

Variant patterns of Subject-Verb agreement in Portuguese: morphological and phonological issues

MARIA ANTÓNIA MOTA

Abstract

In this paper, we start from a basic assumption: agreement is a most relevant operation for the Portuguese language speakers as a whole. The data from different varieties of Portuguese support this view, both in the Subject-Verb domain and within the NP, as well as the fact that speakers don't prefer the simplest solutions: the most productive plural agreement forms are the most complex ones, morphologically and phonologically. We focus the discussion on the interplay between different linguistic factors intervening in and promoting variant overt and covert patterns of agreement (the existence of agreement and the presence of visible marks, as expected in standard canonical schemes, constitute two separate questions). More specifically, we claim that it is worth to take into account the matching between the morphological and the phonological properties and features of the verb cells, in order to fully understand the attested variant outputs and the variant patterns of subject-verb agreement. We assume that the 3rd plural person-number marker has to be described in morphophonological terms, and we conclude that agreement is a morpho-phono-syntactic process, sensitive to the lexical-semantic features and discourse properties of the controller.

Introduction

This paper aims to explore the role of morphological and phonological phenomena in the overt vs. covert agreement attested in Portuguese, and the interplay of these grammatical components in the architecture of the different agreement patterns. We present a condensed survey of the morphological-

-phonological information relevant to agreement and give some tentative explanations for variant patterns of agreement based on that linguistic specific kind of information. This reflexion, when crossed with (morpho)syntactic analysis of the agreement operation, should be helpful to reach an integrated evaluation of the relative weight of the different components of grammar intervening in variant agreement patterns, and, more specifically, of the relevance and limits of concrete features associated with different components. The (non)neutralization of the contrasts between inflected forms in the Subject-Verb domain is our main object of questioning.

The line of reasoning proposed throughout this paper is the result of previous studies, some of them published or presented by the author to international meetings, and developed within the *Comparative study of agreement patterns in African, Brazilian and European varieties* Project.

For our purpose, we use the spoken data of Portugal (EP), Brazil (BP), and Africa (AP) collected for the Project (see more details in the Introduction of this volume), as well as data gathered in the *Atlas Linguístico e Etnográfico de Portugal e da Galiza*, ALEPG (non-standard varieties of European Portuguese), in two corpora of Cape-Verdian and Angolan Portuguese varieties collected by post-graduate students for their doctoral thesis (Amália Melo-Lopes and Afonso Miguel, respectively, to whom we address our special thanks), and, finally, in some interviews collected in little towns around Lisbon by our students of Sociolinguistics, at the University of Lisbon. Those data are used to support our reflexion, but their analysis won't be presented in detail; our specific goal is to put some hypothesis based on the available data.

As shown in other papers in this volume, a set of internal and external independent variables concur to the lack of canonical agreement in a number of sub-varieties of Portuguese. Our particular insight into variant schemes of agreement is thus to be combined with other contributions, in order to bring some light to the non-uniform interpretation of the agreement relations by speakers of Portuguese. Note that we use *canonical agreement* for the overt agreement, the one revealing the full realization of all the relevant features. Corbett (2006) puts that in canonical instances of agreement (the clearest ones), the target shows values compatible with the values of the formal and/or semantic features of the controller; they co-vary, their features specifications match, they share the same number of features. The canonical instances of agreement «fix a point from which occurring phenomena can be calibrated.» (*op. cit.*, p. 9). It is helpful to retain the notions of *compatibility*, *co-variance* and *reference point* associated to *canonical*. The concept of *canonical agreement* is to be distinguished from that of *standard agreement*, since speakers of the standard varieties produce linguistically motivated variant patterns of agreement, which are not canonical. Variant patterns are fully attested in Portuguese as a whole, including European standard

Portuguese, where the covert agreement is rare (see others papers, in this volume). Thus, standard agreement includes grammatical-induced variation.

Finally: even though we concentrate on the verbal morphological/phonological features for agreement, we will mention the NP agreement, on trying to capture possible generalizations on the Number assignment strategies. The reason for that excursus is that, among several other factors studied by other members of the Project, we believe that a parallel with verbal morphological and phonological comportment may be judicious.

Before going on, it is worth saying that the investigation developed within the Project reveals that the degree of covert agreement tends actually to diminish, due to the increase of the level of instruction, a most important extra-linguistic factor (see, for instance, Brandão, in this volume). However, covert agreement is still quite widespread, namely in BP and in AP.

