yourself with that!' (Syea, 2017: 118) | | | | | |

(Mauritian Creole)

Furthermore, Corne (1988: 75) notes that possessive + *lekor* can occur in a handful of fixed expressions like *tuy so lekor* 'to work hard' and *kasyet so lekor* 'to be work-shy'. This can also include other body parts as, for example, the heart in *dir daṅ so leker* 'to say to oneself/in one's heart'. Moreover, it is important to mention that in the specific cases where BPTs are used reflexively, no other reflexivization strategy is possible. In (22c–d) above, no alternative *li/li-mem* or *u/u-mem* can be used. It is also noteworthy that there may arise different nuances in meaning with certain verbs depending on

whether the plain pronominal or the pronominal-*mem* combination is used (Syea, 2017: 118–119).

Another interesting issue is that combinations of *lekor* and *li-mem* seem to be marginally possible, see (23). Here the question arises whether these are complex reflexives, i.e., whether *lekor* as well as the *li-mem* form constitute the reflexive material or whether one is the reflexive and the other form is a mere emphatic pronoun. For a more detailed discussion see Heine (2005).

| (23) | *i* | *ti* | *êvit* | *sô* | *lekor* | *li-mem* |
|---|---|---|---|---|---|---|
| | 3SG | TAM | invite | his | body | him-self |
| | ‘He invited himself’ (Papen, 1978: 398; *apud* Heine, 2005: 211; glosses adapted) | | | | | |

(Seychelles Creole)

A look at the diachronic development of reflexive marking reveals relevant insights: 18th-century Mauritian Creole also has the possibility of marking reflexivity with BPTs; alternatively and for most predicates whose French counterparts take *se*, the plain verb is used. Corne (1988: 80) reports on fossil forms that include French *se* as in *s'en mélé* ‘to get involved in something’, *s'en vanté* ‘to boast’.

The BPR structure, i.e., possessive + *lekor*, is found with verbs that clearly express a physical movement. Corne (1988: 83) indicates that the BPR structure may have developed out of sentences where an actual positioning of the body was expressed, thus contexts that are very similar to the medieval examples in Section 3.2.2. The same author points to the possibility of French *corps* being a main source for paraphrases of that sort (Corne, 1988: 90).[12] Consider the following examples, which constitute important early examples that are similar to reflexive expressions (24a–c):

| (24) | a. | *pas* | *gâter* | *menage* | *de-s* | *aute-s,* | *ni* | *note* | *corps* |
|---|---|---|---|---|---|---|---|---|---|
| | | NEG | spoil.INF | household | of-PL | other-PL | nor | 1SG.POSS | body |
| | | ‘Not to spoil the households of others, nor our own bodies’ (Caulier, 1770; *apud* Bollée, 2007a: 72) | | | | | | | |

12 Corne (1989) as well as Carden (1993) argue, however, that BPRs in Mauritian Creole may be traced back to reflexive uses of *tena* ‘body, trunk (of a person or tree)’ in Malagasy.

(Réunion Creole French)

| | b. | *vous* | *battez* | *mon* | *corps* |
|---|---|---|---|---|---|
| | | 2.PL | beat.2PL | 2.PL.POSS | body |
| | | 'You beat me' (1784, cited in Chaudenson, 1981: 78 *apud* Corne, 1988: 83) | | | |

(Mauritian Creole)

| | c. | *sipa* | *docteir* | *va* | *soulaze* | *mo* | *lécorps* |
|---|---|---|---|---|---|---|---|
| | | perhaps | doctor | go.3SG | ease.INF | 1.SG.POSS | body |
| | | 'Perhaps the doctor will relieve my pain' (Baissac, 1967 [1888]: 323, *apud* Corne, 1988: 83) | | | | | |

(Mauritian Creole)

While early Mauritian Creole does not seem to use a post-verbal pronoun regularly, 19th-century Mauritian shows evidence of the post-verbal pronoun *twa*/*zot* (from affirmative imperatives in French) more often and *mem* may be added to the verb as well. Bare verbs are also an option. Corne (1988: 85) points out the possibility of the post-verbal pronoun and the use of *mem* being characteristic of registers that seem more heavily influenced by French. He advocates the view that 18th–19th century French served as a model for introducing these pronouns, i.e., post-verbal *twa*/*zot* and *mem*. The possessive + *lekor* structure seems to be the only construction that is available from the earliest times onwards, as it is first attested as an anaphor in 1784. For a detailed discussion of the data see Corne (1988: 79–87).

It is noteworthy that, since in Mauritian Creole possessive + *lekor* is only possible with verbs of physical movement and in combination with some verbs that express emotions, the structure does not seem to be fully grammaticalized. As opposed to this, possessive + *lekor* in Seychelles Creole may also co-occur with other verb classes such as verbs of saying, like *esplik*(*e*) 'to explain' (Corne, 1988: 88; Kriegel, 2005: 74).

### 3.3.2 Portuguese-lexifier PCM languages

Only some of the present-day Portuguese-lexifier contact languages under study here have BPRs, either as the only or as an alternative reflexive strategy. This is the case of

Malacca Creole, where BPRs mostly occur with extroverted verbs such as *matá* ‘to kill’ and *pinchá* ‘to commit suicide’ (lit. to throw away) (25a–b):

| (25) | a. | *teng* | *ńgua* | *omi* | *ki* | *ja* | *matá* | *korpu* | *na* | *Muar* |
|---|---|---|---|---|---|---|---|---|---|---|
| | | be | one | man | REL | PFV | kill | body | LOC | Muar |
| | | ‘There was a man who killed himself in Muar’ (Baxter, 1988: 206) | | | | | | | | |

(Malacca Creole)

| | b. | *eli* | *ja* | *pinchá* | *korpu* |
|---|---|---|---|---|---|
| | | 3SG | PFV | throw.away | body |
| | | ‘(S)he commited suicide’ (Michaelis et al., 2013) | | | |

(Malacca Creole)

In Batavia (Tugu) and Bidau (East Timor) Creole, on the other hand, BPRs occur with verbs that could be rather ascribed to the domain of middle voice, that is to say, grooming verbs such as *lava*, *laba* ‘to wash, to bathe’ and motion predicates like *pasa bira bira* ‘to stroll about’ (26a–c):

| (26) | a. | *lava* | *korpu* |
|---|---|---|---|
| | | wash | body |
| | | ‘To wash (oneself)’ (Maurer, 2011: 98) | |

(Batavia (Tugu) Creole)

| | b. | *eu* | *bai* | *‘laba* | *eu* | *sa* | *‘korpu* |
|---|---|---|---|---|---|---|---|
| | | 1SG | go | wash | 1SG | GEN | body |
| | | ‘I am going to bathe’ (Baxter, 1990: 12) | | | | | |

(Bidau (East Timor) Creole)

| | c. | *pasa* | *bira* | *bira* | *korpu* |
|---|---|---|---|---|---|
| | | go | turn | turn | body |
| | | ‘To stroll about’ (Maurer, 2011: 98) | | | |

