

Phases and autonomous features: a case of mixed agreement in European Portuguese

João Costa & Sandra Pereira

Universidade Nova de Lisboa & Centro de Linguística da Universidade de Lisboa

This paper studies a case of mixed agreement in European Portuguese, triggered by the pronoun *a gente*. This pronoun has active non-agreeing referential and grammatical features that emerge in different local domains. In predicative contexts, the two sets of features become visible. We analyze this fact, by proposing that phases are the relevant domains for agreement to operate and that the two types of features are separately represented in the featural hierarchy of pronouns.

1. Introduction and goals.

The pronominal system of European Portuguese includes the expression *a gente*. This expression, which literally means “the people”, is used as a first person plural pronoun, generally occurring in free variation with the first person plural pronoun *nós*. Menuzzi (2000) shows that this expression truly behaves like a pronoun, since it obeys conditions B and C of Binding Theory like the other pronouns of the pronominal paradigm, as illustrated in the examples below.

In (1), it is shown that *a gente* behaves like a pronoun for the sake of principle B effects, since it can only be bound non-locally:

Condition-B:

- (1) a. *[Eu e o Pedro]_i vimos a gente_i.
I and Pedro saw A GENTE
b. [Eu e o Pedro]_i dissemos que a gente_i adorou a festa.
I and Pedro said that A GENTE loved the party

In (2) and (3), it is shown that *a gente* behaves like the first person plural pronoun *nós* for the sake of compliance with principle B and C:

- (2) a. [Eu e o Pedro]_i dissemos que a gente_i adorou a festa.
I and Pedro said that A GENTE loved the party
b. [Eu e o Pedro]_i dissemos que nós_i adorámos a festa.
I and Pedro said that we loved the party
- (3) a. *A gente_i viu [a mim e ao Pedro]_i.
A GENTE saw me and Pedro
b. *Nós_i vimos [a mim e ao Pedro]_i.
We saw me and Pedro

Further evidence for the status of this expression as a pronoun comes from the fact that it behaves like other pronouns as far as the avoid pronoun principle is concerned:

- (4) a. Nós achamos que somos felizes.
We think that are happy
b. A gente acha que é feliz.
A GENTE thinks that is happy

Also, like other pronouns, *a gente* resists adjectival modification, as illustrated in (5):

- (5) a. *Nós felizes fomos à praia.
We happy went to the beach
b. *Eu feliz fui à praia.
I happy went to the beach
c. *A gente feliz foi à praia.
A GENTE happy went to the beach

Finally, apart from its referential reading, like other pronouns, *a gente* allows arbitrary readings, as illustrated in (6):

- (6) A gente tenta sempre fugir ao fisco.
A GENTE tries always to run from taxes

These data, discussed in Menuzzi (2000), provide compelling evidence that *a gente* behaves like other pronouns of the pronominal paradigm of Portuguese.

An aspect in which this form differs from other pronouns, making it particularly interesting, is that its referential and grammatical features do not coincide. In fact, the pronoun *a gente* refers to first person plural, alternating with the pronoun *nós* (we), but it is grammatically specified as third person singular. The grammatical specification becomes obvious in the fact that this pronoun triggers third person singular agreement on the verb. In this respect, as mentioned by Menuzzi (2000), the pronoun *a gente* behaves like the French pronoun *on*.

Menuzzi (2000) shows that both the semantic and the grammatical features are active and relevant for binding purposes. This author shows that the grammatical features emerge in the establishment of local binding relations, as illustrated in (7). (7a) shows that 3rd person singular agreement is triggered on the verb; in (7b), it is shown that in local binding relations, the anaphoric form selected is of the SE-type, which is specified with 3rd person singular features as well:

- (7) a. A gente (vai/*vamos) à praia.
The people goes/*go_{1st plural} to the beach.
b. A gente vê-(se/*nos) no espelho.
The people sees herself/*ourselves in the mirror

From (7), one might suppose that the referential features of *a gente* are just a matter of interpretation without reflexes on the syntax. However, Menuzzi (2000) shows that non-locally, *a gente* selects a pronominal form specified for the referential features. As shown in (8), the pronominal form selected in non-local environments must be specified with 1st person plural features.

