

When corpus analysis refutes common beliefs: the case of interpolation in European Portuguese dialects

Catarina Magro

**Édition électronique**

URL : <http://journals.openedition.org/corpus/1859>

DOI : 10.4000/corpus.1859

ISSN : 1765-3126

Éditeur

Bases ; corpus et langage - UMR 6039

Édition imprimée

Date de publication : 1 novembre 2010

Pagination : 115-136

ISSN : 1638-9808

Référence électronique

Catarina Magro, « When corpus analysis refutes common beliefs: the case of interpolation in European Portuguese dialects », *Corpus* [En ligne], 9 | 2010, mis en ligne le 04 juillet 2011, consulté le 08 septembre 2020. URL : <http://journals.openedition.org/corpus/1859> ; DOI : <https://doi.org/10.4000/corpus.1859>

When Corpus Analysis Refutes Common Beliefs: the Case of Interpolation in European Portuguese Dialects

Catarina MAGRO

Centro de Linguística da Universidade de Lisboa

1. Introduction

It is well known that EP differs from most Romance languages with respect to clitic placement in that enclisis and proclisis occur both in finite and non-finite domains. In standard EP, the two patterns are in complementary distribution in finite clauses: proclisis is triggered by the presence of an overt complementizer, a negation operator, certain adverbs in preverbal position, a displaced *wh*-, quantificational or ‘affective’ phrase; enclisis surfaces in the remaining environments. The two alternative patterns are illustrated below, the sentences in (1) displaying proclisis while that in (2) displays enclisis.

- (1a) O rapaz disse que **lhe** telefonou.
the boy say_{PST.3SG} that him_{CL.DAT} phone_{PST.3SG}
The boy said that he called him.
- (1b) O rapaz não **lhe** telefonou.
the boy not him_{CL.DAT} phone_{PST.3SG}
The boy didn’t call him.
- (1c) Quem **lhe** telefonou?
who him_{CL.DAT} phone_{PST.3SG}
Who called him?
- (1d) Muitos rapazes **lhe** telefonaram.
many boys him_{CL.DAT} phone_{PST.3SG}
Many boys have called him.

- (1e) Até esse rapaz **lhe** telefonou.
 even that boy him_{CL.DAT} phone_{PST.3SG}
Even that boy called him.
- (1f) Todos **lhe** telefonaram.
 all him_{CL.DAT} phoned_{PST.3SG}
All the boys called him.
- (1g) Esse rapaz também **lhe** telefonou.
 that boy also him_{CL.DAT} phone_{PST.3SG}
Also that boy called him.
- (2) O rapaz telefonou-**lhe**.
 the boy phone_{PST.3SG}-him_{CL.DAT}
The boy called him.

In the proclisis contexts, exemplified in (1), the clitic can be separated from the verb by intervening material. The phenomenon of non-adjacency between the proclitic and the verb (known as *interpolation* in the Romance philological literature) was a very productive phenomenon in Old Portuguese. During that period, a wide variety of constituents could be interpolated. The loss of generalized interpolation occurs in the 17th century and from then on only the negation operator *não* (not) can disrupt proclitic-verb sequences (cf. Martins 1994; Fiéis 2003; Namiuti 2008). Nowadays, the proclitic-verb adjacency is mandatory for the majority of speakers of standard EP, although *não* interpolation is still an available option for some others. The sentences in (3) illustrate the marked contrast between Old and Contemporary interpolation:

Old Portuguese – 14th century

- (3a) Se **me** Deus enton a morte não **deu**
 if me_{CL.DAT} god then the death not give_{PST.3SG}
If then God didn't give me death.

Contemporary standard EP

- (3b) O rapaz disse que **lhe** não **telefonou**
 the boy say_{PST.3SG} that him_{CL.DAT} not phone_{PST.3SG}
The boy said that he didn't call him.

*When Corpus Analysis Refutes Common Beliefs: the Case of
Interpolation in European Portuguese Dialects*

Along with the highly restrictive interpolation of standard EP, illustrated in (3b), there are dialectal varieties that behave more freely with respect to this phenomenon. The dialectal interpolation has been interpreted as the remains of the Old Portuguese system and, therefore, the grammars of these dialects have been regarded as conservative grammars (cf. Martins 1994; Barbosa 1996; Fiéis 2003). In this article I will claim that this is an erroneous idea, which stems from insufficient empirical support.