(Batavia (Tugu) Creole)

As illustrated by examples (25a–b) and (26a–c), BPRs in Malacca, Bidau and Batavia Creole tend to occur as direct objects and not to bear a possessive pronoun (though see 26b). This suggests that the grammaticalization of these constructions is more advanced than in the Romance superstrates and than in the French-lexifier contact varieties discussed so far. Stated differently, reflexive markers have in this case reached Stage 2 of the semantic cline and are probably in transition between Stages 2 and 3 of the formal dimension of the grammaticalization path laid out in Section 2.2. At any rate, in all three languages alternative reflexive strategies are available, some of which combine BPRs with distinct markers such as emphatic pronouns (27a–b). As opposed to this, other strategies make do without BPRs altogether (27c):

| (27) | a. | *Maria* | *ja* | *olá* | *onsong* | *sa* | *rosto* | *na* | *spelu* |
|---|---|---|---|---|---|---|---|---|---|
| | | Maria | PFV | see | EMPH | 3.POSS | face | in | mirror |
| | | 'Maria saw herself in the mirror (lit. Maria saw self's face in the mirror)' (Baxter and de Silva, 2004: 67) | | | | | | | |

(Malacca Creole)

| | b. | *eli* | *ja* | *matá* | *onsong* | *sa* | *korpu* |
|---|---|---|---|---|---|---|---|
| | | 3SG | PFV | kill | EMPH | GEN | body |
| | | 'He killed himself (lit. he killed self's body)' (Baxter and de Silva, 2004: 67) | | | | | |

(Malacca Creole)

| | c. | *engena* | *sua* | *mesmu* |
|---|---|---|---|---|
| | | betray | GEN | self |
| | | 'To betray oneself' (Maurer, 2011: 98) | | |

(Batavia (Tugu) Creole)

By contrast, BPRs are absent from present-day grammatical descriptions of Indian Creole Portuguese varieties. These include Korlai and Sri Lanka Creole, where a noun meaning 'self' (28a) and verbal suffixes (28b) next to oblique personal pronouns (28c) serve as reflexive markers, respectively:

| (28) | a. | *Lwidz* | *ki* | *sota* | *mato* |
|---|---|---|---|---|---|
| | | Lwidz | OBJ | self | kill.3SG.PST |
| | | ‘Lwidz killed himself’ (Clements, 2007: 169) | | | |

(Korlai Creole)

| | b. | *eli* | *jaa-cucaa-taam* | *faaka* | *vɔɔnda* |
|---|---|---|---|---|---|
| | | 3SG | PST-stab-REFL | knife | by |
| | | ‘He stabbed himself with a knife’ (Michaelis et al., 2013) | | | |

(Sri Lanka Creole)

| | c. | *pasa* | *aanu* | *eev* | *permi* | *jaa-kuziɲaa* |
|---|---|---|---|---|---|---|
| | | past | year | 1SG | 1SG.DAT | PST-cook |
| | | ‘Last year I cooked for myself’ (Michaelis et al., 2013) | | | | |

(Sri Lanka Creole)

There are, however, good reasons to believe that the word *corp*(*o*) ‘body’ had anaphoric and possibly also reflexive uses in earlier Indian Portuguese PCM languages. Such uses are found in late-19th and early-20th century texts across varieties, including Mangalore Creole (29a), Sri Lanka Creole (29b), Diu Creole (29c), Kochi Creole (29d) and Kannur Creole (29e):

| (29) | a. | *deixá* | *entra* | *dentro* | *corpo* | *delicad* |
|---|---|---|---|---|---|---|
| | | let.2SG.IMP | enter.INF | inside | body | delicate |
| | | ‘Let me (lit. this delicate body) come inside’ (Schuchardt, 1883b: 886) | | | | |

(Mangalore Creole)

| | b. | *assie* | *alme* | *corpoe* | *botta* | *per* | *re* | *ne* | *chaan* |
|---|---|---|---|---|---|---|---|---|---|
| | | thus | soul | body | throw.PST | by | king | to | ground |
| | | ‘Thus soul and body threw the king on the ground’ (Anonymous, 1865; Jackson, 1990: 253) | | | | | | | |

(Sri Lanka Creole)

| | c. | *muito* | *sust* | *tomá* | *meu* | *corp* | *na* | *viaz* |
|---|---|---|---|---|---|---|---|---|

| | | great | fright | take.PST | 1SG.POSS | body | on | voyage |
|---|---|---|---|---|---|---|---|---|
| | | ‘I was greatly frightened during the voyage (lit. my body took great fright)’ (Schuchardt, 1883a: 9) | | | | | | |

(Diu Creole)

| | d. | *eu* | *envistio* | *meu* | *corpo* | *com* | *bunito* | *vestimento-s* |
|---|---|---|---|---|---|---|---|---|
| | | 1SG | dress.1SG.PST | 1SG.POSS | body | with | beautiful | clothes-PL |
| | | ‘I dressed my body/myself with beautiful clothes’ (Schuchardt, 1882b: 808) | | | | | | |

(Kochi Creole)

| | e. | *escusado* | *cançar,* | *Maquita,* | *este* | *vosso* | *corpo* |
|---|---|---|---|---|---|---|---|
| | | needless | tire.INF | Maquita | this | 2SG.POSS | body |
| | | ‘(It is) needless to tire yourself, Maquita (lit. to tire this body of yours)’ (Schuchardt, 1889: 522) | | | | | |

(Kannur Creole)

These anaphoric and reflexive-like uses of BPTs —which have arguably shifted to Stage 2 of the semantic cline of grammaticalization— involve predicates that belong to the middle domain, including movement verbs like *entra* ‘to enter’ (29a) and *botta* ‘to throw’ (29b) as well as predicates denoting bodily process such as *tomá sust* ‘to frighten, take fright’ (29c) and *cançar* ‘to tire, get tired’ (29e). BPTs co-occurring with middle predicates are likewise no rarity in 19th and 20th-century Southeast Asian contact varieties, including Bidau (East Timor) Creole (30a) and Batavia (Tugu) Creole (30b):

| (30) | a. | *elli* | *sua* | *corpo* | *tinha* | *assi* | *limpo* |
|---|---|---|---|---|---|---|---|
| | | 3SG.M | 3SG.POSS | body | have.3SG.PST | so | clean |
| | | ‘He was so clean (lit. he had such a clean body)’ (Dalgado, 1900: 122) | | | | | |

(Bidau (East Timor) Creole)

| | b. | *à* | *pöena* | *grande* | *pode* | *non* | *carta* | *un* | *criis* |
|---|---|---|---|---|---|---|---|---|---|
| | | under | punishment | great | may | NEG | carry | a | cross |

| | | *sua* | *corpe* |
|---|---|---|---|
| | | 3PL.POSS | bodies |
| | | ‘Under great penalty, they may not carry a cross with themselves (lit. on/with their bodies)’ (Schuchardt, 1890: 12) | |

(Batavia (Tugu) Creole)

This suggests that in Indian Creole PCM languages, where at present reflexivity is encoded by other means, BPRs used to be more widespread. This change may be accounted for by contact, since the dominant Dravidian languages adjacent to Indo-Portuguese varieties of the area have no reflexive markers derived from BPTs (Andronov, 2003: 169–170). Furthermore, this data situation also indicates that anaphoric and reflexive-like uses of BPTs used to be more common with verbs belonging to the middle domain than with others.