The next sections lay out some basic topics concerning Number and Person categories (sections 1, 1.1) in the verb and their formal realization. Afterwards, we discuss the possible origin of the non-converging interpretations of those categories by the speakers (sections 2, 2.1), and, for doing so, we dress some questions concerning (i) the particular nature of the 3rd plural marker of person-number (PN), as opposed to the 3rd singular marker, (ii) the interplay between the morphological and the phonological structure of the 3rd pl. (P6) – 3rd sing. (P3) forms, (iii) the way speakers match their properties and features and the consequently predictable phonetic properties outputs. Expanding this discussion, (iv) we draw a picture of the morphophonological results of the adjunction of agreement suffixes. Additional comments are made about (v) other verbal forms than the 3rd, which may not be involved in person-number paradigmatic oppositions but which are pertinent to our analysis and have clear consequences on agreement. Next, we describe the effects of the person-number inflection on verbs (section 2.2) and of the number inflection on nouns (section 2.3), in order to compare the relevance of morphological and phonological properties in the verbal and nominal plural forms; on doing so, we try a brief co-related interpretation and present a tentative picture of the possible combinations, under the form of an agreement scale. In section 3, we present a few final remarks.

1. The linguistic categories involved in the Subject-Verb agreement

Our basic assumption is that agreement is an operation which is relevant in the grammar of the Portuguese language speakers. The data support this view, both within the NP and in the subject-verb domain; they also show that the degree and the typology of the (non)canonical agreement depend namely on the fact that speakers have Portuguese as L1 or L2. In the first case, the type of transmission, regular vs. irregular is a relevant factor (see Lucchesi (2003), Baxter (2004), Lucchesi & Baxter (2009) for the BP subvarieties

situation); in the second case, a typical irregular transmission case (see Lopes & Baxter, 2011), the maintenance of the contact with other L1 languages vs. the diminishing or loss of the contact seems to be identically important. Therefore, the focus of the discussion seems not to be the absence of the agreement operation, in the grammars of some varieties of Portuguese, but the interplay between different level linguistic factors intervening in and promoting variant patterns of agreement. The existence of agreement and the presence of visible marks, as expected in standard canonical schemes, constitute two separate questions.

As demonstrated by current theory and by experimental studies concerning other languages, for an agreement operation to be done, speakers scan the properties of the controller, which shows different types of inherent consistency, and they project those properties into the verb, through the harmonization of the specific features of the controller and those of the target. If the controller is a NP or a referential pronoun argument (null or fully expressed) – and not a weak subject, like the non nominalized infinitives –, the status of subject-verb agreement is robust, no matter the output form the person-number markers take; agreement is a constant, the different way of marking it being the result of a combination of factors, among which are the morphological and phonological ones.

Agreement is known to be an asymmetric feature-sharing relation (Corbett, 1998a), involving properties and features of different levels, as well as suffixes with different values and forms. Agreement is triggered by syntax, but its overt canonical realization and its variant outputs (still canonical or non canonical) depend, in the first place, on the morphological possibilities of the grammar to support the oppositions related to the relevant categories in the subject-verb domain. As stated above, the subject element and the verb have to share features, even though in different ways, depending on their inner, lexically determined, categories: «Agreement morphology is the ‘prerequisite’ for agreement. If the target does not have the means to realize the agreement features, then we have no evidence for agreement» (Corbett (2006:78). In general, Portuguese lexical classes that are relevant to agreement have the possibility to mark agreement features through inflectional suffixes overtly realized. But, concerning the NP domain of agreement, some athematic nouns and adjectives ending in *-s* in their singular morphological form don't show, phonetically, the plural mark (e.g. *vírus* +/-sing. ('virus') or *simples* +/-sing. ('simple')); in the subject-verb domain, part of the verbal paradigm cells don't dispose of an overt agreement suffix (e.g. *chega*, 3rd sing (he/she arrives); see below). It is then imperative to analyse the inflectional (agreement) suffixes of the controller and of the target to understand the different kinds of possible visible realization of the agreement features (Corbett (1998b) presents an interesting survey of questions concerning the agreement morphology in several languages).