I will discuss this issue taking into account a large amount of unreported interpolation data drawn from the *Syntax-oriented Corpus of Portuguese Dialects – CORDIAL-SIN*. On the basis of this new evidence I will review the traditional perspective about this topic.

My analysis of dialectal interpolation accounts for (i) the specific properties of dialectal interpolation (contrasting with both the wide interpolation of Old Portuguese and the restrictive interpolation of standard contemporary EP) and (ii) the connection between the interpolation phenomenon and other clitic related phenomena displayed by dialectal varieties in the very same syntactic contexts where interpolation surfaces: free variation between proclisis and enclisis and clitic duplication flanking an interpolated element or a verbal form. My fundamental claim is that interpolation, clitic duplication, and unexpected enclisis are special instances of a single displacement operation taking place in the Morphology component of grammar, namely the operation of *metathesis* as formulated by Harris & Halle (2005).

This article is organized as follows. Section 2 describes the defining properties of the dialectal interpolation construction. Section 3 examines the challenges posed by the reported dialectal data and it evaluates earlier accounts of the interpolation phenomenon in the light of the new data. In Section 4 I present an alternative analysis of dialectal interpolation. Section 5 discusses the status of the dialectal phenomenon in the history of the construction. Section 6 closes the paper making the explanatory power of my approach clear.

2. The interpolation in contemporary EP dialects

2.1 *The considered dialectal corpus*

The empirical support for this work has been drawn from the *Syntax-oriented Corpus of Portuguese Dialects – CORDIAL-SIN*. This corpus is being built up since 1999, at the Linguistics Center of University of Lisbon (CLUL), within the scope of a research project aimed at promoting the study of European Portuguese dialect syntax, a fairly recent area of interest in Portuguese linguistic research.¹

CORDIAL-SIN is a corpus of spoken dialectal EP that collects a geographically representative body of excerpts of spontaneous and semi-directed speech, selected from the oral interviews gathered by the Linguistic Variation Team at CLUL in the course of several Dialect Geography projects. The corpus covers 42 locations within the (continental and insular) territory of Portugal and it compiles about 600 000 words.



Map I. Geographical distribution of CORDIAL-SIN locations

¹ The CORDIAL-SIN project is supported by national and European funding (PRAXIS XXI/P/PLP/13046/1998; POSI/1999/PLP/33275; POCTI/LIN/46980/2002; PTDC/LIN/71559/2006).

*When Corpus Analysis Refutes Common Beliefs: the Case of
Interpolation in European Portuguese Dialects*

Map I shows the geographical distribution of the CORDIAL-SIN locations. The compilation of CORDIAL-SIN represents a turning point in the history of Portuguese dialectology, because for the first time a corpus of dialectal syntactic data was set up and systematically studied.

As in the case of other Romance languages, traditional dialectal studies do not focus on syntax; thus, data gathering methods and instruments (like questionnaires) were not designed to elicit and collect syntactic information.

However, the archives of the CLUL Linguistic Variation Team contain raw syntactic data in those cases where spontaneous speech has been recorded during the fieldwork sessions (mostly to serve ethnographic purposes). These data have been retrieved, transcribed and annotated by the CORDIAL-SIN team. The resulting material constitutes the first corpus of dialectal Portuguese (and the only existing one so far).

The lack of syntactic data available to researchers explains why some features of Portuguese dialects remained unnoticed until recently, while others have been described and explained without enough empirical support, as in the case of the interpolation phenomenon.

On the other hand, dialectal variation became a central issue to syntacticists studying linguistic variation within the Principles & Parameters framework, following the proposals by Chomsky (1981) and subsequent work. This led to a need for an accurate and enlarged empirical basis on Portuguese dialects fulfilling the demands of dialect syntax comparative inquiry.

2.2 The dialectal interpolation data

The collection of data drawn from CORDIAL-SIN provides a general picture of contemporary dialectal interpolation. Besides the negation operator *não* (that can also be interpolated in standard contemporary EP), there is a wide range of constituents that dialectally may intervene between the proclitic and the verb. Table I lists the interpolated elements of the considered dialectal corpus.