#### 3.3.3 Spanish-lexifier PCM languages

In present-day Philippine Creole varieties BPRs are used with all kinds of predicates, including physical verbs such as *mata* ‘to kill’ and *controlá* ‘to control’, mental-state predicates like *rabyá* ‘to hate’ and verbs belonging to the middle domain such as *introdusí* ‘to introduce’ (31a–d):

| (31) | a. | *ya* | *mata* | *el* | *rey* | *con* | *su* | *mismo* | *cuerpo* |
|---|---|---|---|---|---|---|---|---|---|
| | | PFV | kill | DEF | king | OBJ | 3SG.POSS | same | body |
| | | ‘The king killed himself’ (Michaelis et al., 2013) | | | | | | | |

(Cavite Chavacano)

| | b. | *ta-rabyá* | *‘le* | *su* | *kwérpo* |
|---|---|---|---|---|---|
| | | IPFV-hate | 3SG | 3SG.POSS | body |
| | | ‘(S)he hates herself/himself’ (Michaelis et al., 2013) | | | |

(Zamboanga Chavacano)

| | c. | *ya* | *puede* | *controla* | *disuyo* | *cuerpo* | *si* | *Jose* |
|---|---|---|---|---|---|---|---|---|
| | | PFV | be.able.3SG.PRS | control | 3SG.POSS | body | DEF | Jose |
| | | ‘Jose can control himself’ (Beck, 2001: 18) | | | | | | |

(Zamboanga Chavacano)

| | d. | *keré* | *yo* | *introdusí* | *mi* | *kwérpo* |
|---|---|---|---|---|---|---|
| | | want | 1SG | introduce | 1SG.POSS | body |
| | | 'I want to introduce myself' (Michaelis et al., 2013) | | | | |

(Zamboanga Chavacano)

Since the constructions in (31a–d) are still ambiguous between a literal and a reflexive meaning, one might argue that they represent Stage 2 of the semantic level of grammaticalization. In turn, no reduction in the morphosyntactic behavior of these constructions can be observed other than the fact that the object marker *con* is rarely used with reflexive markers.

Spanish-lexifier PCM languages likewise present alternative reflexive strategies, some of which combine BPRs with distinct markers such as emphatic pronouns (32a). Others do not make use of BPRs (32b):

| (32) | a. | *kwándo* | *yo* | *góra* | *ta-pená,* | *ta-pená* | *mi* |
|---|---|---|---|---|---|---|---|
| | | when | 1SG | now | IPFV-suffer | IPFV-suffer | 1SG.POSS |
| | | *saríli,* | *mi* | *kwérpu* | | | |
| | | self | 1SG.POSS | body | | | |
| | | 'Now, when I was suffering, my self, my body suffered' (Sippola, 2011: 261) | | | | | |

(Ternate Chavacano)

| | b. | *tyeni* | *lastima* | *yo* | *konmigo* |
|---|---|---|---|---|---|
| | | have | pity | 1SG | 1SG.OBJ |
| | | 'I feel bad for myself' (Pérez, 2015: 115) | | | |

(Cavite Chavacano)

There are also good reasons to believe that anaphoric and reflexive uses of the word *cuerpo*, *cuelpo*, *kwérpo* 'body' can be traced back to older stages of Philippine Creole varieties. This is suggested by late-19th and early-20th century attestations of BPRs (33a–b):

| (33) | a. | *mira* | *vos* | *bonito* | *cuerpo* |
|---|---|---|---|---|---|
| | | look.2SG.IMP | 2SG.POSS | beautiful | body |
| | | 'Look at your beautiful self/body' (Tombo, 1860: 258) | | | |

(Zamboanga Chavacano)

| | b. | *este* | *mi* | *cuerpo* | *sin* | *el* | *alma* | *ya* | *quedá,* |
|---|---|---|---|---|---|---|---|---|---|
| | | this | 1SG.POSS | body | without | DEF | soul | PFV | become |
| | | *blando* | *blando,* | *lengue* | *lengue,* | *ni* | *está* | *podé* | *caminá* |
| | | soft | soft | sickly | sickly | nor | stay | be.able | walk |
| | | 'I have lost my soul, I have become very soft and sickly, I cannot even walk (lit. this body of mine has lost its soul, has become very soft and sickly, it cannot even walk)' (Schuchardt, 1884: 150) | | | | | | | |

(Zamboanga Chavacano)

| | c. | *un* | *bida* | *que* | *ya* | *sentí* | *ele* | *palpitá* | *na* | *dentro* |
|---|---|---|---|---|---|---|---|---|---|---|
| | | INDF | life | REL | PFV | feel | 3SG | throb | in | inside |
| | | *de* | *su* | *cuelpo* | | | | | | |
| | | of | 3SG.POSS | body | | | | | | |
| | | 'A life which he felt throbbing inside himself (lit. inside his body)' (Balmori, 1917: 71) | | | | | | | | |

(Cavite Chavacano)

These anaphoric and reflexive-like early uses of BPTs in Philippine Creole varieties involve predicates that belong to the middle domain such as *senti* 'to feel' and *quedá* 'to become', as well as other kinds of verbs like *mirá* 'to look'. This data situation, which is largely identical to the present-day languages, indicates that despite the coexistence of other markers, BPRs have been a major reflexive strategy in Philippine Creole varieties.

To summarize this section, first attestations of BPTs in Romance-lexifier contact varieties often involve the concrete positioning of the body in an action that is related to physical movement. These instances are quite similar to the data that has been found concerning medieval and Renaissance Romance varieties. While in some PCM languages BPTs can occur with different verb classes and can serve as a major

reflexivization strategy, they seldom constitute the most generalized or most grammaticalized reflexive form. Rather, it is mostly the case that BPTs are restricted to predicates indicating physical action or movement as well as to verbs that convey an emotional state. There are almost always other, more recent and more productive reflexivization strategies at hand, which can often be traced back to emphatic pronouns.