- (8) A gente disse que a Maria (nos/*a) viu.
The people said that Maria us/*her saw

These data provide the necessary evidence to posit that the two sets of features play a role in understanding the distribution and behavior of this pronominal form. According to Menuzzi (2000), there is a correlation between the locality of the binding relation established with *a gente* and the set of features that becomes salient.

An interesting puzzle arises when predicative contexts are considered. Since adjectival and participial predicates bear agreement morphology, it is interesting to know whether these forms will agree with the grammatical features of the pronoun or with its referential features. Although, at first sight, there seems to be a great deal of variation in the data, a quantitative study carried out in Pereira (2000, 2003)¹ reveals that the most predominant pattern is the one illustrated in (9):

- (9) a. A gente está cansados.
A GENTE is tired-masc-pl
b. A gente está cansadas.
A GENTE is tired-fem-pl
c. Minha mãe que Deus haja, ia para a casa de uma tia
minha, defronte ao hotel Porto Santo, e
a gente ia, **pequeninas**, a gente ia.
A GENTE went-3sg little-fem-pl
d. Se ele pôr mais, arrebenta com **a gente** todos, **marítimos**
A GENTE all-masc-pl sailors-
masc.pl

The data obtained in Pereira (2000, 2003) reveal that *a gente* triggers 3rd person singular agreement on the verb, but 1st person plural agreement on the predicate. The variation in gender on the predicate depends on the referent for the pronoun (masculine or feminine). The conclusion we can draw from these results is that there is a case of mixed agreement within one single IP. This is obviously problematic for Menuzzi's conclusion that the activation of grammatical or referential features was dependent on the locality of the relation between *a gente* and the agreeing form, since, in the cases illustrated in (9), the relation between the predicates and *a gente* is local.

Given the facts just described, the goal of this paper is to answer the following questions:

¹ The methodology used in Pereira (2000, 2003) for analyzing these data was as follows: at a first stage, a collection of occurrences of the relevant type of construction in different types of *corpora* was carried out (CORDIAL-Sin and CRPC); at a second stage, the data obtained from the *corpora* were completed so as to include all types of relevant syntactic environments. After the data was compiled, a grammaticality judgement task was conducted, in which informants were asked to complete participial and adjectival forms by adding the relevant agreement markers.

- a) How can the mixed agreement in predicative sentences be accounted for?
- b) How can the grammatical and referential features be set apart?

The paper is organized as follows: in section 2, we spell out some assumptions regarding the hierarchical organization of the feature specification of pronouns, and concerning the locality of agreement relations. In section 3, we present the analysis based on the idea that words are inserted at Morphological Structure (Halle and Marantz 1993). Finally, in section 4, some predictions made by this hypothesis are explored.

2. Assumptions.

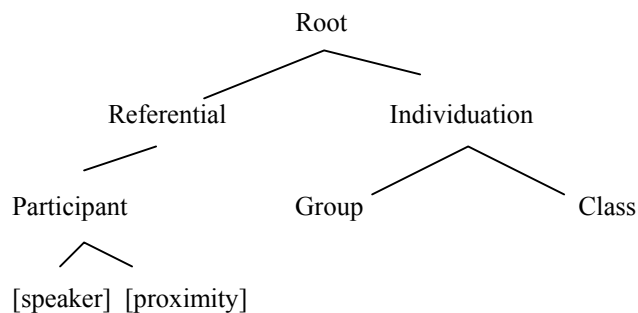
The analysis we are about to propose for the pattern of mixed agreement described above is based on the assumption that the features of the pronominal forms do not work as a bundle, and that the Phase is a relevant domain for syntactic agreement to be established.

Let us review each of these assumptions individually.

- a) The feature specification of pronouns is hierarchically organized. Different sets of features group together independently of other features.

This assumption is independently motivated in work by Harley and Ritter (1999), Duarte et al. (2002, among others). According to these authors, the features of a pronoun are subject to an internal hierarchy. It is not particularly crucial for our goals what the exact internal hierarchy of the pronoun *a gente* is as long as the grammatical features and the referential features are set apart. This is achieved in the following feature specification for the pronominal *a gente*, adapted from Duarte et al. (2002):

- (10) Feature specification of *a gente*:



As can be observed in the information given in (10), the grammatical (Individuation) and referential features (Referential) are separate. We will further assume that each of the nodes in the hierarchical representation can be activated in different domains. For instances, assuming this type of hierarchy, and coming back to Menuzzi's observations regarding the

binding facts, one must assume that only the referential node is active in non-local domains.