Table I. The interpolated elements of *CORDIAL-SIN*

pronouns	<i>eu</i> (I) <i>ele/ela</i> (he/she) <i>nós/a gente</i> (we/the people[=we]) <i>eles/elas</i> (they) <i>esta</i> (this [feminine]) <i>isso</i> (that [neuter]) <i>isto</i> (this [neuter])
adverbs	<i>aqui</i> (here [+close to speaker]) <i>aí</i> (there [-close to speaker/+close to addressee]) <i>ali</i> (there [-close to speaker/addressee]) <i>cá</i> (here [+close to speaker]) <i>lá</i> (there [-close to speaker/addressee]) <i>agora</i> (now) <i>depois</i> (after) <i>então</i> (then) <i>hoje</i> (today) <i>ontem</i> (yesterday) <i>ainda</i> (still) <i>já</i> (already) <i>assim</i> (like this)
prepositional phrases	<i>para lá</i> (to there [=there]) <i>para aí</i> (to there [=there]) <i>a nós</i> (to we [=to us])
negation operator	<i>não</i>

These interpolated constituents –which belong to different morphosyntactic categories– play diverse syntactic and discursive functions, as it is presented at Table II.

*When Corpus Analysis Refutes Common Beliefs: the Case of
Interpolation in European Portuguese Dialects*

Table II. Categories and functions of the interpolated constituents

morphosyntactic category	syntactic/discursive function
pronouns	subjects
	peripheral expletives
	topics
adverbs	modifiers
	displaced objects
	displaced small-clauses predicates
prepositional phrases	modifiers
	displaced objects
	displaced clitic-doubling phrases

Examples (4) through (13) illustrate the interpolation of the different kinds of constituents.

pronominal subject

- (4) E a gente, para comer a carne da arca,
and the people to eat_{INF} the meat from.the freezer

sabe o que **lhe** a gente **faz**?
know_{PRS.3SG} what it_{CL.DAT} the people do_{PRS.3SG}

Do you know how we manage to eat the frozen meat?
(Covo, CORDIAL-SIN COV07)

peripheral expletive

- (5) Os carros é tudo em ferro
the cars be_{PRS.3SG} all in iron

como **se** ele **vê**
as SE_{CL} EXPL see_{PRS.3SG}

All the cars' components are made of iron, as it can be seen.
(Sto André, CORDIAL-SIN STA36)

pronominal topic

- (6) Eu até **me** eu **aborreço**
I even me_{CL.REFL} I bore_{PRS.3SG}

de ouvir aquelas coisas!
to listen_{INF} such things

To listen such things really upsets me!
(Carrapatelo, CORDIAL-SIN CPT54)

adverbial modifier

- (7) É o que mais se agora cria
be_{PRS.3SG} what more SE_{CL} now grow_{PRS.3SG}

é isto.
be_{PRS.3SG} this

Nowadays, this is mostly what grows here.
(Figueiró, CORDIAL-SIN FIG17)

- (8) Antigamente havia muita moagem, como eu
in.the.past there.be_{PST.3SG} many mill as I

lhe há um bocado expliquei,
you_{CL.DAT} there.be_{PRS.3SG} a while explain_{PST.1SG}

não é?
NEG be_{PRS.3SG}

In the past, there were many mill engines, as I told you a while ago, you know?
(Fiscal, 1BR0501A – 29:48)

adverbial object

- (9) Também os aqui meto.
also them_{CL.ACC} here put_{PST.1SG}

I also put them in here.
(Monsanto, CORDIAL-SIN MST29)

adverbial small-clause predicate

- (10) Quando me cá apanhei, disse assim:
when me_{CL.ACC} here get_{PST.1SG} say_{PST.1SG} like.this

When I got here, I said: [...]
(Unhais da Serra, CORDIAL-SIN UNS21)

*When Corpus Analysis Refutes Common Beliefs: the Case of
Interpolation in European Portuguese Dialects*

prepositional modifier

- (11) Eu não sei o que **me** para aí **fizeram**.
I not know_{PRS.1SG} what me_{CL.DAT} to there do_{PST.3PL}
I don't know what they did to me around there.
(Granjal, CORDIAL-SIN GRJ26)

prepositional object

- (12) Todos os dias **se** para lá **ia** achar túbara.
all the days se_{CL} to there go_{PST.3SG} find_{INF} mushrooms
Everyday we would go there to look for mushrooms.
(Lavre, CORDIAL-SIN LVR24)