In general terms, the fact that reflexive-like uses of BPTs are common with verbs denoting body positioning and physical action in the historical Romance lexifiers as well as the present-day contact varieties suggests a link between both. According to the hypothesis laid out in Sections 1 and 2.2, this indicates that speakers of contact varieties may have identified, recapitulated and replicated reflexive-like uses of BPTs in the three bridging contexts described in Section 3.2. There are, however, also differences, the main one being that BPRs tend to be more grammaticalized in the contact varieties than in the lexifiers: in the former, reflexive constructions display slightly constrained morphosyntactic behavior such as the possibility to omit possessive pronouns. This is not the case in the lexifiers. Furthermore, anaphoric uses of BPTs are commonly attested in the lexifiers, whereas they are all but lost in the current contact varieties. These differences as well as the historical and sociolinguistic contact setting are discussed in Section 4.

# 4 Replication and Grammaticalization of BPRs in Romance-lexifier PCM languages: Textual, Historic and Sociolinguistic Factors

The relevance of the Bible and other religious texts in the formation process of the pidgins and creoles under study has been mentioned at various points in Section 3. The idea behind this view rests on two assumptions: (i) that religious texts were used to evangelize non-native speakers of Romance languages in the colonies during the 17th, 18th and 19th centuries, and (ii) that these texts were written in a particular style that favors metonymy of BPTs, a mechanism which plays an important role in the grammaticalization of reflexive markers. The last point is also supported by the fact that many translations of the Bible into major European languages stem from the Vulgate, itself a translation into LML (Nunn, 1922). Let it be remembered at this point that in LML metonymic uses of the BPTs *caput* 'head' and *corpus* 'body' are particularly

prevalent (cf. Section 3.2.1), which may have enabled their transfer (or inheritance) into the Romance languages.

The first of these assumptions seems quite uncontroversial: the relevance of religion in the spread and use of colonial languages is directly or indirectly addressed in studies on Indian Ocean Creole French (Bollée, 2007a), Indian Creole Portuguese (Jackson, 1990) and Philippine Creole Spanish (Fernández, 2010). Jackson (1990: 13) points out that religion played a very important role in the life of colonial Portuguese cities in 17th-century Sri Lanka, and the Philippine scholar Pardo de Tavera, in a 19th-century letter to Hugo Schuchardt, even states that speakers of "kitchen Spanish" barely have anything else to talk about beyond language and religion (Fernández, 2010: 249). Furthermore, many 18th- and 19th-century PCM-language texts from the area are religious in nature (Schuchardt, 1883a; Bollée, 2007a).

The second assumption is more difficult to substantiate. However, examples provided in Section 3 illustrate that metonymic uses of BPTs are indeed widespread in religious (as well as more secular) Romance texts from the late Middle Ages, the Renaissance and subsequent centuries as well as in the older documented PCM-language texts. There is thus no reason not to believe that the particular writing style of religious texts and the repeated exposure to this textual material in the course of evangelization influenced the emergence of PCM languages throughout the Indian-Pacific Ocean areas.

The role of nonstandard and dialectal Romance varieties in the creation of BPRs in Romance-lexifier PCM languages has also been highlighted previously. As in the former point, the evidence is most clear for French-lexifier pidgins and creoles. More specifically, Kriegel et al. (2019: 332–336) suggest a link between BPRs in western varieties of 17th-to-19th-century French such as Berrichon and Poitevin-Saintongeais, on the one hand, and in French-lexifier contact varieties, on the other. Specific examples of the former are reminiscent of BPRs discussed in Section 3 (34a–b):

| (34) | a. | *moué* | *que* | *le* | *corps* | *m='en* | *tremble* |
|---|---|---|---|---|---|---|---|
| | | 1SG | SUB | DEF | body | 1SG=3SG | shiver.3SG.PRS |
| | | '(Oh) my, I am shivering (lit. my body is shivering on me)' (*Couplets de la Fête des Rois*, De la Salle, 1875: 25; *apud* Kriegel et al., 2019: | | | | | |

| | | 305) |
|---|---|---|

(Berrichon French)

| | b. | *avoir* | *a=u* | *darriée* | *de* | *son* | *corps* |
|---|---|---|---|---|---|---|---|
| | | have.INF | on=DEF | back | of | 3SG.POSS | body |
| | | 'To have something behind oneself (lit. behind one's own body)' (Jaubert, 1869: 180; *apud* Kriegel et al., 2019: 333) | | | | | |

(Berrichon French)

As illustrated by (34b), the link between non-standard Romance varieties and Romance-lexifier PCM languages is particularly evident in proverbs, sayings and fixed expressions, which as mentioned in Section 3.1 tend to preserve linguistic archaisms. This suggests that BPRs might have been replicated by the individual contact varieties early on.[13] The importance of proverbs, sayings and fixed expressions for the topic at hand has also been made evident on various occasions in Section 3.

This leads to the question of how the emergence of BPRs in Romance-lexifier PCM languages may have ensued from a sociolinguistic perspective. In this respect Guy (1990: 48) classifies three major sociolinguistic types of change: (i) 'spontaneous' change, which ensues within a single speech community without an external linguistic model, (ii) change 'from above', which is induced by native speakers drawing on an external pattern, and (iii) 'imposition', where it is non-native speakers who are the agents of change:

TABLE 1 *The major sociolinguistic types of language change* (Guy, 1990: 48)

In cases of change 'from above', native speakers import into their language features from another language (Guy, 1990: 49). This fits well with the proposal at hand: native speakers of early colonial creoles might have identified and recapitulated metonymic, anaphoric and reflexive-like uses of BPTs in the relevant bridging contexts extant in the Romance lexifiers, thus creating a basis for replicating reflexive markers. This situation

13 If it were assumed, as the data under discussion here suggest, that at least in some cases Romance-lexifier PCM languages of the Indian and Pacific Ocean areas replicated BPRs from substandard varieties of their lexifiers, then this would be in line with the view that creoles do not develop from pidgins, but from basilectal varieties approximating their lexifiers (Chaudenson, 1992; Mufwene, 2008). However this may be, we steer away from that discussion, since it is not central to the arguments laid out here.

gives rise to an opposition: as shown in Section 3, there are many anaphoric but few uncontroversial reflexive uses of BPTs in the lexifier languages, whereas anaphoric and reflexive functions are abundant in Romance-based contact varieties to varying degrees. Also, BPTs display more constrained morphosyntactic behavior in the pidgins and creoles than in the source languages. This implies that reflexive constructions are more grammaticalized in the former than in the latter.

There could be many reasons to explain this contrast, one of which may relate to ecology. Specifically, most lexifier languages were spoken in complex, heterogeneous societies by comparatively large numbers of speakers during the 16th, 17th and 18th centuries: Grégoire (1794, *apud* Lodge, 1993: 198–200), for example, estimates the population of France to have reached 26 million between 1790 and 1794, slightly above half of which might have been fluent in French, as opposed to the other half, which was probably more proficient in closely related *Langue d'Oïl* varieties, or other languages such as Breton, Occitan and Alsatian.

