

The Portuguese Infinitive and the Nature of Linguistic Explanation

A. Carlos Quicoli

University of California, Los Angeles

1. INTRODUCTION

The central problem of linguistic theory is to explain, in scientific terms, how human beings develop Grammar. In the process of Grammar development (“language acquisition process”), children are exposed only to a small, finite number of utterances. Yet all over the world, children ‘instinctively’ develop a *recursive* Grammar that allows them to produce and understand an infinite number of sentences of their language. Now, this ‘instinct’ that guides children to develop recursive Grammars, rather than some other kind of system, shows that the human mind must have an innate notion of the *form* of Grammar; that is, human beings are biologically endowed with specific mental structures for developing Grammar. Or, to put it in psychological terms, the human mind-brain must possess an innate *language faculty*, which determines the general *form* of Grammars that human languages must have.

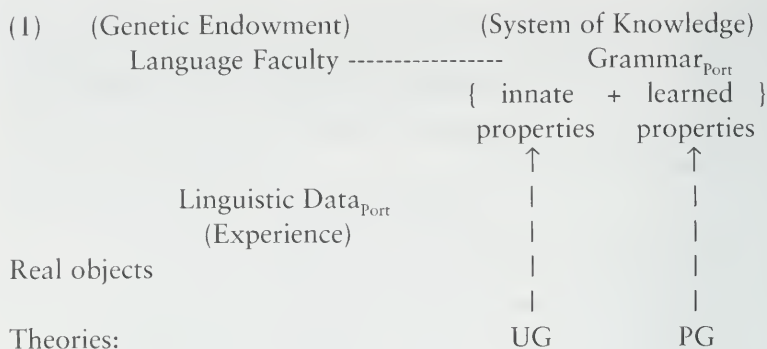
Thus, the development of Grammar must involve an essential *innate component*—which is commonly referred to (after Descartes) as the “language faculty.” At the same time, it is quite evident also that in order for the child to develop a Grammar appropriate to Portuguese, the child must be exposed to a sample of linguistic data of Portuguese. So experience with the data of a specific language is also necessary. Hence, in addition to the innate component, the development of Grammar must also involve a *learned component*, which is also essential, although it appears to play only a secondary role to that of the innate component in the process of Grammar development. These are the basic assumptions underlying much of current work in linguistics since Chomsky (1965).

More generally, the investigation of the contribution of the genetic endowment and of experience in the development of linguistic abilities

provides a framework for the investigation of the development of human cognitive abilities (i.e. “mind faculties”) in general, and can be regarded as a step toward understanding the nature of the human mind—a point that is summarized in the following statement by Nobel Prize biologist Salvador Luria:

From thinking of language as a dual entity consisting of a genetically determined component inscribed in the structure of the brain and a learned component derived from experience it is an easy step to a more general conception of the human mind [. . .]. To the biologist it makes eminent sense to think that, as for language structures, so also for logical structures there exist in the brain network some patterns of connection that are genetically determined and have been selected by evolution as effective instruments for dealing with the events of life [. . .]. Perfecting of these cerebral structures must have depended on their becoming progressively more useful in terms of reproductive success. For language this must have meant becoming a better instrument in formulation and communication of meaning through a usable grammar and syntax. (1973: 140-1; qtd. in Lightfoot 1982: 12)

Viewed in these terms, the study of Grammar constitutes an integral part of modern scientific investigation, and is best understood when cast in this context. The Grammar developed by the individual speaker represents a real object present in some form in the individual’s mind-brain. The linguist wants to *scientifically* understand, at the psychological level, the basic structural properties and the principles of organization of this real object. Within this general context, we can recast the problem of explaining Grammar development (i.e. “language acquisition”) as follows:



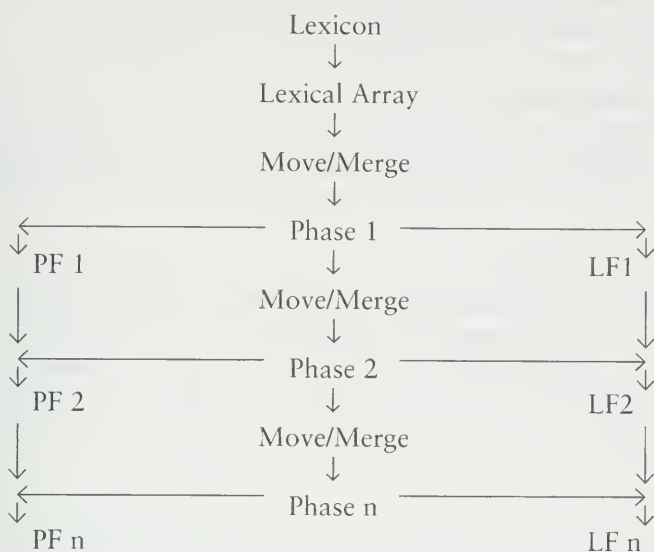
According to (1), the Grammar internalized by a Portuguese speaker can be regarded a dual entity that contains: a) innate properties, representing the contribution of the innate *language faculty*—i.e. the genetically determined component Grammar; and b) learned properties that reflect the contribution of the *learned component* of Grammar, which is derived from the speaker's experience with the linguistic data of Portuguese. Accordingly, we must develop two theories: a Universal Grammar (UG) to account for the innate properties of Grammar; and a Particular Grammar (PG) to account for the learned properties that reflect the language specific features that distinguish Portuguese from all other languages.

A linguistic explanation consists in demonstrating how the facts that are part of the Particular Grammar of Portuguese can be deduced from the principles of Universal Grammar. In the remaining of this article, we attempt to exemplify by a few concrete examples how this can be accomplished. We will examine a range of facts involving Portuguese inflected infinitives—clearly a language specific phenomenon. We then proceed to show how this range of facts of Portuguese inflected infinitives can be deduced from some specific principles that have been proposed as part of a substantive theory of Universal Grammar—the principles of the theories of Binding and Case.