prepositional clitic-doubling phrase

- (13) Normalmente era sempre com as charruas de ferro,
usually be_{PST.3SG} always with the plows of iron
que **nos** a nós **lembra**.
that us_{CL.INH} to we remember_{PRS.3SG}
*Usually, as far as we remember, it was always with plows
made of iron.*
(Sto André, CORDIAL-SIN STA21)

negation operator

- (14) Embem **me** não **agradando** a fatia,
In well me_{CL.DAT} not please_{GER} the slice
venho-me embora.
go_{PRS.1SG} away
When the job doesn't satisfy me, I go away.
(Cabeço de Vide, CORDIAL-SIN CBV15)

2.3 The class of the interpolated elements

CORDIAL-SIN provides a wide and varied inventory of interpolated elements. Actually, the interpolated constituents vary in morphosyntactic class, syntactic structure, grammatical and discursive function and metrical structure. However, contemporary dialectal interpolation is far from the generalized

and highly permissive interpolation of Old Portuguese.² We need thus to identify the grammatical property shared by all the interpolated elements, that is, the property that makes it possible for these elements (and not for any other) to be interpolated.³

I claim that these dialectal interpolated elements share a semantic property: all of them are referentially deficient elements with a deictic interpretation (elements bearing a [+dependent] feature, in formal terms).

Table III lists the dialectal interpolated elements according to their type of deixis. As the table shows, the traditional categories of deixis are evenly covered (and in some cases almost filled) by the dialectal interpolated elements.

Table III. Interpolated elements – types of deixis

personal deixis	<i>eu</i> (I) <i>ele/ela</i> (he/she) <i>nós/a gente</i> (we/the people[=we]) <i>eles/elas</i> (they) <i>a nós</i> (to we [=to us])
spatial deixis	<i>aqui</i> (here [+close to speaker]) <i>aí</i> (there [-close to speaker/+close to addressee]) <i>ali</i> (there [-close to speaker/addressee]) <i>cá</i> (here [+close to speaker]) <i>lá</i> (there [-close to speaker/addressee]) <i>para lá</i> (to there [=there]) <i>para aí</i> (to there [=there]) <i>esta</i> (this [feminine]) <i>isso</i> (that [neuter]) <i>isto</i> (this [neuter])
temporal deixis	<i>agora</i> (now) <i>depois</i> (after) <i>então</i> (then) <i>hoje</i> (today) <i>ontem</i> (yesterday) <i>ainda</i> (still) <i>já</i> (already)
manner deixis	<i>assim</i> (like this)

² Up to the 16th century, all kinds of subjects and IP-scrambled constituents could occur between the proclitic and the verb (Martins 1994, 2003, 2005).

³ In this respect, Barbosa (1996) and Fiéis (2001, 2003) propose, respectively, that dialectal interpolated elements are monosyllabic prosodic words and verbal adjunct heads. CORDIAL-SIN data don't confirm their proposals.

*When Corpus Analysis Refutes Common Beliefs: the Case of
Interpolation in European Portuguese Dialects*

With the exception of the negation operator *não*, all the dialectal interpolated elements are classified as deictics in a straight way. The special case of *não* will be addressed in section 5.

3. Previous interpolation analyses and new dialectal data

3.1 The syntactic approaches

The word order found in interpolation structures excludes a purely syntactic treatment of this phenomenon (Martins 1994, 2003, 2005; Fiéis 2001, 2003). There are several arguments against a syntactic approach to dialectal interpolation. Here, I will sketch three of them.

In the first place, it is problematic to associate the clitic with an invariable structural position. In interpolation constructions within a single dialect, the clitic either (i) precedes elements of the high left periphery, like peripheral expletives (cf. Carrilho 2005), as in (15a), or topics, as in (16a), or (ii) is preceded by elements of the IP domain, like aspectual adverbs, as in (15b)-(16b), or subjects:

- (15a) Assim é que se isto **poderia**
so be_{PRS.3SG} that SE_{CL} this_{EXPL} can_{COND.3SG}
fazer alguma coisa.
do_{INF} something

This is how one could do something.
(Melides, CORDIAL-SIN MLD08)

- (15b) Parece que já **se** não **damos**
seem_{PRS.3SG} that already_{ADV} SE_{CL} not go_{PRS.1PL}
com as mezinhas, não é?
with the traditional.medicines NEG be_{PRS.3SG}?