In turn, the PCM language-speaking communities in question have generally been small and relatively homogeneous, albeit often as parts of much larger, linguistically diverse societies. Malacca Creole, for example, is currently the language of a small Christian ethnic minority of approximately 1000 speakers (out of nearly half a million inhabitants in Malacca), and this state of affairs cannot have changed significantly since the time of colonization (Baxter, 1988: 5–6). The situation of Sri Lanka Creole is similar (Jackson, 1990: 3–4), and much the same is true of many contact varieties discussed so far such as Diu Creole (Cardoso, 2009: 13–15), Ternate Chavacano (Sippola, 2011: 12) and Cotabato Chavacano (Fernández, 2012: 300–301). Other contact varieties, particularly Mauritian Creole (Seuren, 1995: 531–533) and Zamboanga Chavacano (Lorenzino, 2001: 11–12), are the exception rather than the rule in that they have been widely used in the communities in which they coexist with other languages.

In sum, most of the creole-speaking communities have been small and relatively tight-knit since the time of colonization. As pointed out by Ansaldo (2009: 111), community size largely determines the rate at which languages change. To the extent that replication of reflexive-like uses of BPTs by native speakers of Romance-lexifier PCM languages must have been an innovation, diffusion of this innovation throughout

the small creole-speaking communities must have ensued at a fast rate. In comparison, the propagation of anaphoric and reflexive-like BPTs in the larger societies in which the colonial languages were spoken must have been slower, which explains the difference concerning the degree of grammaticalization of reflexive markers between the lexifiers and the contact varieties.

Additionally, the divergent degree of grammaticalization in contact varieties and in their lexifiers may be traced back to processes of standardization, which begin in the lexifier languages during the 17th–19th centuries, whereas such processes have been implemented, if at all, much later in most pidgins and creoles. Indeed, standardization processes have been argued to slow down or halt cases of language change or, alternatively, to favor conservative patterns at the expense of innovative ones. Laitinen (2004: 253–259), for example, suggests that grammaticalization of the Finnish non-agreeing negative verb *ei* into a negative particle was halted in 19th-century written language by the gradual reintroduction of agreement with the subject; agreement became generalized by the mid-20th century.

In general terms, standardized varieties are considered to change at a lower rate than non-standardized varieties (Deumert, 2004: 7). This restraining effect obtains because standardization involves the reduction of dialect differences, which ensues via two mechanisms: dialect leveling or the avoidance of salient features of particular dialects, on the one hand, and simplification, i.e., the reduction in inventory and regularization of alternations occurring in one or more varieties (Ferguson, 1997 [1987]: 70), on the other hand. As a consequence of these two unifying mechanisms, structures of dialects that have become salient or given rise to complex alternations may be 'retracted' or 'withdrawn' by the speaker community in order to comply with inconspicuous or less complex features of the variety that serves as the model for standardization.

This seems to have occurred with anaphoric and reflexive-like uses of BPTs in 18th-19th-century Renaissance varieties. Normative tendencies that arose within the standardization of French and the creation of the *bon usage* are deeply intertwined with the establishment of supra-regional norms favoring a certain language use on the one hand (usually exemplified by the use at court and by most praised writers of the time) and the regression of regional features and the language use of the people in more rural settings on the other hand (Lodge, 1993: 85, 176ss.). Within these processes, French

shows one of the richest bodies of normative works and most intensive periods of codification in Europe with the expressed goal to make French as eloquent as Latin or Italian in all areas (Lodge, 1993: 157ss). Especially in the 17th and 18th centuries, French should convey logic and reason with uttermost precision and efficiency (Lodge, 1993: 177ss).

These tendencies seem to discourage anaphoric and reflexive-like uses of BPTs, which have never been the primary source of reflexivization. The much more grammaticalized and generalized reflexive strategy was with forms of Romance *se* and, while BPTs were widely used in the Middle Ages, they are later restricted to a more regional use, as was shown in Section 3. More recently, Kriegel et al. (2019: 335) explicitly connect the retraction with the standardization processes that gain momentum in the 17th century and also show that the use of BPTs seems to live on in dialects like Berrichon and Saintongeais. They also hint at the fact mentioned above that the French language should express logic and reason with great precision as a contributive factor for the more expressive variant with *cors* 'body' to be relegated.

> Cette expression du réfléchi – qui n'est, bien entendu, qu'une alternative au pronom réfléchi simple par rapport auquel elle apporte toujours, malgré ce début de grammaticalisation, une composante sémantique expressive – se perd dans le processus d'élaboration du « bon usage ». Et, alors que les concepts physiques deviennent plus abstraits dans l'évolution du français, ce sont les dialectes qui semblent hériter de cette attention accrue aux parties du corps. [335] [...] [Cettes structures] ont un statut accessoire, à côté de l'encodage pronominal traditionnel du réfléchi, et c'est sous cette forme que la structure réfléchie avec *corps* a fait partie des registres du français oral transmis aux Antilles par les colons (Kriegel et al., 2019: 336)

This seems to show that although BPTs have never been as generalized as Romance *se*, there is nonetheless a continuous thread of use attested in regional varieties, which has not survived in today's standard French due to normative effects.

## 5 Conclusions

The cross-linguistic analysis of BPRs has shown that there seems to be a prevalence of these structures in pidgins, creoles and mixed languages that are connected to Indo-European lexifiers. The view that reflexives based on body-part terms should be traced

back to a West African substrate, as has often been suggested for contact varieties of the Atlantic area, is countered by the fact that a handful of contact languages including Réunion, Batavia, Bidau and Malacca Creole, Cavite, Ternate and Zamboanga Chavacano display body-part reflexives, yet they have had little or no demonstrable influence from African languages throughout their histories, they are not spoken in the Atlantic area as most other Romance-lexifier contact languages with BPRs, and they are in many cases in contact with languages that have no such reflexive constructions.

Accordingly, and drawing on previous studies on the topic, it has been argued that BPRs in the PCM languages under analysis originate in anaphoric and reflexive-like uses of body-part terms in the lexifiers. Specifically, the data show that such uses represented an alternative, yet never primary reflexive strategy in medieval, Renaissance and later Romance languages. Anaphoric and reflexive-like uses of BPTs were particularly prevalent in three contexts: (a) those involving an opposition between *body* and *soul*, (b) those where the use of *body* is involved in battle situations, and (c) those where *body* co-occurs with actions that typically fall into the realm of the middle domain. These contexts might have functioned as bridges for BPRs to be replicated into the contact varieties; the affinity between the lexifiers and the contact varieties concerning contexts (a)–(c) is illustrated by the fact that such contexts are likewise found in most PCM languages, particularly in the oldest texts, cf. Tables 2–5.

TABLE 2 *Contexts of semantic spread of BPTs in Indian Ocean Creole French varieties*

TABLE 3 *Contexts of semantic spread of BPTs in Southeast Asian Creole Portuguese varieties*

TABLE 4 *Contexts of semantic spread of BPTs in Indian Creole Portuguese varieties*

TABLE 5 *Contexts of semantic spread of BPTs in Philippine Creole Spanish varieties*

The co-occurrence of anaphoric and reflexive uses of BPTs in the lexifiers, which represent an early and a later stage of grammaticalization respectively, probably allowed speakers of the contact varieties to identify and recapitulate the steps in the reflexives' grammaticalization cline. This situation then enabled replication of anaphoric and reflexive structures of the superstrates into the contact languages —a change that could be characterized in sociolinguistic terms as carried out 'from above' (in Guy's (1990) terminology)— via the three bridging contexts described earlier. Replication may also have been driven by the more functionally transparent needs of contact situations.