As shown below (Table 2), the inflectional paradigms may be fully filled with inflectionally contrasted morphophonological forms or may present partial syncretism. This fact may have notorious implications on the patterns of the subject-verb agreement: the abstract-level syncretism reduces the evidence for phonetically visible agreement marks; if an increase of phonetic syncretism takes place (due to particular morphology-phonology matching systems), the covert agreement will prevail.

About the role of morphology on agreement, Stump (1998:16) puts it in the following terms:

«... the logic of inflection entails that distinct members of a lexeme's paradigm carry distinct set of morphosyntactic properties; (...) [given that] such properties are by definition syntactically relevant, it follows that inflectional morphology must itself be syntactically relevant in the indirect sense that it spells out a paradigm's syntactically contrasting word-forms.»

From our point of view, it is obviously indispensable to consider the morphological properties of the target (the verb), but we assume that it is necessary to go further and take also into account the phonological properties of the different cells in the verbal paradigms. As a matter of fact, the analysis of spoken data can't ignore the necessary matching between a morphological structure, along with its inherent morphological properties, and their phonological counterpart. Thus, complementarily to the morphology-syntax interaction, the phonological form of the inflectional suffixes and the phenomena that their adjunction to the verbal base implies are central to the understanding of the diverging agreement outputs. That is to say that, in Portuguese, it is important to consider the morphology-phonology interface in the agreement markers of the verb, in order to explain its variant outputs and the shape of variant agreement patterns.

Agreement is also sensitive to lexical-semantic properties (like animacy) and to discourse properties of the controller: the most frequent subjects in Portuguese – count NPs and Nominative pronouns (full or null) – trigger 3rd singular or plural agreement, according to their referential number values, while the other referential pronouns trigger specific agreement inflection (see, again, Table2). But some weak subjects trigger a default 3rd sing. agreement (for instance, non nominalized infinitives, exemplified by *nadar **diverte-o***, 'to swim amuses him'). Syntactic phenomena (see other papers in this volume), like complex NP subjects including a collective-type noun, or a quantifier, and a *de-N* complement may lead to a competing 3rd singular/plural agreement: e.g., *o grupo/a maioria_i dos turistas_{ii} vem_i / vêm_{ii} no Verão*, 'the group/the majority of tourists come [sing/pl] in Summer' (see Peres & Móia (1995); for other contexts favouring sing./pl. *canonical and non canonical* agreement in EP, see Carrilho (2003), Colaço (1999), and the publications of members of the Project, indicated in the references of the

other papers in this volume). This complex set of agreement cues provided by the subject is necessarily to be taken into account when analyzing agreement variant schemes (see Duarte & Varejão, in this volume).

1.1. Person and Number categories

Depending on the lexical class involved, either number or person features present inner differences and distinct ways of representation, as it is briefly summarized below (see Mota, Miguel & Mendes (2012), for further details):

- (i) in the noun, nucleus of a NP, number is referentially controlled, morphologically inflectional and marked by the $-\emptyset$ or $-/S/$ suffixes (e.g. *palavra\emptyset*, *palavras* ('word', 'words')); person is 3rd by default (person is not a nominal category proper);
- (ii) in the 3rd Nominative pronouns: similarly to nouns, number is referentially controlled and inflectional, morphologically marked by the same suffixes (*ele\emptyset*, *eles* ('he', 'they')); person is lexically inherent, with discourse-level information;
- (iii) in the verb paradigms, number and person are inflectional *agreement* features inherited from the subject element, and they occur under the form of PN amalgams (morphologically and phonologically): one suffix, bound to the verb stem (root+thematic vowel) or to the stem+TAM (tense-aspect-mood) morphological constituent, congregates both kinds of information, e.g. *mostra*]_{STEM S}]_{PN}]_{2nd +SING}, *mostra*]_{STEM va}]_{TAM S}]_{PN}]_{2nd +SING} ('you (sing.) show', 'you (sing.) used to show', respectively). In these examples, $-s$ corresponds to the pronominal 2nd rank person feature and to the number of that person, singular (from now on, we will use P1 to P6 symbols to represent the series of the 6 cells of the paradigms, in spite of the fact that the P5 cell is not considered in our analysis; see note 3).

The marked values of these two categories – plural number and plural person-number – are the predictable ones to undergo covert agreement. Giving apart the nature of those categories in nouns and verbs, we concentrate first on the morphological-phonological/prosodic characteristics of verb forms and, later on, on those of the noun forms.