*It seems that we can't cure ourselves with the traditional
medicines anymore, isn't it?*
(Melides, CORDIAL-SIN MLD31)

- (16a) Eu até **me** eu **aborreço**
I even me_{CL.REFL} I bore_{PRS.3SG}

de ouvir aquelas coisas!
to listen_{INF} such things

To listen such things really upsets me!
(Carrapatelo, CORDIAL-SIN CPT54)

- (16b) até que lhe dei uma palavra que eu
until that him_{CL.DAT} give_{PST.1SG} a word that I

até já me não lembra
even already me_{CL.INH} not remember_{PRS.3SG}

como ela foi.
how it be_{PST.3SG}

At a certain point, I told him something but I can't remember my exact words anymore.
(Carrapatelo, CORDIAL-SIN CPT19)

On the other hand, the fact that a clitic may intervene between the elements of a phrase is incompatible with an analysis that derives interpolation by clitic movement across the interpolated element to a higher functional category.

- (17) [Ainda os ontem] meti no carroço.
[just them_{CL.ACC} yesterday] put_{PST.1SG} in.the cart

It was just yesterday that I put them in the cart.
(Monsanto, CORDIAL-SIN MST29)

Finally, constituents with similar syntactic properties are not affected by interpolation in the same way. The contrast between pronominal subjects and regular DP-subjects interpolation, illustrated below, reveals the particular shape of the dialectal construction. This contrast does not hold for Old Portuguese interpolation.

- (18) O rapaz disse que lhe ela telefonou.
the boy say_{PST.3SG} that him_{CL.DAT} she_{SBJ.PRON} phone_{PST.3SG}
The boy said that she called him.

- (19) *O rapaz disse que lhe a Maria telefonou.
the boy say_{PST.3SG} that him_{CL.DAT} Mary_{SBJ} phone_{PST.3SG}
The boy said that Mary called him.

3.2 The prosodic approaches

Likewise, the metrical structure of interpolated elements and some other aspects concerning clitics prosodization disprove the predictions made by an analysis that accounts for interpolation in prosodic terms (Barbosa 1996). Barbosa (1996) attributes dialectal interpolation to a restructuring process at the prosodic level that enables the clitic to form a unit with a prosodic phrase that exhaustively dominates a monosyllabic prosodic word. Her analysis predicts that interpolated elements are monosyllabic prosodic words and that leftward phonological cliticization is a ruled out option in these dialects' grammars. None of these predictions is supported by CORDIAL-SIN data. The list in (20) shows that interpolated elements have a variable metrical structure. In (21), the mid central vowel is produced as a glide. This means that this segment doesn't occur at the prosodic word boundary and concomitantly that the clitic is left cliticized.

- (20) a gente, agora, depois, então, para lá, há um bocado
the people now after then to therethere is a while
we now after then there a while ago

- (21) O linho quanto mais fininho **a gente** o
 [ɐ'ʒẽtju]
 the flax.thread as.much more thin the.people it_{CL.ACC}
 pudesse fazer, mais fininho o fazia.
 can_{SBJV.3SG} make_{INF} more thin it_{CL.ACC} make_{PST.3SG}

We made the flax-thread as thin as possible.

(Fiscal, CORDIAL-SIN FIS15)

4. An alternative analysis

My analysis of contemporary dialectal interpolation assumes the organization of grammar as envisioned by the theory of Distributed Morphology (Halle & Marantz 1993, 1994 and subsequent work), namely a Morphology component in PF with the ability to change the syntactic output.

Under my proposal, contemporary dialectal interpolation is the result of a morphological readjustment rule that manipulates a linearized syntactic output. Specifically, I claim

that this phenomenon is derived by a metathesis process, in the sense of Harris & Halle (2005), whereby a designated subsequence in a given phonological string is duplicated and the peripheral elements in each of the generated copies are deleted afterwards.

According to Harris & Halle (2005), metathesis is just a special instance of a more general duplication process in which a designated substring in a base form can be repeated in its entirety (*full duplication*) or in part (*partial duplication*) in a derived form.

The formalism of Harris & Halle (2005) is presented below. The abstract derivation in (26) displays the metathesis device as conceived by the authors.