The aforementioned events seem to have given rise to an opposition: anaphoric uses of BPTs are more frequent in the lexifiers, whereas reflexive uses prevail in the contact languages. Furthermore, reflexive constructions display more constrained morphosyntactic behavior in the latter. This implies that grammaticalization has gone further in the PCM languages than in the Romance lexifiers. Two reasons have been adduced for this asymmetry: on the one hand, most of the contact languages under scrutiny are or were spoken in relatively small and tight-knit communities, whereas the lexifier languages were part of larger, rather heterogeneous societies. Consequently, diffusion of an innovative feature such as reflexive uses of BPTs must have occurred more rapidly in societies speaking PCM languages.

On the other hand, it has been argued that normative tendencies, which the lexifiers were subjected to earlier and to a greater extent than the contact languages, contributed to eliminating features from standard varieties which had become salient and given rise to relatively complex oppositions. One such opposition may have involved the reflexive pronoun *se* vs. reflexive uses of BPTs like 'body' and 'head'. In turn, regional varieties less affected by normative efforts such as Berrichon and Poitevin-Saintongeais in the case of French, as shown by Kriegel et al. (2019), preserved reflexive-like uses of BPTs until recently. Moreover, our study has shown that traces of such uses are preserved in conservative contexts like proverbs, sayings and fixed expressions.

**Abbreviations**

1 1st person; 2 2nd person; 3 3rd person; ABL ablative; ACC accusative; AUX auxiliary verb; BPR body-part reflexive; BPT body-part term; DAT dative; DEF definite; CND conditional; EMPH emphatic; F feminine; FUT future; GEN genitive; GER gerund; IMP imperative; INDF indefinite; INF infinitive; IPFV imperfective; LML late and medieval Latin; M masculine; MID middle voice; NEG negative; NOM nominative; OBJ object marker; OBL oblique; PASS passive; PCM pidgins, creoles and mixed languages; PFV perfective; PL plural; POSS possessive; PROG progressive; PRS present; PST past; PTCP participle; REFL reflexive; REL relative; SBJ subjunctive; SG singular; SUB subordinator; TAM tense-aspect-mood marker.

## References


Andronov, Mikhail S. 2003. *A comparative grammar of the Dravidian languages*. Wiesbaden: Otto Harrassowitz.

Ansaldo, Umberto. 2009. *Contact languages: Ecology and evolution in Asia*. Cambridge: Cambridge University Press.

Arana-Ward, Maria. 1977. *A synchronic and diachronic investigation of Macanese: The Portuguese-based creole of Macao*. PhD dissertation, University of Hong Kong.

Baschet, Jérôme. 2000. Âme et corps dans l'occident médiéval: Une dualité dynamique entre pluralité et dualisme. *Archives de Sciences Sociales des Religions* 112: 1–24.

Baxter, Alan N. 1988. *A grammar of Kristang (Malacca Creole Portuguese)*. Canberra: Pacific Linguistics.

Bollée, Annegret. 2007a. *Deux textes religieux de Bourbon du 18e siècle et l'histoire du créole réunionnais*. London: Battlebridge.

Bollée, Annegret. 2007b [1982]. Die Rolle der Konvergenz bei der Kreolisierung. In Per Sture (ed.), *Die Leistung der Strataforschung und der Kreolistik: typologische Aspekte der Sprachkontakte. Akten des 5. Symposiums über Sprachkontakt in Europa*, 391–405. Tübingen: Niemeyer.

Bruyn, Adrienne. 2009. Grammaticalization in creoles: Ordinary and not-so-ordinary cases. *Studies in Language* 33(2): 312–337.

Carden, Guy. 1993. The Mauritian Creole 'lekor' reflexive: Substrate influence on the target-location parameter. In Francis Byrne and John Holm (eds.), *Atlantic meets Pacific: A global view of pidginization and creolization*, 105–117. Amsterdam/Philadelphia: John Benjamins.

Cardoso, Hugo C. 2009. *The Indo-Portuguese language of Diu*. PhD dissertation, Utrecht University.

Cardoso, Hugo. 2020. Contact and Portuguese-lexified creoles. In Raymond Hickey (ed.), *The handbook of language contact* (2nd edition), 469–488. Hoboken: Wiley-Blackwell.

Cennamo, Michela. 1993. *The reanalyisis of reflexives: A diachronic perspective*. Naples: Liguori.

Cennamo, Michela. 2009. Argument structure and alignment variations and changes in

Late Latin. In Jóhanna Barðdal and Shobhana L. Chelliah (eds.), *The role of semantic, pragmatic and discourse factors in the development of case*, 307–346. Amsterdam/Philadelphia: John Benjamins.

Chaudenson, Robert. 1973. Pour une étude comparée des créoles et parlers français d'outre-mer: Survivance et innovation. *Revue de Linguistique Romane* 147-148: 342–371.

Chaudenson, Robert. 1992. *Des îles, des hommes, des langues: Essais sur la créolisation linguistique et culturelle*. Paris: L'Harmattan.

Corne, Chris. 1988. Mauritian Creole reflexives. *Journal of Pidgin and Creole Languages* 3(1): 69–94.

Corne, Chris. 1989. On French influence in the development of creole reflexive patterns. *Journal of Pidgin and Creole Languages* 4(1): 103–115.

Dayley, Jon P. 1981. *Tzutujil grammar*. PhD dissertation, University of California.

Detges, Ulrich and Richard Waltereit. 2002. Grammaticalization vs. reanalysis: A semantic-pragmatic account of functional change in grammar. *Zeitschrift für Sprachwissenschaft* 21(2): 151–195.

Deumert, Ana. 2004. *Language standardization and language change: The dynamics of Cape Dutch*. Amsterdam/Philadelphia : John Benjamins.

Ernout, Alfred and Antoine Meillet. 2001. *Dictionnaire étymologique de la langue latine: Histoire des mots*, 4th edition. Paris: Klincksieck.

Evans, Nicholas and David Wilkins. 1985. *The knowing ear: An Australian test of universal claims about the semantic structure of sensory verbs and their extension into the domain of cognition*. Cologne: Universität zu Köln.

Evseeva, Natalia and Iker Salaberri. 2018. Grammaticalization of nouns meaning "head" into reflexive markers: A cross-linguistic study. *Linguistic Typology* 22(3): 385–435.

Faltz, Leonard M. 1977. *Reflexivization: A study in universal syntax*. PhD dissertation, University of California.