2. THEORY OF UNIVERSAL GRAMMAR: BINDING AND CASE

While there is no consensus among linguist about the exact form of Grammar that human languages may have, there is some evidence from recent studies related to the concept of “syntactic phase” (Chomsky 2001, and related work) that the form of grammar determined by UG has the following general design:

(2)



The basic idea of the Phase model above is that the rules of Syntax (Merge and Move) organize a set of lexical items (Lexical Array) into clause-like structures (“Phases”), which are then processed in parallel by the phonological component (“Phonetic Form,” or PF-component) and by the semantic component (“Logical Form,” or LF-component), as each Phase is completed by the syntax and submitted to the two interpretive components. In other words, the Phonetic Form or PF (i.e. “sound representation”) of a sentence, and the Logical Form or LF (i.e. “meaning representation”) of a sentence are built incrementally “phase by phase,” as the syntax completes each Phase, and submits it to the Phonology and to the Semantics. The Phase-level substitutes with advantage the two previous levels of D-Structure and S-Structure of earlier theories (cf. Chomsky 2001; Quicoli 2002; 2005).

In addition, each component of the Grammar in (2), and the representations that they produce are regulated by general principles that are assumed to be part of UG. The main principles of UG that are of concern here are the principles of Binding Theory (Chomsky 1981, and related work), and the principles of Vergnaud’s Case Theory (cf. Chomsky and Lasnik 1977).

The principles of Binding Theory regulate the range of anaphoric relations among nominal phrases (NP’s) in a sentence. The standard

formulation of this theory (Chomsky 1981, 1982, 1986, 1995; Chomsky and Lasnik 1993) is summarized in (3) and (4) below:

(3) Binding Theory¹

Principle A: An anaphor must be bound in a Local Domain

Principle B: A pronominal must be free in a Local Domain

Principle C: An R-expression is free.

(4) Local Domain

β is a Local Domain for α , if and only if, β is the minimal category containing α and a SUBJECT accessible to α .

(SUBJECT is: a) an NP in subject position; b) the Agreement ('AGR') element of inflected verbs)

The Binding principle that is of more relevance here is Principle A. Essentially, 'anaphors' are elements such as English "himself" and Portuguese reflexive pronouns such as "se," which do not have reference of their own and, hence, must be associated with a referential NP (i.e. an 'antecedent') in order to be semantically interpreted. The effects of Principle A can be illustrated by the grammatical contrasts observed in (5):

- (5) a. *Nós nos barbeamos.* ('We shaved ourselves.')
- b. **Nós se barbeamos.* (*'We shaved himself.')

In (5a) the reflexive anaphor *nos* 'ourselves' is interpreted as coreferential with the subject NP—*nós* 'we.' Thus, the anaphor is "bound" as required by Principle A. Since no principle (or "law") is violated, the resulting sentence is predicted to be well-formed. By contrast, in (5b), the anaphor *se* 'himself' cannot be interpreted as coreferential with *nós* 'we' because of number mismatch. Hence, the anaphor in (5b) is not bound, which is a violation of Principle A. Since a grammatical principle (or "law") is violated, the ungrammaticality of (5b) is predicted.

However, anaphors cannot be bound by an antecedent just anywhere in the sentence. Rather, they must be bound within the Local Domain in which they occur. The Local Domain for an anaphor is essentially the minimal clause containing the anaphor and a

SUBJECT—i.e. an explicit syntactic Subject, or the Agreement morphology ('AGR') expressing the syntactic subject (e.g. the agreement marker *-mos* in *cantamos* 'SING-1ST PL.,' which acts as a "proxy" for the subject it expresses). The role of Local Domain is illustrated below:

- (6) a. Pedro viu [nós nos barbearnos].
 b. *Pedro viu [nós se barbearnos].

In (6), the Local Domain for the anaphor is the embedded clause, which contains a syntactic subject. In the well-formed (6a) the reflexive anaphor *nos* 'ourselves' is bound by its antecedent *nós* 'we' within its Local Domain, as required by Principle A. However, in the ill-formed (6b) the anaphor *se* is "free" within its Local Domain. It cannot be bound by *nós* 'we' inside the Local domain, and *Pedro* is outside the Local Domain. Thus (6b) is in violation of Principle A, which explains why it is ungrammatical.

Moreover, Binding Theory is not restricted to "lexical" anaphors such as the Portuguese or English reflexives. It is known that "traces" of certain moved items function like anaphors. Such "trace-anaphors" pattern like "lexical anaphors," and their distribution can also be explained by Principle A of Binding Theory. Thus consider the English facts below:

- (7) a. Joe believes [himself to be rich].
 b. *Joe believes [himself is-AGR rich].
 (8) a. Joe seems [t to like Los Angeles].
 b. *Joe seems [that t likes-AGR Los Angeles].

As we can see, the trace "t" left by NP-movement in (7) behaves in the same way as the lexical anaphor *himself* in (7). In the well-formed (7a) and (8a), the lexical anaphor *himself* and the trace-anaphor "t" are bound in their respective Local Domains (the full sentence structure). Since these structures are in compliance with Principle A, the resulting sentences (7a) and (8a) are grammatical, as predicted. However, in (7b) and in (8b) the lexical anaphor and the trace-anaphor are not bound in their respective embedded clauses—each a Local Domain created by the AGR element of the inflected verb. Since this violates Principle A, the resulting sentences (7b) and (8b) are ill-formed, as predicted.