- (22) ABCD linearized string
- full duplication
- (23) A[BC]D duplicating subsequence
A-BC-BC-D result of duplication
- partial duplication
- (24) A[B>C]D duplicating subsequence + deletion instructions
A-~~B~~C-BC-D result of duplication + deletion effects
A-C-BC-D final result
- (25) A[B<C]D
A-BC-B~~C~~-D
A-BC-B-D
- partial duplication with metathesis
- (26) ABCD linearized string
A[B><C]D duplicating subsequence + deletion instructions
A-~~B~~C-B~~C~~-D result of duplication + deletion effects
A-C-B-D final result

In my analysis, the metathesis operation is triggered when the deictic-clitic sequence occurs in a linearized string and is motivated by the [+dependent] feature borne by the two elements involved (note that the clitic is a referentially dependent element itself). My hypothesis is that the grammar of EP dialects that exhibit interpolation includes the following rule:

*When Corpus Analysis Refutes Common Beliefs: the Case of
Interpolation in European Portuguese Dialects*

(27) in a string of the form $X Y_{dep} cl Z$, insert

- [to the immediate left of Y_{dep}
-] to the immediate right of cl
- < to the immediate right of Y_{dep}
- > to the immediate left of cl

The application of such a rule will produce an output where the deictic and the clitic reverse their linear order:

(28) $X Y_{dep} cl Z$
 $X [Y_{dep} > <cl] Z$
 $X - \overline{Y_{dep}} cl - Y_{dep} \overline{cl} - Z$
 $X - cl - Y_{dep} - Z$

Since metathesis and duplication are just different instances of the same operation, we should expect to find clitic duplication data in the same syntactic contexts in which these informants produce interpolation. In fact, CORDIAL-SIN data validate this prediction. Examples (29)-(30) illustrate the interpolation with clitic duplication construction. Note that in these cases the two copies of the clitic are flanking the interpolated element.

(29) Se não se fizesse muito, como é que
 if not SE_{CL} work_{SBJV.3SG} enough how be_{PRS.3SG} that
me eu me safava?!
 me_{CL.REFL} I me_{CL.REFL} make_{PST.3SG}
If we hadn't worked enough, how would I have made it?
 (Melides, CORDIAL-SIN MLD21)

(30) O pão tem de ser posto com uma
 the bread have_{PRS.3SG} to be_{INF} put with a
 vasilha que **se lá se** põe.
 bowl that SE_{CL} there SE_{CL} put_{PRS.3SG}
*We have to put the bread [on the oven] with a bowl that we
 put there.*
 (Serpa, 1B03b07b, 11:40)

Clitic duplication will thus be the Spell-Out of a partial duplication operation:

- (31) $X Y_{\text{dep}} \text{cl} Z$
 $X [Y_{\text{dep}} > \text{cl}] Z$
 $X - \text{cl} - Y_{\text{dep}} \text{cl} - Z$
 $X - \text{cl} - Y_{\text{dep}} \text{cl} - Z$

5. A different history of interpolation

From what has been said so far, we are naturally led to formulate the following question: what is the connection between contemporary dialectal interpolation and Old Portuguese interpolation, given the apparent post-syntactic nature of the first one and the pure syntactic status of the last one? I claim that dialectal interpolation doesn't have its direct origins in Old Portuguese interpolation, corresponding, on the contrary, to a recent innovation in the history of Portuguese.

Two facts can be invoked to support such a hypothesis: (i) texts from the 19th century start to attest interpolation data, similar to the dialectal interpolation data, after a gap of one hundred and fifty years (texts from the second half of the 17th century and from all the 18th century don't attest interpolation data of any constituent but *não*) and (ii) dialectal EP contrasts with dialectal Galician in what concerns the interpolation phenomenon, what suggests the late emergence of dialectal interpolation in the history of Portuguese (happening at a time in which Portuguese and Galician are distinct linguistic systems).

Under my hypothesis, these morphological displacement operations emerge in Portuguese grammar in the 17th century as the outcome of a reanalysis process of *não* interpolation. In this reanalysis process the clitic-*não*-verb sequence is interpreted as the output of a metathesis operation involving the clitic and the functional category Σ (Laka 1990; Martins 1994 and subsequent work). The relevant contexts for the reanalysis to take place are those in which Σ is endowed with a [+dependent] feature, that is, the contexts where the requirement concerning the strong nature of Σ is satisfied by syntactic merger between Σ and C (cf. Costa & Martins 2003, 2004). The metathesis operation then spreads over other contexts where the clitic is preceded by an element bearing a [+dependent] feature as well, that is, a deictic element. Under

*When Corpus Analysis Refutes Common Beliefs: the Case of
Interpolation in European Portuguese Dialects*

this alternative view, the deictic interpolation data are the manifestation of a recent and innovating change.