Ferguson, Charles A. 1997 [1987]. Standardization as a form of language spread. In Kirk Belnap and Niloofar Haeri (eds.), *Structuralist studies in Arabic linguistics: Charles A. Ferguson's Papers, 1954–1994*, 69–80. Leiden: Brill.

Fernández, Mauro. 2010. Las cartas de Pardo de Tavera a Schuchardt sobre el "español

de cocina” de las Islas Filipinas. *Grazer Linguistische Studien* 74: 239–272.

Fernández, Mauro. 2012. Leyenda e historia del chabacano de Ermita (Manila). *UniverSOS: Revista de Lenguas Indígenas y Universos Culturales* 9: 9–70.

Geeraerts, Dirk. 1985. Cognitive restrictions on the structure of semantic change. In Jacek Fisiak (ed.), *Historical semantics - historical word formation*, 127–154. Berlin/New York: De Gruyter.

Guy, Gregory G. 1990. The sociolinguistic types of change. *Diachronica* 7(1): 47–67.

Hagège, Claude. 1993. *The language builder: An essay on human signature in linguistic morphogenesis*. Amsterdam/Philadelphia: John Benjamins.

Haiman, John. 1983. Iconic and economic motivation. *Language* 59(4): 781–819.

Harris, Alice C. and Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.

Haspelmath, Martin. 1993. More on the typology of inchoative/causative verb alternations. In Bernard Comrie and Maria Polinsky (eds.), *Causatives and transitivity*, 87–120. Amsterdam/Philadelphia: John Benjamins.

Haspelmath, Martin. Forthcoming. Comparing reflexive constructions in the world’s languages. In Katarzyna Janic, Nicoletta Puddu and Martin Haspelmath (eds.), *Reflexive constructions in the world’s languages*. Berlin: Language Science Press.

Heine, Bernd. 1999. Polysemy involving reflexive and reciprocal markers in African languages. In Zygmunt Frajzyngier and Traci S. Curl (eds.), *Reciprocals: Forms and functions*, 1–29. Amsterdam/Philadelphia: John Benjamins.

Heine, Bernd. 2005. On reflexive forms in creoles. *Lingua* (115), 201–257.

Heine, Bernd and Mechthild Reh. 1984. *Grammaticalization and reanalysis in African languages*. Hamburg: Helmut Buske.

Heine, Bernd and Tania Kuteva. 2003. On contact-induced grammaticalization. *Studies in Language* 27(3): 529–572.

Heine, Bernd and Tania Kuteva. 2005. *Language contact and grammatical change*. Cambridge: Cambridge University Press.

Hopper, Paul J. 1991. On some principles of grammaticization. In Elizabeth C. Traugott and Bernd Heine (eds.), *Approaches to grammaticalization*. vol. I: *Focus on theoretical and methodological issues*, 17–35. Amsterdam/Philadelphia: John Benjamins.

Hopper, Paul J. and Elizabeth C. Traugott. 2003. *Grammaticalization*. Cambridge: Cambridge University Press.

Jackson, Kenneth D. 1990. *Sing without shame: Oral traditions in Indo-Portuguese creole verse: With transcription and analysis of a nineteenth-century manuscript of Ceylon Portuguese*. Amsterdam/Phildadelphia: John Benjamins, Instituto Cultural de Macau.

Jensen, Frede. 1986. *Syntaxe de l'ancien occitan*. Tübingen: Niemeyer.

Jensen, Frede. 1990. *Old French and comparative Gallo-Romance syntax*. Tübingen: Niemeyer.

Kemmer, Suzanne. 1993. *The middle voice*. Amsterdam/Philadelphia: John Benjamins.

König, Ekkehard and Peter Siemund. 2000. Intensifiers and reflexives: A typological perspective. In Zygmunt Frajzyngier and Traci S. Curl (eds.), *Reflexives: Forms and functions,* 41–74. Amsterdam/Philadelphia: John Benjamins.

Kriegel, Sybille. 2005. Quelques évolutions syntaxiques en créole seychellois. *La Linguistique* 41(1): 67–82.

Kriegel, Sybille, Ralph Ludwig and Stefan Pfänder. 2019. Dialectes – créolisation – convergence: Quelques hypothèses à partir du berrichon et du poitevin-saintongeais. In Andreas Dufter, Klaus Grübl and Thomas Scharinger (eds.), *Des parlers d'oïl à la francophonie: Contact, variation et changement linguistiques*, 299–348. Berlin/New York: De Gruyter.

Kuteva, Tania. 2017. Contact and borrowing. In Adam Ledgeway and Ian Roberts (eds.), *The Cambridge handbook of historical syntax*, 163–186. Cambridge: Cambridge University Press.

Kuteva, Tania, Bernd Heine, Bo Hong, Haiping Long, Heiko Narrog and Seongha Rhee. 2019. *World lexicon of grammaticalization*. 2nd edition. Cambridge: Cambridge University Press.

Laitinen, Lea. 2004. Grammaticalization and standardization. In Olga Fischer, Muriel Norde and Harry Perridon (eds.), *Up and down the cline – The nature of grammaticalization*, 247–262. Amsterdam/Philadelphia: John Benjamins.

Lefebvre, Claire. 1998. *Creole genesis and the acquisition of grammar: The case of Haitian creole*. Cambridge: Cambridge University Press.

Lipski, John M. 1987. Modern Spanish once-removed in Philippine Creole Spanish: The

case of Zamboangueño. *Language in Society* 16, 91–108.

Lodge, Raymond A. 1993. *French: From dialect to standard*. London/New York: Routledge.

Lorenzino, Gerardo A. 2001. *The morphosyntax of Spanish-lexified creoles*. Munich: Lincom Europa.

Matras, Yaron and Jeanette Sakel. 2007. Investigating the mechanisms of pattern replication in language convergence. *Studies in Language* 31(4): 829–865.

McWhorter, John H. 2002. The rest of the story: Restoring pidginization to creole genesis theory. *Journal of Pidgin and Creole Languages* 17(1): 1–48.

Michaelis, Susanne M. 2020. Avoiding bias in comparative creole studies: Stratification by lexifier and substrate. *Isogloss* 6: 1–35.

Michaelis, Susanne M. and Martin Haspelmath. 2020. Grammaticalization in creole languages: Accelerated functionalization and semantic imitation. In Walter Bisang and Andrej Malchukov (eds.), *Grammaticalization scenarios: Areal patterns and cross-linguistic variation. A comparative handbook*, 1109–1128. Berlin: De Gruyter.

Michaelis, Susanne M., Philippe Maurer, Martin Haspelmath and Magnus Huber (eds.). 2013. *Atlas of pidgin and creole language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://apics-online.info (accessed on 29/01/2020)

Mufwene, Salikoko S. 2008. *Language evolution: Contact, competition and change*. Wiltshire: Cromwell Press.