2. The matching between the P6 morphological and phonological structures and its phonetic results

The observation of the Portuguese spoken data showing defective realization of the relation <subject, plural – 3rd verbal form, plural (P6)> suggests that there are different interpretations of the morphological structure of the P6 form. In P6 forms, the person-number feature involves an oral vowel and $-/N/$, the inflectional person-number marker (in the following examples, 6PN represents the PN of the 3rd plural form):

(1) *chegam* (they arrive)

CHEG]_{ROOT} a]_{thematic vowel /N/]}_{6PN} | cheg/aN/ | cheg[ẽw̃]

(2) *chegavam* (they used to arrive)

CHEGA]_{STEM} va]_{TAM /N/]}_{6PN} | chegav/aN/ | chegav[ẽw̃]

/N/ is a fluctuant auto-segment which anchors to the syllable rhyme (see Bisol (1998), Mateus & Andrade (2000)); the results of this process are, first, the nasalization of the precedent vowel (the TV (thematic vowel) or (part of) the TAM), which is most of the time unstressed (stress is morphologically predictable), and, secondly, a phonetic nasal diphthong (a false diphthong): cf. cheg[ẽw̃], chegav[ẽw̃].

(i) The canonical matching

In this kind of matching, there is a full correspondence between morphological structure, phonological form, and phonetic output:

Table 1 – Morphological and phonological structure of P6 and P3 and canonical phonetic outputs

	morphological structure	phonological form (stress, syllabic structure)	phonetic output
	⇔	full matching	⇔
P3	CHEG] _{ROOT} a] _{TV} Ø] _{6PN}	/jẽ.ga/	[jẽgɛ]
	CHEGA] _{STEM} va] _{TAM} Ø] _{6PN} (Ø : null position)	/jɛ.gá.va/	EP (AP) [jígávɛ] BP (AP) [jɛgávɛ]
P6	CHEG] _{ROOT} a] _{TV} -] _{6PN}	/jẽ.gaN/	[jẽgẽw̃]
	CHEGA] _{STEM} va] _{TAM} -] _{6PN} (- : empty position)	<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;"> $\begin{array}{c} \sigma \\ \swarrow \quad \searrow \\ A \quad R \text{---} /N/ \\ \quad \\ /g \quad \tilde{a}/ \end{array}$ </div> <div style="text-align: center;"> $\begin{array}{c} \sigma \\ \swarrow \quad \searrow \\ A \quad R \text{---} /N/ \\ \quad \\ /v \quad \tilde{a}/ \end{array}$ </div> </div>	EP (AP) [jígávẽw̃] BP (AP) [jɛgávẽw̃]

As indicated in Table 1, we assume that the morphological structure of P3 has a final null position: no PN constituent can attach to the TV or to the

TAM (and there is no suffix available to do it, excepted one, marked for preterit tense, *-u*: CHEGA]_{STEM u}]_{3PN} *chegou*, ‘he/she arrived’; TEME]_{STEM u}]_{3PN} *temeu*, ‘he/she feared’. The preterit person-number markers are exceptional). Differently, we posit that for the P6 it is possible to admit a morphological final empty position of a quite particular nature: this position won’t be filled with any explicit consonantal segment, since the floating nasal segment *-N/* attaches to the Rhyme. So, after considering a tentative morphological structure (first column) for the P6 and once analysed its phonological structure (second column), we assume that P6 can’t be fully explained except if we consider it in morphophonological terms, and we come to consider its morphophonological structure as follows:

Table 1.1 – Morphophonological structure of P6 (cf. canonical phonetic outputs)

	<i>chegam</i>	<i>chegavam</i>
P6	CHEG] _{ROOT} [[a] _{TV} [/N/] _{6PN}]] or CHEG] _{ROOT} [a _{TV} ^{/N/ 6PN}]	CHEGA] _{STEM} [[va] _{TAM} [/N/] _{6PN}]] or CHEGA] _{STEM} [va _{TAM} ^{/N/ 6PN}]