Note that we are aware of the fact that, by itself, this is not an explanatory device. For the purposes of this paper, what is relevant for us is that the two sets of features are isolated, and that one does not consider that the features of pronominal form work as a non-structured bundle.

- b) Phases are the domains for agreement operations to take place (Chomsky 1998, 1999).

The second assumption we make is inspired on Chomsky's (1998, 1999) proposal that Agree operates phase-internally. This assumption is independently motivated by the observation that agreement, when it occurs, is the morphological reflex of a local relationship between two agreeing elements.²

- c) The small clause is a Phase.

Related to the previous point, we propose that the Small Clause is a phase-boundary, as independently proposed in Svenonius (2003). Although there is little consensus regarding the exact definition of a Phase, three aspects motivate this assumption:

- First, Chomsky (1998, 1999) proposes that a phase is a coherent unit in terms of reference. As is well known, Small Clauses constitute a complete domain of predication, in which theta-roles are fully saturated.
- Second, small clauses are prosodically coherent, which is another characteristic of phases, according to Chomsky (1998,1999). The example in (11) illustrates this property, since each of the coordinated small clauses may be argued to constitute an independent prosodic unit:

- (11) Eu considero [a Maria bonita], [a Olga gorda] e [o Pedro feio].
I consider Maria pretty, Olga fat and Pedro ugly

- Finally, it is well-known that small clause predicates do not scramble, as the contrast between DP-scrambling and small-clause predicate scrambling illustrated in (12) attests:

- (12) a. O Pedro fala francês sempre.
Pedro speaks French always.
b. *O Pedro está cansado sempre.
Pedro is tired always

² It is important to note that we are not assuming that lack of morphological agreement signals the lack of establishment of an Agree relation between two elements. The assumption just states that whenever morphological agreement is visible, it reflects the existence of a local Agree relation.

If short-distance-scrambling must be internal to a Strong Phase, the difference between nominal objects and small clause predicates follows. The predicate of the small clause cannot be scrambled, because it would have to cross the strong phase boundary, not being able to use the edge of the phase as an escape hatch.

- d) Words are inserted in the morphological component. Syntax only handles features (Halle and Marantz 1993).

A final assumption we need to make is that, following Halle and Marantz (1993) lexical insertion is post-syntactic. This assumption ensures that the selection of the pronominal form agreeing with *a gente* in the case of binding, or the selection of the agreement marker in the predicative contexts will be sensitive to the features activated in each domain, and not to the whole set of features present in the feature specification of *a gente* given in (10). Without this assumption, it becomes difficult to explain why only some features are activated in certain domains, and why not all features are checked in all steps of a derivation, as argued in Costa, Moura and Pereira (2001).

3. Analysis.

Given the assumptions spelled out in the previous section, we have the ingredients for deriving the pattern of mixed agreement observed.

The subject of the sentence is generated as the subject of the small clause predicates. Within the small clause domain, the semantic features locally probe the predicate for agreement, specifying its features for 1st person plural. It is important to clarify why the referential features are the first features to be activated. Following Costa and Figueiredo Silva (2002), we assume that referential number agreement is a dissociated morpheme in the sense of Embick and Noyer (2001) and adjoins to each potential root. Costa and Figueiredo Silva (2002) present evidence for this claim based on a comparison of the agreement patterns in three registers of Portuguese. It is argued by these authors that referential number agreement is a dissociated morpheme in European Portuguese, and a singleton in Brazilian Portuguese. This proposal is based on the following facts:

- i) In European Portuguese (EP), there is full DP-internal agreement;
- ii) In the two registers of Brazilian Portuguese (BP1 and BP2) studied, number within the DP is marked either on the D head or in all pronominal elements;
- iii) Pronominal adjectives and possessives differ in BP, in the sense that only the latter may carry number marking in the absence of plurality on the definite article;
- iv) In EP, there is full subject-verb agreement, independently of the position of the subject, except in the case of unaccusative verbs in which agreement is optional with inverted subjects in colloquial speech;
- v) In EP, there is full number agreement with passive participles and SC predicates;
- vi) In BP1, there is full subject-verb agreement, except for the case of inverted subjects;
- vii) In BP1, there is no number agreement with passive participles and SC predicates;
- viii) In BP2, there is no number subject-verb agreement;

- ix) In BP2, there is no number agreement with passive participles and SC predicates.