This brings us to the problem of determining the precise level of representation at which Binding principles apply. Consider in this regard the English examples below:

- (9) a. The ambassadors appeared to each other [*t* to contradict the secretary].
 b. The ambassadors appeared [*t* to contradict each other].
 c. The ambassadors appeared to each other [*t* to contradict themselves].

All three examples in (9) involve application of NP-movement (an instance of “Move”) that raises *the ambassadors* to the main clause, leaving a trace in its original position. One might think based on (9a), that Principle A must apply after the NP-movement (i.e. at the S-structure level). However, the evidence of (9b) may suggest that Principle A must apply before NP-movement (i.e. at the level of D-structure), while (9c) seems to suggest that Principle A must apply both before and after NP-movement—i.e. both at D-structure and S-structure, a seemingly paradoxical result. Under previous theories, such facts were problematic. However, none of these problems arise if we assume the Phase Model (2), and that Principle A applies at the Phase-level. The derivation of (9c) by “phase” is as follows:

- (10) a. [The ambassadors_i to contradict themselves_i].
 b. [The ambassadors_i appeared to each other_i [*t*_i to contradict themselves_i]].

The Merge rules of the Syntax produce (10a). Since this is a Phase, it is submitted to the semantics (LF-component), where Binding applies. Application of Principle A at this Phase binds the anaphor *themselves* to *the ambassadors*. The structure goes into its second syntactic Phase. Move (i.e. NP-movement) moves *the ambassadors* to the main clause, and Merge embeds (10a) to form (10b). The second Phase is complete and (10b) is submitted to the semantics (LF-component). Principle A applies in the second Phase and binds the anaphor *each other* to *the ambassador* (which also binds the trace-anaphor “*t*”). No principle is violated and the result is the grammatical (9c). A similar analysis can be extended in a straightforward manner to (9a) and (9b). There is no need for D-structure or S-structure, which appear to be reflexes

of the more basic level of the Phase—application of Principle A in the first phase gives the impression of D-structure application, while application of Principle A in the second phase gives the impression of S-structure application (Quicoli 2002; 2005).

Similarly, the Phase hypothesis also resolves a problem noted with respect to the interpretation of “trace-anaphors” in examples such as (11):

- (11) a. **John** is likely [t to win].
 b. [_{CP} [How likely t to win] is [John ____]]?

The problem here is that normally an anaphor must be “lower”—i.e. “c-commanded” by its antecedent for the sentence to be grammatical (cf. **himself shaved John*). This happens in (11a), where the trace-anaphor in the lower clause is “c-commanded” by its antecedent *John*. However, in (11b), after NP-movement has applied, leaving a trace-anaphor in the embedded clause, Wh-movement moves the adjectival phrase and the subordinate clause with the trace-anaphor to the CP-position of the main clause, so that the trace-anaphor is actually ‘higher’ than its antecedent. Yet, surprisingly (11b) is also well-formed—a serious problem for previous theories, since neither application of Principle A at D-structure or at S-structure (or at LF-structure) seem to plausibly account for such facts.

However, the problem posed by (11b) can be resolved if we assume that Principle A applies at the Phase level. The derivation of (11b) is then as follows:

- (12) a. [_{vP} John to win] Phase 1 (vP Phase): Binding (inapplicable)
 b. [_{vP} **John**_i is [_{AP} how likely [t_i to win]]] Phase 2 (vP Phase)
 = Binding
 c. [_{CP} [_{AP} how likely [t_i to win] [is]] [John_i ____ [_{AP} ____]]]
 Phase 3 (CP phase)

In the first vP phase (see Chomsky 2001; Quicoli 2005), Principle A is not applicable. But in the second vP phase (12b) after Move/Merge, Principle A applies (after NP-movement) and correctly binds the trace-anaphor to *John*. In the third phase (12c), a CP-phase, Wh-movement moves the adjectival expression (AP) into the Spec-CP position, while Aux-movement moves the auxiliary *is* into the Head-position of CP.

However, since the trace-anaphor was already bound to an antecedent (“John”) in a previous Phase, structure (12c) is in compliance with Principle A, which explains the grammaticality of (11b). Thus, we have now two pieces of evidence to show that Binding Theory must apply at the Phase level, which supports the Phase Model given in (2).

Let us now turn to Vergnaud’s Case Theory (cf. Chomsky and Lasnik 1977). According to this theory, certain NPs must be marked for (abstract) Case. This requirement is guaranteed by the Case Filter, which requires that NPs that have “phonetic content” (i.e. are pronounced) must have Case. However, there is evidence that the original Case Filter must be extended to include also the “phonetically null” pronominal *pro*² (Quicoli 1996). Accordingly, we revise the Case Filter as follows:

(13) Extended Case Filter (ECF)

Noun Phrases containing *personal features* must have Case.

The class of NP’s containing ‘personal features’ (e.g. first person, second person, etc.) includes “Referential expressions” (e.g. *Maria, os médicos* ‘the doctors,’ etc.), overt pronominals (e.g. *ele* ‘he,’ *me* ‘me’), and their corresponding null counterparts (i.e. the null pronominals represented by “little *pro*”). The Extended Case Filter (ECF) interacts with the principles of Case Marking, which assign Abstract Case (i.e. Case which may or may not be overtly expressed by the noun morphology). With a great deal of simplification we may assume the following rules of Case Marking:

(14) Case Marking³:

Rule 1: NP head-governed by AGR is marked Nominative.

Rule 2: NP head-governed by a Verb is marked Accusative.

Rule 3: NP head-governed by a Preposition is marked Oblique.