In this picture, the case of the interpolation of *não* plays a central role in the history of interpolation in Portuguese. In fact, its course along the time is the key to explain the contrast between: (i) the generalized syntactic interpolation of Old Portuguese, (ii) the highly restricted interpolation of contemporary standard Portuguese and (iii) the less constrained and lexically-oriented interpolation of contemporary dialectal varieties.

6. Other cases of metathesis and duplication in EP dialects

Such an analysis is descriptively adequate as it thoroughly covers the identified interpolation data (a welcome result in itself), and allows us to account for seemingly disparate phenomena as special cases of a single grammatical device.

In addition to interpolation, let's consider two other phenomena displayed by dialectal varieties in the very same syntactic contexts (standard proclisis contexts) where interpolation surfaces: (i) free variation between proclisis and enclisis and (ii) clitic duplication flanking a verbal form or an interpolated element. This cluster of constructions is illustrated in (32)-(36) with data produced by a speaker of the dialect of Lavre (a southern EP variety from Alentejo).

- (32) É daí que se **gera** as arrãs.
be_{PRS.3SG} from.there that SE_{CL} originate_{PRS.3SG} the frogs
It is from there that frogs come.
- (33) E era assim que se a gente **vivia**.
and be_{PST.3SG} like.this that SE_{CL} the.people live_{PST.3SG}
And it was like this that we lived.
- (34) Ainda hoje se lá se **conserva** aquele
even today SE_{CL} there SE_{CL} stay_{PRS.3SG} that
bocadinho.
piece
Even today that piece stays there.

- (35) É aí que eu que
 be_{PRS.3SG} at.that.point that I that

ia-lhe para contar...
 be.going_{PST.1SG-you_{CL.DAT}} to tell_{INF}
It was at that point that I was about to tell you...
- (36) É essa tira que **lhe** **chamam-lhe**
 be_{PRS.3SG} that strip that it_{CL.DAT} call_{PRS.3PL-it_{CL.DAT}}

 o baço.
 the spleen
It is that strip that is called the spleen.

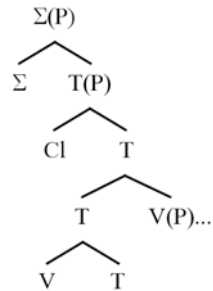
Besides standard proclisis with adjacency, as in (32), this speaker produces, in typical proclisis contexts, interpolation, as in (33), interpolation with clitic duplication, as in (34), regular enclisis, as in (35), and pre and post verbal clitic duplication, as in (36). All these constructions coexist in the speech of a single individual and they are not occasional performance errors but systematic realizations. To the best of my knowledge, this is an unattested paradigm across the Romance languages.

I take these coexistent constructions to be closely related phenomena and I propose a formal treatment that captures their relatedness appropriately, which previous interpolation analyses have not been able to do. My fundamental claim is that interpolation, unexpected enclisis and clitic duplication are special instances of a single displacement operation taking place in the Morphology component of grammar.

Assuming Costa & Martins (2003, 2004) analysis of clitic placement in EP, I propose that the object derived in the syntactic component of the grammar corresponds to that in (37), in which the clitic is left-adjoined to the maximal projection of T (clitics left-adjoin to the highest functional head targeted by verb movement, that, in most cases, corresponds to T in EP) (cf. Kayne 1991).

*When Corpus Analysis Refutes Common Beliefs: the Case of
Interpolation in European Portuguese Dialects*

(37)



The linearization of this object in PF will originate the following string:

(38) [Σ * [Cl * [V * T]]]

I hypothesize that whenever the visibility requirement of Σ is satisfied in the syntactic component (where Σ is lexicalized under syntactic merger or under Σ-to-C movement),⁴ the clitic is free to enter metathesis operations with an adjacent element both to its left and to its right. That is, I propose that the grammar of dialects with interpolation includes a metathesis/duplication rule that operates in both directions.