According to this analysis, the first step of the derivation proceeds as in (13): the features of *a gente* are merged at the edge of the small clause, and due to the dissociated nature of number in Portuguese, the number features of the predicate become specified and valued.

- (13) a. $[_{SC} DP \ [[3^{rd} \text{ sg}] \ [pl]] \ AP \ [pl]]$
 b. $[_{SC} DP \ [[3^{rd} \text{ sg}] \ [\cancel{pl}]] \ AP \ [\cancel{pl}]]$

Assuming that the number marking on the small clause predicate partially follows from the dissociated nature of the morpheme is not crucial for the analysis, but makes it possible to dispense with a theoretical apparatus in which it would be necessary to posit independent processes for establishing agreement relations within the same language.

Since the small clause is a strong phase, only its edge is accessible for further operations. This enables movement of the subject to Spec,IP in order to check EPP-features, as illustrated in (14):

- (14) $[_{IP} DP \ [[3^{rd} \text{ sg}] \ [\cancel{pl}]] \ I \ [_{SC} DP \ [[3^{rd} \text{ sg}] \ [\cancel{pl}]] \ AP \ [\cancel{pl}]]$

Due to the Phase Impenetrability Condition, the only features that can be accessed within the IP domain are the unvalued ones: the grammatical features. This explains why 3rd person singular verbal agreement is found on the verb. In other words, the claim is that checked features are not available for further agreement operations in other domains. Since the referential number features were valued within the small clause, they will no longer be valued in the Infl domain.

This analysis has the advantage of bringing the case of mixed agreement close to the generalization obtained by Menuzzi (2000) for the binding facts. For both sets of facts, the explanation lies on the observation that there are two different domains deriving the activations of different parts of the feature specification of the pronoun *a gente*. The case of mixed agreement discussed in this paper reproduces the binding pattern, in the sense that in both cases, the grammatical features emerge locally, while the referential features emerge non-locally. The assumption that the small clause is a strong phase derives the non-local facet of agreement in predicative contexts. Agreement in this context is similar to the selection of 1st person plural pronouns in non-local domains, because the pronoun *a gente* and the agreeing predicate surface in a non-local relationship.

It is important to emphasize that without late lexical insertion, there would be a problem for checking theory, as argued in Costa, Moura and Pereira (2001), since the pronoun would be inserted in the derivation endowed with the complete feature specification. Since the predicate and the verb should be inserted already inflected, there would be a clash between the features of the verb and the predicate and part of the features of the pronoun.

4. Further Predictions.

The analysis put forward in the preceding section makes two additional predictions:

a) *Variation in agreement on the verb and a difference between European and Brazilian Portuguese:*

Although in some dialects of European Portuguese, *a gente* triggers 1st person plural verbal agreement, there is no dialect in which the pronoun triggers plural agreement on the verb, and singular agreement on the predicate, given the independent nature of the two types of features.

- (15) a. A gente estamos contentes.
A GENTE are-1st-pl happy-pl
b. *A gente estamos contente.
A GENTE are-1st-pl happy-sg

As mentioned in the previous section, we contend that the plural referential agreement visible on the predicate is a consequence of the dissociated status of the referential number morphology, as argued in Costa and Figueiredo Silva (2002). According to these authors, in European Portuguese, plural dissociated morpheme attaches to all potential hosts. On the other hand, the grammatical agreement follows only from the feature specification of the pronoun. This derives the ungrammaticality of (15b). Since the plural morphology on the verb is a consequence of the dissociated status of the number morpheme, (15b) is ruled out because the number morpheme is not attaching to all roots, not surfacing on the predicate.