In simple words, the principles as stated in (14) will ensure that an NP in a Subject-Verb relation with an inflected verb (i.e. with the element AGR) is head-governed by AGR, and receives Nominative Case. On the other hand, an NP inside a VP-node is head-governed by the Verb and it is assigned Accusative Case, while an NP inside a PP-node is head-governed by the Preposition and is, thus, assigned Oblique Case.

Restricting the discussion to essentials, normally the subject position of a finite verb will be marked Nominative by the AGR element of the verb. But the subject position of a regular infinitive (i.e. the “non-inflected infinitive” of most languages) normally cannot receive any Case at all—which raises the potential for the structure to be found in violation of the ECF. This can be best illustrated by the facts of languages such as Spanish/English, which have only non-inflected infinitives, compare:

- (15) a. Este muchacho-NOM parece [t odiar Las Vegas].
 b. This guy-NOM seems-AGR [t to hate Las Vegas].
 (16) a. *Parece [este muchacho odiar Las Vegas].
 b. * It-NOM seems-AGR [this guy to hate Las Vegas].

In the grammatical (15), the (non-inflected) infinitive cannot give Case to *this guyleste muchacho*. However, since NP movement raised the embedded subject to the subject position under *seems/parece*, the raised NP receives Case from the AGR element of the main verb, so as to satisfy the ECF. However, in the ill-formed (16), *this guyleste muchacho* remained in subject position of the infinitive. Since in English/Spanish, infinitives do not have AGR, *this guyleste muchacho* are not Case marked. But this is a violation of the ECF, so the examples in (16) are ill-formed, as predicted by the ECF.

3. SOME CONSEQUENCES OF CASE THEORY

In this section we examine some predictions of the Extended Case Filter (ECF), the central principle of Case Theory, for the data of Portuguese infinitives. In Portuguese, infinitives may occur without AGR (non-inflected infinitive), or with AGR (inflected infinitive). Thus, Case Theory predicts that when the inflected infinitive occurs, its AGR element should give Case to its subject, allowing the structure to satisfy ECF, which should result in a different pattern than that found in languages that only have non-inflected infinitives. The first pattern to consider in this regard is represented by cases where the infinitive has an overt NP as its subject, and the predicate of the main clause cannot assign Case to it. In such cases, the inflected infinitive must occur:

- (17) a. É bom os estudantes/eles estudarem latin.
 b. *É bom os estudantes/eles estudar latin.

But in Spanish, no grammatical sentence is possible:

(18) *Es bueno los estudiantes estudiar latín.

The facts are as predicted by Case Theory. In (17a) the AGR of the inflected infinitive gives Case to its subject. Since the ECF is satisfied, (17a) is predictably grammatical. However, in (17b) (like Spanish (18), the infinitive has no AGR, and the structure is ruled out by the ECF, just like Spanish (18)).

To account for such facts, many traditional pedagogical, and descriptive grammars give an informal rule, or descriptive generalization, like (19)⁴:

(19) The inflected infinitive is used “whenever the infinitive is accompanied by a nominative subject, noun or pronoun.”
(Ali 1964: 175)

The rule in (19) would account for the differences in grammaticality observed in (17). In fact, one fruitful way to advance linguistic theory is to take descriptive statements like (19) seriously, and then attempt to derive their effects from independently motivated principles of UG. Thus, given (19), we may ask a further question: “Why must the inflected infinitive occur when it is “accompanied by a nominative subject”?” Someone interested in language teaching might simply say: “Because that is the way it is,” or something to this effect. This may be fine, if the goal is the practical teaching of the language. But from the point of view of understanding how human language works—the goal of linguistics as a scientific discipline—we must strive to find principled explanations for the facts. Ideally, the facts of Portuguese—even the facts of an idiosyncratic construction as the inflected infinitive—ought to be explained by general principles of UG. In this case, as shown above, the generalization (19) follows from the ECF—an independently motivated principle of UG, a desirable result.

Consider now a second pattern, represented by examples where the infinitive has an overt NP subject and the main clause contains an ECM Verb (‘exceptional case-marking’ verbs such as *ver* ‘to see,’ *ouvir* ‘to hear,’ and *causatives* such as *fazer* ‘to make/to cause,’ *mandar* ‘to order,’ *deixar* ‘to allow/to let’). With such verbs, either the non-inflected, or the inflected infinitive may occur (cf. Maurer 1968: 239):

- (20) a. Lula via os problemas crescer à sua volta.
 b. Lula via [os problemas-ACC crescer à sua volta].
 ‘Lula saw the problems grow around him.’
- (21) a. Lula via os problemas crescerem à sua volta.
 b. Lula via [os problemas-NOM crescerem-AGR à sua volta].
 ‘Lula saw the problems grow-3RD PL. around him.’

These results are as predicted. They follow from the way the ECF interacts with ECM verbs and with inflected infinitives. Accordingly, there are in Portuguese two ways in which the embedded subject *os problemas* ‘the problems’ may be Case-marked to satisfy the ECF. If the infinitive is inflected, its AGR will assign Nominative Case to its subject. This would satisfy the ECF (and, at the same time, block Case assignment by the main verb, since inflected infinitive is the “closest” Case-marker), so that (21a) can be derived. Alternatively, in (20a) the infinitive is not inflected, so it cannot assign Case to its subject, but the ECM Verb in the main clause can. So it will assign Accusative Case to *os problemas* ‘the problems.’ This would allow the structure to also satisfy the ECF (just like in its Spanish counterpart), so that (21a) is also grammatical, as expected.

As a third situation, consider now the problem posed by the facts in (22):

- (22) a. É melhor dizer sempre a verdade
 (‘It is better to always tell the truth.’)
 b. É melhor dizermos sempre a verdade.
 (‘It is better to always tell-1ST PL. the truth.’)