When none of these rules applies, standard proclisis is displayed, as in (39). The rule of left metathesis/duplication will derive plain interpolation and interpolation with clitic duplication, as in (40); similarly, the rule of right metathesis/duplication will derive regular enclisis and clitic duplication in pre and post verbal position, as in (41).

(39) clitic-verb

(40a) clitic-interpol-verb left metathesis

(40b) clitic-interpol-clitic-verb left duplication

(41a) verb-clitic right metathesis

(41b) clitic-verb-clitic right duplication

⁴ According to Costa & Martins (2003, 2004) proposal, Σ can be licensed in PF, as a last resource, under morphological merger between Σ and the verb. For the morphological merger to take place, Σ and the verb must be adjacent. In these cases, Local Dislocation with inversion applies and the clitic becomes enclitic.

Finally, I suggest that this kind of operations affects adjacent elements bearing the same feature. In this particular case, it is the fact that the clitic bears a dependent-feature and V-features that makes it possible for the clitic to enter metathesis/duplication operations with a deictic element that may precede it and with the verb that always follows it.

References

- Barbosa P. (1996). « Clitic placement in European Portuguese and the position of subjects », in A. Halpern & A. Zwicky (eds), *Approaching Second: Second Position Clitics and Related Phenomena*. Stanford, California : CSLI Publications, 1-40.
- Carrilho E. (2005). *Expletive Ele in European Portuguese Dialects*. PhD Dissertation, Universidade de Lisboa.
- CORDIAL-SIN – Syntax-oriented Corpus of Portuguese Dialects*, http://www.clul.ul.pt/english/sectores/variacao/cordialsin/projecto_cordialsin.php
- Costa J. & Martins A. M. (2003). « Clitic placement across grammar components », *Going Romance 2003*. Nijmegen University, November 2003.
- Costa J. & Martins A. M. (2004). « What is a strong functional head ? », *Lisbon Workshop on Alternative Views on the Functional Domain*. Universidade Nova de Lisboa, July 2004.
- Fiéis A. (2001). « Interpolação no português medieval como adjunção a XP », *Actas do XVI Encontro Nacional da Associação Portuguesa de Linguística*. Lisboa : APL, 197-211.
- Fiéis A. (2003). *Ordem de Palavras, Transitividade e Inacusatividade. Reflexão Teórica e Análise do Português dos Séculos XIII a XVI*. PhD Dissertation, Universidade Nova de Lisboa.
- Halle M. & Marantz A. (1993). « Distributed Morphology and the pieces of inflection », in K. Halle & S. J. Keyser (eds), *The View from Building 20 : Essays in Linguistics in Honor*

- When Corpus Analysis Refutes Common Beliefs: the Case of Interpolation in European Portuguese Dialects*
of Sylvain Bromberger. Cambridge, Massachusetts : MIT Press, 111-176.
- Halle M. & Marantz A. (1994). « Some key features of Distributed Morphology », in A. Carnie & H. Harley (eds), *MITWPL 21 : Papers on Phonology and Morphology*. Cambridge, Massachusetts : MIT Press, 275-288.
- Harris J. & Halle M. (2005). « Unexpected plural inflections in Spanish : reduplication and metathesis », *Linguistic Inquiry* 36, 2 : 195-222.
- Laka I. (1990). *Negation in Syntax. On the Nature of Functional Categories and Projections*. PhD Dissertation, MIT.
- Magro C. (2007). *Clíticos : Variações sobre o Tema*. PhD Dissertation, Universidade de Lisboa.
- Martins A. M. (1994). *Clíticos na História do Português*. PhD Dissertation, Universidade de Lisboa.
- Martins A. M. (2003). « From unity to diversity in Romance syntax : Portuguese and Spanish », in K. Braunmuller & G. Ferraresi (eds), *Aspects of Multilingualism in European Language History*. Amsterdam & Philadelphia : John Benjamins, 201-233.
- Martins A. M. (2005). « Clitic placement, VP-ellipsis and scrambling in Romance », in M. Batllori, M-L Hernanz, C. Picallo & F. Roca (eds), *Grammaticalization and Parametric Change*. Oxford : Oxford University Press, 175-193.
- Namiuti C. (2008). *Aspectos da História Gramatical do Português: Interpolação, Negação e Mudança*. PhD Dissertation, Universidade Estadual de Campinas.