Costa and Figueiredo Silva (2002) hypothesize that number is a singleton morpheme in Brazilian Portuguese. This correctly predicts that sentences like (15a) are less frequent in Brazilian Portuguese than in European Portuguese. On the contrary, in European Portuguese, (15a) is becoming more and more frequent. Accepting this difference in the status of the morpheme in the two languages further predicts that in Brazilian Portuguese, *a gente* never triggers plural agreement on the predicate of the small clause.

b) *Behavior of a gente in null subject contexts.*

The behavior of this pronoun with respect to the null subject parameter is different from other pronouns. In section 1, it was observed that normally *a gente* freely alternates with the first person plural pronoun *nós* (we). However, in null subject contexts, it can be observed that *nós* can be dropped, but *a gente* cannot, as illustrated in (16):

- (16) a. Vamos embora. (we) go away.
b. *Vai embora. (A GENTE) goes away.

According to the analysis outlined above, the problem with (16b) is that the referential features of the subject are not present, which prevents recoverability of its reference. In other words, it may be said that the null subject in (16b) is not referential.

Incidentally, this fact may indicate that Agr is not necessarily pronominal in null subject languages, casting doubt on recent views on the null subject parameter, such as Barbosa (1995) and Alexiadou and

Anagnostopoulou (1998), who claim that Agr is pronominal in null subject languages. These facts indicate that the pronominal nature of Agr is not defineable for a whole verbal paradigm, but it may vary depending on the type of subject associated with it.³

5. Conclusion.

The pattern of agreement analyzed in this paper was never explained. Combining phase theory and some of the assumptions of Distributed Morphology provides a straightforward analysis for this problem. Therefore, this type of phenomenon is the kind of empirical evidence supporting these recent developments in syntactic theory.

References

- Alexiadou, Artemis and Elena Anagnostopoulou. 1998. Parametrizing Agr: Word order, verb-movement and EPP-checking. *Natural Language and Linguistic Theory*. 16:3, 491-539
- Barbosa, Pilar. 1995. Null Subjects. Doctoral dissertation, MIT.
- Chomsky, Noam. 1998. "Minimalist Inquiries", MIT Occasional Papers in Linguistics
- Chomsky, N. 1999. "Derivation by Phase". MIT Occasional Papers in Linguistics. Number 18.
- CORDIAL-SIN: *Corpus of Portuguese Dialects Syntactically Annotated* http://www.clul.ul.pt/english/sectores/cordialsin/projecto_cordialsin.html
- Costa, João. 2001. "Null vs Overt Spec,TP in European Portuguese" to appear in J. Schroten (ed) *Romance Languages and Linguistic Theory 2001*, Johns Benjamins, Amsterdam
- Costa, João & M^a Cristina Figueiredo Silva. 2002. "Nominal and verbal agreement in Portuguese: an argument for Distributed Morphology". Ms, Universidade Nova de Lisboa
- Costa, João., Denilda Moura & S. Pereira. 2001. *Concordância com a gente: um problema para a teoria de verificação de traços*. In Actas do XVI Encontro Nacional da Associação Portuguesa de Linguística. Coimbra. Setembro de 2000.
- Duarte, Inês, M. João Freitas, Anabela Gonçalves, Matilde Miguel & Celeste Rodrigues. 2002. "Geometria de traços e distribuição de pronomes sujeito em PE e em PB". Paper presented at the 3rd Workshop do Projecto PE-PB. Lisboa, September 23-25.
- Embick, D. and Noyer, Rolf 2001. "Movement operations after Syntax". *Linguistic Inquiry* 32: 555-595.
- Halle, M. e A. Marantz. 1993. *Distributed Morphology*. In Keyser, J. (ed) The View from Building 20. Cambridge. MIT Press.
- Harley, Heidi & Ritter. 2000. "Person and number in pronouns: a feature-geometric analysis", ms, University of Arizona
- Menuzzi, Sérgio. 2000. *First Person Plural Anaphora in Brazilian Portuguese: chains and constraint interaction in binding*. In João

³ See also Costa (2001) and Costa and Duarte (2003) for other arguments against a characterization of null subject languages as languages in which Agr is pronominal.

- Costa (ed) Portuguese Syntax. New Comparative Studies. Oxford University Press.
- Pereira, Sandra. 2000. "*A gente*: acordos e desacordos". Ms, Universidade de Lisboa
- Pereira, Sandra. 2003. *Gramática compara de a gente. Variação no português europeu*. MA dissertation, Universidade de Lisboa
- Svenonius, Peter. 2003. "Phases at the interface". Paper presented at MIT Workshop on Phases and EPP, January, MIT