The Portuguese sentences in (22) are not synonymous (cf. also Maurer 1968: 148). In sentence (22a), the covert subject of the infinitive is an impersonal, or unspecified human subject, with no specific reference—(22a) means something like “It is better for people to tell the truth,” (just like in Spanish and English). However, in (22b) the inflected infinitive refers to a null, personal subject, corresponding to *nós* “we;” so the sentence means “It is better for us to always tell the truth.”

Traditional grammars state that in such cases “the infinitive will agree with the subject that we have in mind.” (Ali 1964: 175). Again, this may be justified in a pedagogical grammar. However, in a formal

grammar the goal is precisely to make explicit what kind of “subject we have in mind” that causes the infinitive to behave differently in such examples. That is, under a formal (i.e. “generative”) approach, it is necessary to provide theoretical assumptions about the types of null subjects that the speaker “has in mind,” and to show how the different choice of subjects correlate with the occurrence of the inflected vs. the non-inflected infinitive in such examples.

Under standard generative analyses, the ‘unspecified subject’ that occurs in Portuguese examples such as (22a) (and its counterpart in other languages) is theoretically represented by *PRO-*arb** (Chomsky 1981)—a phonetically null element that is “arbitrary” in reference in the sense that it does not refer to any specific individual. In the theory advanced here, *PRO* is ‘non-personal.’ That is, *PRO* is assumed to contain only features that identify it as an animate *pro*-element. But unlike personal pronouns, *PRO* does not contain ‘personal features’ (e.g. first person, second person, etc) (cf. Quicoli 1996, but see Safir 1996 for discussion). Hence, *PRO* is not subject to the ECF. In fact, *PRO* cannot occur in a Local Domain (which implies that it cannot occur in a Case-marked position) for independent reasons, having to do with Chomsky’s (1981) ‘PRO-Theorem.’⁵ So the occurrence of *PRO* is strictly limited to the subject position of (non-inflected) infinitives—the only position that is normally not a Local Domain. Thus, the structure underlying (22a) is essentially (23), where the subject of the non-inflected infinitive is *PRO-*arb**:

(23) É melhor [PRO-*arb* – dizer sempre a verdade].

By contrast, the pronominal ‘little *pro*’ is simply the phonetically null variant of a personal pronoun. Therefore, it has features for person and, hence, it requires Case in order to satisfy the ECF (just like overt pronouns). Thus, when *pro* occurs, the infinitive must be inflected to give Case to it. Thus, the structure underlying (22b) is (24):

(24) É melhor [(*pro*-1st PL.)-NOM dizermos-AGR sempre a verdade].

Under these assumptions, we can provide a principled explanation for the facts above in terms of the ECF. When the subject is *pro*, the infinitive must be inflected in order to assign Nominative Case to

it (just like in (17a) with an overt pronoun). By contrast, when the subject is *PRO*-arb, the infinitive must always be non-inflected. First, because *PRO*'s are not 'personal' and, hence, are not subject to the ECF. Second, because *PRO*'s cannot occur in a Case-marked position due to Chomsky's (1981) "PRO-Theorem."

The explanation above can be extended also to infinitives embedded under "control verbs"—i.e. structures where the subject of the infinitive is a 'controlled' *PRO* which is obligatorily interpreted as coreferential with a NP in the main clause. This is the case, for example, of infinitival clauses embedded under verbs such as *preferir* 'to prefer,' *tentar* 'to try,' among others. In such structures, only the common non-inflected infinitive is possible (cf. Cegalla 2000: 551). This is evidenced by the examples in (25), which are associated with their respective derivation in (26):

- (25) a. Os prisioneiros tentaram escapar.
 b. *Os prisioneiros tentaram escaparem.
 'The prisoners tried to escape.'
- (26) a. Os prisioneiros-NOM_i tentaram-AGR [*PRO*_i escapar].
 b. Os prisioneiros-NOM_i tentaram-AGR [(*PRO*-NOM)_i escaparem-AGR].

As is clear from the above, only sentence (25a) with structure (26a), which has a non-inflected infinitive, is well-formed. Sentence (25b), associated with (26b), which has an inflected infinitive, is ill-formed. This is as expected. Under the present analysis, *PRO* does not have 'personal features' of its own, and it is exempt from the ECF. Thus, no principle is violated in (25a), which is a well-formed sentence. However, in the derivation of sentence (25b) the infinitive is inflected, causing the controlled *PRO* to be in a Local Domain. But *PRO* cannot be in a Local Domain because of the PRO-Theorem. Hence (25b) is excluded by Binding Theory, under the assumptions related to the PRO-Theorem, which explains its ungrammaticality.

Consider now the problem posed by the class of "semi-control" verbs (Quicoli 1996). Such verbs allow two constructions with infinitives: one in which the subject of the infinitive is a controlled *PRO*, and another in which the subject of the infinitive is a 'personal' noun phrase. This is the case for example of verbs like *afirmar* 'to assert,' *dizer* 'to say,' *crer* 'to believe.' Compare:

- (27) a. Os guardas afirmam **terem** eles visto o ladrão.
 b. Os guardas afirmam **terem** visto o ladrão.
 c. Os guardas afirmam **ter** visto o ladrão.
 'The guards assert (they) saw the thief.'

According to the present analysis, their respective underlying structures are as follows:

- (28) a. Os guardas afirmam [eles-NOM **terem**-AGR visto o ladrão].
 b. Os guardas afirmam [(pro-3RD PL.)-NOM **terem**-AGR visto o ladrão].
 c. Os guardas afirmam [PRO **ter** visto o ladrão].