Muysken, Pieter. 1993. Reflexes of Ibero-Romance reflexive clitic + verb combinations in Papiamentu: Thematic grids and grammatical relations. In Francis Byrne and Donald Winford (eds.), *Focus and grammatical relations in creole languages: Papers from the University of Chicago conference on focus and grammatical relations in creole languages*, 285–301. Amsterdam/Philadelphia: John Benjamins.

Muysken, Pieter and Norval Smith. 1994. Reflexives. In Jacques Arends, Pieter Muysken and Norval Smith (eds.), *Pidgins and creoles: An introduction*, 271–288. Amsterdam/Philadelphia: John Benjamins.

Naro, Anthony J. 1978. A study on the origins of pidginization. *Language* 54(2): 314–347.

Nunn, Henry P. V. 1922. *An introduction to Ecclesiastical Latin*. Cambridge: Cambridge

University Press.

Pietsch, Lukas. 2009. Hiberno-English medial-object perfects reconsidered: A case of contact-induced grammaticalisation. *Studies in Language* 33(3): 528–568.

Police-Michel, Daniella, Arnaud Carpooran and Guilhem Florigny. 2011. *Gramer kreol morisien*. vol. I: *Dokiman referans*. Port Louis: Akademi Kreol Morisien, Ministry of Education and Human Resources.

Schachter, Paul and Fe T. Otanes. 1972. *Tagalog reference grammar*. Berkeley/Los Angeles: University of California Press.

Schladt, Mathias. 2000. The typology and grammaticalization of reflexives. In Zygmunt Frajzyngier and Traci S. Curl (eds.), *Reflexives: Forms and functions*, 103–124. Amsterdam/Philadelphia: John Benjamins.

Schuchardt, Hugo. 1883a. Kreolische Studien III: Über das Indoportugiesische von Diu. *Sitzungsberichte der kaiserlichen Akademie der Wissenschaften* 103: 3–18.

Seuren, Pieter A. 1995. Notes on the history and the syntax of Mauritian Creole. *Linguistics* 33(3): 531–577.

Sippola, Eeva. 2011. *Una gramática descriptiva del chabacano de Ternate*. PhD dissertation, Helsinki University.

Syea, Anand. 2017. *French creoles: A comprehensive and comparative grammar.* London/New York: Routledge.

Viti, Carlotta. 2009. Anaphorische und reflexive Strukturen im Altgriechischen und im Lateinischen. *Indogermanische Forschungen* 114(1): 143–172.

Weiss, Michael. 2009. *Outline of the historical and comparative grammar of Latin*. Ann Arbor: Beech Stave Press.

Wolfsgruber, Anne C. 2017a. *Medieval Romance* se. *Its grammaticalization and syntactic status. New data from Medieval Spanish, French, Catalan and Occitan*. PhD dissertation, Universität Salzburg.

Wolfsgruber, Anne C. 2017b. *On*, *se* and related valency alternations in Medieval French. *Lingvisticæ Investigationes* 40(1): 59–80.

Wright, Roger. 2017. Late Latin: The evidence. In Alfonso G. Leal and Clara E. Prieto (eds.), *Latin vulgaire – latin tardif XI: Actes du XI Colloque international sur le latin vulgaire et tardif (Oviedo, 1–5 septembre 2014)*, 129–145. Hildesheim: Olms Weidmann.

Ziegeler, Debra. 2000. *Hypothetical modality: Grammaticalisation in an L2 dialect*. Amsterdam/Philadelphia: John Benjamins.

Ziegeler, Debra. 2014. Replica grammaticalisation as recapitulation: The other side of contact. *Diachronica* 31(1): 106–141.

Ziegeler, Debra. 2017. Historical replication in contact grammaticalization. In Daniel Van Olmen, Hubert Cuyckens and Lobke Ghesquière (eds.), *Aspects of grammaticalization: (Inter)subjectification and directionality*, 311–352. Berlin/New York: De Gruyter.

Ziegeler, Debra and Sarah Lee. 2019. Lexical retention in contact grammaticalisation: *Already* in Southeast Asian Englishes. *Journal of Language Contact* 12(3): 737–783.

Zúñiga, Fernando and Seppo Kittilä. 2019. *Grammatical voice*. Cambridge: Cambridge University Press.

## Appendix: Corpus Materials Used

The Appendix can be accessed through the following anonymous link:
https://zenodo.org/record/6363504#.YjI2_jUo-5c

## Tables

TABLE 1 *The major sociolinguistic types of language change* (Guy, 1990: 48)

| | Change types | | |
|---|---|---|---|
| | Internally induced | Externally induced | |
| Alternative terms | Spontaneous, untargeted, natural, 'from below' | Borrowing, targeted, 'from above', recipient language agentivity | Imposition, substratum, source language agentivity |
| Language contact involved? | No | Yes | Yes |
| Agents of change | Native speakers | Native speakers | Non-native speakers |

TABLE 2 *Contexts of semantic spread of BPTs in Indian Ocean Creole French varieties*

| New contexts (NCs) for BPRs | Mauritian CF | Réunion CF | Rodrigues CF | Seychelles CF |
|---|---|---|---|---|
| *body* vs. *soul* dichotomy (NC #1) | - | - | - | - |
| possessive pronoun + | + | + | + | + |

| | | | | |
|---|---|---|---|---|
| *body* in battle situations (NC #2) | | | | |
| possessive pronoun + *body* expressing middle contexts (NC #3) | + | + | + | + |

TABLE 3 *Contexts of semantic spread of BPTs in Southeast Asian Creole Portuguese varieties*

| New contexts (NCs) for BPRs | Batavia (Tugu) CP | Bidau (East Timor) CP | Malacca CP |
|---|---|---|---|
| *body* vs. *soul* dichotomy (NC #1) | - | - | - |
| possessive pronoun + *body* in battle situations (NC #2) | - | - | + |
| possessive pronoun + *body* expressing middle contexts (NC #3) | + | + | + |

TABLE 4 *Contexts of semantic spread of BPTs in Indian Creole Portuguese varieties*

| New contexts (NCs) for BPRs | Diu CP | Kannur CP | Kochi CP | Korlai CP | Mangalore CP | Sri Lanka CP |
|---|---|---|---|---|---|---|
| *body* vs. *soul* dichotomy (NC #1) | - | - | - | - | - | + |
| possessive pronoun + *body* in battle situations (NC #2) | - | - | - | - | - | - |
| possessive pronoun + *body* expressing middle contexts (NC #3) | + | + | + | - | + | + |

TABLE 5 *Contexts of semantic spread of BPTs in Philippine Creole Spanish varieties*

| New contexts (NCs) for BPRs | Cavite CS | Cotabato CS | Ternate CS | Zamboanga CS |
|---|---|---|---|---|
| *body* vs. *soul* dichotomy (NC #1) | + | - | - | + |
| possessive pronoun + *body* in battle situations (NC #2) | + | - | + | + |
| possessive pronoun + *body* expressing middle contexts (NC #3) | + | - | + | + |