As shown in structures (28a) and (28b), sentences (27a) and (27b) contain "personal" subjects—the pronominal elements *eles* 'they,' and 'little *pro*,' respectively. This is clear since the subject of the infinitive in both instances is 'free' to refer to *os guardas* 'the guards,' or to another individual understood in the discourse, as is typical of pronominals. Thus, they are 'personal' NPs, and they need Case. Since the infinitive is inflected in these structures, it assigns Case to their respective subjects. This satisfies the ECF, so that the grammaticality of (27a) and (27b) is explained. However, as shown in structure (28c), in (27c), the subject is a controlled *PRO*, which cannot be in a Local Domain (because of the PRO-Theorem) and, hence, cannot occur with an inflected verb. Since in (28c) the infinitive is not inflected, the embedded *PRO* is not in a Local Domain. Therefore, neither Principle A nor the ECF is violated, and (27c) can surface as a grammatical sentence. Thus, as we can see, the facts in (28) follow from both the theories of Case and Binding.

4. SOME CONSEQUENCES OF BINDING THEORY

In this section, we examine some of the predictions of Binding Theory for the data pertaining to the Portuguese inflected vs. non-inflected infinitive. The Binding effects that are of relevance here are those involving 'trace-anaphors' that are created by Movement rules. Essentially, such trace-anaphors are subject to Principle A of Binding Theory, and must be 'bound' by an antecedent within the Local Domain in which they occur. Since AGR creates Local Domains, the theory predicts that inflected infinitives (unlike non-inflected infinitives) should

create Local Domains. We should then expect different results related to movement of the subject of an inflected infinitive vs. movement of the subject of a non-inflected infinitive. We shall examine this question with respect to four types of movement rules: Clitic movement, NP-movement, Wh-movement, and Topicalization.

Consider first Clitic movement. A well-known fact about Portuguese infinitives is that cliticization of the subject is possible with a non-inflected infinitive but not with an inflected infinitive. The following example is typical:

- (29) a. Não nos deixeis cair em tentação.
 b. *Não nos deixeis cairmos em tentação.
 ‘Do not let us fall into temptation.’

Because of this, many traditional grammars (Cegalla 2000: 551; Lima 1972: 382) must assume a specific ‘rule’ for cases where the subject of the infinitive is a clitic. Compare:

- (30) [I]f the infinitive has as its subject an oblique pronoun with which it constitutes the object of the verbs *deixar* ‘let,’ *fazer* ‘to make,’ *mandar* ‘to order,’ *ver* ‘to see,’ and *sentir* ‘to feel,’ it is not inflected. (Cegalla 2000:551)

Of course, a great deal of generality can be gained if the facts pertaining to clitic subjects did not require a special rule just for clitics such as (30). Ideally, such facts should, rather, be deduced from independently motivated principles. In fact, we see that the facts in (29) can be explained in a straightforward manner under Binding Theory. Thus, the respective structures corresponding to (29) are essentially as shown in (31):

- (31) a. [pro não nos_i deixeis [t_i cair em tentação]].
 b. [pro não nos_i deixeis [t_i cairmos-AGR em tentação]].

In both structures, Clitic movement left a trace-anaphor. In (31a), with the non-inflected infinitive, the Local Domain for the trace-anaphor movement is the whole structure. Since in this configuration the trace-anaphor is bound by the clitic *nos* ‘us,’ Principle A is satisfied. This explains the grammaticality of (29a). However, in (31b), the AGR

of the inflected infinitive narrows the Local Domain to the embedded clause. Since the trace-anaphor is not bound by an antecedent in this domain, the structure violates Principle A of Binding Theory. This explains the ungrammaticality of (29b). Thus, the Portuguese facts in (29) can be explained in terms of Principle A, a general principle of UG. No special statement like (30) is needed in the Particular Grammar of Portuguese—a significant theoretical result.

As a second set of data, consider now NP-movement. As shown below, NP movement is possible with non-inflected infinitives but not with inflected infinitives (cf. Maurer 1968: 109 fn; Quícoli 1976, 1982):

- (32) a. Os rapazes pareciam odiar o filme.
 b. [Os rapazes₁-NOM pareciam-AGR [t₁ odiar o filme]].
 ‘The guys seem to hate the movie.’
- (33) a. *Os rapazes pareciam odiarem o filme.
 b. [Os rapazes₁-NOM pareciam-AGR [t₁ odiarem-AGR]].
 ‘The guys seem-agr to hate-AGR the movie.’

These facts can also be explained in a straightforward manner by Principle A of Binding Theory. In both instance, the embedded subject *os rapazes* ‘the guys’ is raised by NP-movement to the position of subject under *parecer*. This movement leaves a trace-anaphor, as before. In structure (32b), the infinitive is not inflected, so the Local Domain for the trace-anaphor is the whole structure. Since the trace is bound by *os rapazes* in this domain, the structure satisfies Principle A, and the resulting sentence (32a) is predictably grammatical. In structure (33b), however, the presence of the inflected infinitive narrows the Local Domain to the subordinate clause. Since the trace-anaphor is not bound by an antecedent in this Local Domain, the structure violates Principle A, which explains why the resulting sentence (33a) is ungrammatical. Also, here we have a situation where the inflected infinitive gives Case to its subject, so that the derivation of (33a) satisfies Case Theory. However, the inflected infinitive also creates Local Domains, and since the movement of the subject in (33a) left a trace-anaphor ‘free’ in the Local Domain created by the inflected infinitive, the structure violates Binding Theory, which correctly excludes the sentence as ungrammatical. In other words, both Case Theory and Binding Theory must be satisfied if the sentence is to be grammatically well-formed.

As a third set, let us now examine the pattern involving Wh-movement. We see now that when the subject is moved by Wh-movement the pattern is just the opposite of that when the subject is moved by NP movement discussed above. Now the sentences with inflected infinitive are well-formed, while the sentences with non-inflected infinitive are ill-formed (cf. Raposo 1987; Quicoli 1996):

- (34) a. Que jogadores ele afirma terem abandonado o time?
 b. [_{CP2} Que jogadores₁-NOM [ele afirma [_{CP1} t₁ [t₁ AGR-terem abandonado o time]]]]?
 ‘Which players does he affirm (that they) have abandoned the team?’
- (35) a. *Que jogadores ele afirma ter abandonado o time?
 b. * [_{CP2} Que jogadores₁ [ele afirma [_{CP1} t₁ [t₁ ter abandonado o time]]]]?
 ‘*Which players does he affirm to have abandoned the team?’

Wh-movement involves movement of a wh-phrase (i.e. an interrogative phrase or a relative pronoun) into a position under the CP-node—the phrasal category at the beginning of a clause. Thus, in both derivations (34b) and (35b) Wh-movement moved the wh-phrase *que jogadores* ‘which players’ first under the CP1 of the subordinate clause, and then to the CP2 of the main clause. In both structures, the trace of the moved subject is properly bound by the trace under CP1, and the trace under CP1, in turn, is bound by *que jogadores* ‘which players’ under CP2—so Binding Theory is satisfied. However, the difference here is due to Case Theory. In (34a) the infinitive is inflected, while in (35a) the infinitive is not inflected. In the grammatical (34a) the inflected infinitive gives Case to the moved subject, so that the Case Theory (i.e. ECF) is satisfied. However, in the ungrammatical (35a), the infinitive is not inflected and cannot give Case to its subject. Since Wh-movement moves the embedded subject into CP’s, which are not Case-marking positions either, the moved subject is not assigned Case. As a result, (35a) is in violation of Case Theory, which explains its ungrammaticality. Again, as we see, it is not sufficient for structures to satisfy only Binding Theory but not Case Theory. Rather, both Binding Theory and Case Theory must be satisfied for the result to be well-formed.

As a fourth, and final case, consider the pattern involving Topicalization. It is known since Chomsky (1977) that Topicalization behaves

much in the same way as Wh-movement. So we would expect Topicalization to display the same pattern with the Portuguese inflected vs. non-inflected infinitives as that observed with Wh-movement. That is, due the interaction of Binding Theory and Case Theory above, we would expect the resulting sentences to be grammatical with the inflected infinitive and ungrammatical with the non-inflected infinitive. As shown by the facts below, these predictions hold:

- (36) a. Estes jogadores, o técnico afirma terem chegado tarde.
 b. [_{CP2} Estes jogadores₁-NOM [o técnico afirma [_{CP1} t₁ [t₁ AGR-terem chegado tarde]]]].
 ‘These players, the coach asserts (that they) have arrived late.’
- (37) a. *Estes jogadores, o técnico afirma ter chegado tarde.
 b. * [_{CP2} Estes jogadores₁ [ele afirma [_{CP1} t₁ [t₁ ter chegado tarde]]]].
 ‘*These players, the coach asserts to have arrived late.’

Topicalization, like Wh-movement, moves the topicalized noun phrase into CP's. This would satisfy Binding Theory, since the traces left by Topicalization in both structures would be properly bound, as required by this theory. However, since CP is not a position where an NP can receive Case, the moved subject must receive Case elsewhere in the derivation. In the case of (36a) the infinitive is inflected and its AGR gives Case to its subject, so the resulting sentence satisfies also Case Theory. No principle is violated and the resulting sentence (36a) is grammatical. However, in (37a) the infinitive is non-inflected and cannot give Case to its subject. Since the subject of the infinitive is moved into CP by Topicalization, it cannot receive Case in this position either. As a result, structure (37a) is in violation of Case Theory, which explains its ungrammaticality. Thus, the pattern produced by Topicalization is identical to that produced by Wh-movement, and both can be explained from the interaction of the principles of the theories of Case and Binding.

5. CONCLUSION

In the course of the discussion above, we have examined a range of empirical facts related to the Portuguese inflected vs. non-inflected infinitives. We have argued that the facts of Portuguese infinitives—particularly the facts of the idiosyncratic inflected infinitives—can be

explained in a natural manner by some proposed principles of Universal Grammar—the principles of Binding Theory and Case Theory. If this were on the right tract, we would have here an illustration of how the facts of Particular Grammars can be deduced—and, hence, explained—from general principles of Universal Grammar. Thus, we would have a desirable interaction. The proposed principles of Universal Grammar can be relied upon to explain the particular facts of the Portuguese infinitive. At the same time, the explained facts of Portuguese infinitives can be taken as empirical evidence attesting to the general character of the proposed principles, further justifying their inclusion as part of a substantial theory of Universal Grammar. If these results could be confirmed, they would provide a concrete example of how the facts of particular languages can be deduced and, hence, explained by general principles of language and, at the same time, how the range of explained data of particular languages can provide empirical support in favor of the explanatory power of the linguistic theory that incorporates them.

Notes

1. Chomsky and Lasnik (1993) (reproduced in Chomsky 1995) re-stated the Binding principles as “interpretative rules.” However, both the standard version of Binding Theory given in the text, and the interpretative version of Binding given by Chomsky and Lasnik require the concept of Local Domain. So, for the facts discussed in this article, it does not seem to matter which version of Binding Theory we adopt, although they may yield different empirical results with respects to other facts. For presentation purposes, I have adopted the more familiar standard version.

2. This is the ‘little *pro*’ subject that appears in Portuguese sentences such as *cantamos* ‘(we) sing-1ST PL,’ which, according to current theory, is to be analyzed as [pro-1ST PL. –cantamos] (cf. Chomsky 1981; 1982).

3. The precise relation required for Case Assignment is somewhat unclear. I have assumed, for presentation purposes, that Case Assignment is assigned under the notion “head-government,” which involves the relation of “m-command” (Aun and Sportiche 1983; Chomsky 1986). But this is a controversial move (see Chomsky 1995 for a different approach to Case marking based on the concept of “Spec-Head Agreement”). The notion “m-command” says essentially that the “head,” or nucleus of a phrasal category

“m-commands” all categories inside the phrase category labeled after it. Thus, an NP in ‘direct object’ position in a Verb Phrase (VP) is m-commanded by the Verb, which is the “head” of the VP; an NP inside a Prepositional Phrase (PP) is m-commanded by the Preposition, which is the head of the PP. Likewise, as it is widely assumed (since Pollock 1989), inflectional morphemes such as Tense and Agreement are the heads of their own Phrasal Categories. Thus, an NP said to be ‘in subject position’ of a Verb containing the AGR element (i.e. the Verb-Agreement morphology) is assumed to be in the Specifier-position of an AGR Phrase, and hence, it would be m-commanded by AGR, which would assign Nominative Case to it by Rule 1 of (14).

4. A similar rule is given in Maurer (1968: 145): “[W]hen the infinitive has its own subject—explicit or not—it is always inflected.” See also Cegalla (2000: 552), Bechara (1968: 346), Lima 1972: 382) for similar statements.

5. The argument based on the PRO-Theorem is as follows. According to Chomsky (1981), *PRO* has the features [+pronominal/+anaphor]. Thus, if *PRO* occurred with an inflected infinitive, the AGR of the infinitival clause would constitute a Local Domain for it. But this would cause a conflict between Principles A and B of Binding Theory; since *PRO* is, in part, an anaphor, Principle A would require it be bound. However, since *PRO* is also a pronominal, the resulting structure would violate Principle B, which requires pronominals to be free. Conversely, if *PRO* is free, the structure would satisfy Principle B, but now it would violate Principle A, since the anaphor must be bound in a Local Domain. This conflict can only be avoided if *PRO* does not occur in a Local Domain. It follows then that *PRO* cannot occur with inflected infinitives, since inflected infinitives create Local Domains.

Works Cited

- Ali, Manuel Said. *Gramática secundária da língua portuguesa*. Rio de Janeiro: Livraria Acadêmica, 1964.
- Aoun, Yusef and Dominique Sportiche. “On the Formal Theory of Government.” *The Linguistic Review* 2.3 (1983): 211–36.
- Bechara, Evanildo. *Moderna gramática portuguesa*. São Paulo: Nacional, 1968.
- Cegalla, Domingos P. *Novíssima gramática da língua portuguesa*. São Paulo: Nacional, 2000.
- Chomsky, Noam. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT, 1965.

- . “Derivation by Phase.” *Ken Hale: A Life in Language*. Ed. Michael Kentowicz. Cambridge, MA: MIT, 2001.
- . *Knowledge of Language: Its Nature, Origin, and Use*. New York: Praeger, 1986.
- . *Lectures on Government and Binding*. Dordrecht, Netherlands: Foris, 1981.
- . “On Wh-Movement.” *Formal Syntax*. Ed. Peter W. Culicover, Thomas Wasow, and Adrian Akmajian. New York: Academic, 1997. 71–132.
- . *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge, MA: MIT, 1982.
- . *The Minimalist Program*. Cambridge, MA: MIT, 1995.
- Chomsky, Noam and Howard Lasnik. “Filters and Control.” *Linguistic Inquiry* 8 (1977): 425–504.
- . “The Theory of Principles and Parameters.” *Syntax: An International Handbook of Contemporary Research*. Ed. Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann. Vol. 1. Berlin: Walter de Gruyter. 506–69.
- Lightfoot, David. *The Language Lottery: Towards a Biology of Grammars*. Cambridge, MA: MIT, 1982.
- Luria, Salvador. *Life: The Unfinished Experiment*. New York: Scribner, 1973.
- Maurer Jr., Theodoro H. *O infinitivo flexionado português*. São Paulo: Nacional, 1968.
- Pollock, Jean-Yves. “Verb-Movement, UG and the Structure of IP.” *Linguistic Inquiry* 20 (1989): 365–424.
- Quicoli, A. Carlos. “Anaphora by Phase.” Unpublished article, Department of Spanish and Portuguese, UCLA. Los Angeles, 2002.
- . “A Phase Approach to Anaphoric Relations.” Unpublished article, Department of Spanish and Portuguese, UCLA. Los Angeles, 2005.
- . “Inflection and Parametric Variation: Portuguese vs. Spanish.” *Current Issues in Comparative Grammar*. Ed. Robert Freidin. Dordrecht, Netherlands: Kluwer, 1996. 46–80.
- . “On Portuguese Impersonal Verbs.” *Readings in Portuguese Linguistics*. Ed. Jürgen Schmidt-Radefeldt. Dordrecht, Netherlands: North Holland, 1976. 63–91.
- . *The Structure of Complementation*. Ghent, Belgium: Story-Scientia, 1982.

- Raposo, Eduardo. "Case Theory and Infl-to-Comp: The Inflected Infinitive in European Portuguese." *Linguistic Inquiry* 18 (1987): 85-100.
- Rocha Lima, Carlos E. *Gramática normativa da língua portuguesa*. Rio de Janeiro: José Olympio, 1972.
- Rouveret, Alain. "Sur la notion de proposition finie: gouvernement et inversion." *Recherches Linguistiques* 9 (1980): 76-140.
- Safir, Ken. "PRO and pro: Comments on Quicoli." *Current Issues in Comparative Grammar*. Ed. Robert Freidin. Dordrecht, Netherlands: Kluwer, 1996. 81-90.