Two main arguments for the differentiation of the P3 – P6 structures may be advanced. First, no doubt that the P6 is marked for person-number: the inter-paradigmatic organization shows that, in 7 over 10 verbal Portuguese paradigms, P6 is marked for PN (as well as P2, P4, P5; e.g., imperfect *chegavam*_{P6}, *chegavas*_{P2}, *chegávamos*_{P4}, *chegáveis*_{P5}); this PN gives systematically rise to a nasal diphthong, and there is no canonical syncretism between P3 and P6. In other Romance languages, like Spanish, Catalan or Italian, P6 is clearly marked by a nasal consonant, phonetically visible: respectively, P3 *llegaba*, P6 *llegaban*; P3 *arribava*, P6 *arribaven*; P3 *arrivava*, P6 *arrivavano* – that is to say, the nasality is present in Romance P6 forms and the P3-P6 contrast is effective, in spite of the phonological differences between those languages and Portuguese. Secondly, the same tautosyllabic *-N/* occurs in Portuguese simple nouns, also giving rise to final nasal diphthongs (3a) or to nasal vowels (3b-c) but being phonetically visible as a nasal consonant in derived words formed with the same base as the simple nouns:

- (3) a. *armazém* [ɐʁmɐˈzɛ̃j], *armazenar*
storehouse, to store
b. *patim* [pɐˈtĩ], *patinar*
skate, to skate
c. *tom* [tõ], *tonal*
tone, tonal

Phonologically, final *-N*¹ functions identically in 3rd verbal plural forms and in the simple nouns, as expected and as far as it nasalizes the precedent vowel, but in the verbs it carries a morphological feature (PN), and this feature must have a correspondent structural position (remember our first argument, above): it is, in our view, a morphologically empty particular position, as long as it is not independent of the precedent thematic or ATM position. A position “attached” to the precedent vowel position (or to a part of a position, in CV TAMs), and phonologically marked to be filled with the auto-segment. That is to say, as commented above, a morphophonological position: ultimately, the morphophonological PN of the verbal P6 is revealed by the group /VowelN/ – [nasal diphthong].

The “perfect” matching presented in Table 1 induces an extensive marking of the agreement features on the target.

(ii) Non-canonical matching

In some Portuguese sub-varieties, spoken productions show that there are cases of medium or almost total syncretism of verbal agreement forms (as well described for BP, see Duarte (1993) and Duarte & Varejão, in this volume). We may put several hypotheses to explain that phenomenon (the examples below have been produced by illiterate or less literate informants from Angolan (ANGP), São Tomé (STP) and Brazilian Portuguese (BP), but we find parallel data in other non-European varieties; some examples show the concurrence of two patterns of agreement, like (4c)):

(a) the syncretism results from a morphological reduction, internal to the verbal paradigms (the speakers would have a impoverished inflectional morphology, in their internal grammar); the symbol “–” represents non-syncretism; “±, +” stand for increasing degrees of syncretism:

¹ The particular nature of the *-N*/ auto-segment in Portuguese is also visible when we compare verbs formed through the morpholexical conversion of nominal roots and the nouns which bear those roots, like in the *jejum* → *jejuar* case (fasting, to fast). This word formation process (the nominal root, once converted in verbal, receives the verbal thematic vowel *-a*) shows that the root of the noun *jejum* JEJU/N/-, occurs like JEJU- in the verb *jejuar*; *-N*/ doesn’t anchor in this context (**jejuar*). These examples show very clearly that *-N*/ behaves differently according to the fact that the words involved are morphologically simple, derived or formed by root conversion, a non derivational process.

Table 2 – Degrees of markedness on the indicative present paradigm

1P – 3P	P1 – P3			4P – 6P	P4 – P6		
	–	±	+		–	±	+
eu (<i>I</i>)	fio	fio	fia	nós (<i>we</i>)	fiamos	fiamo	fia
tu (<i>you</i>)	fias	fia		---			
ele/você (<i>he/you</i>)	fia			eles/vocês (<i>they/you</i>)	fiam		

(1P corresponds to 1st person singular pronoun, and so on; P1 stands for 1st sing. form of the verb, and so on)

- (4) a. eu_{1P} fia_{p3} mesmo (ANGP)
‘I do gives credit’, *I do give credit*
- b. que eu_{1P} faz_{p3} parte de uma das famílias que permaneceu lá até hoje (STP)
‘because I is part of one the families’ ..., because I belong to one of the families ...
- c. porque lá que []_{1P} conheci_{p1} o meu namorado, depois []_{1P} arranjou_{p3} filhos, []_{1P} tem_{p3} três filhos lá (...) vivem aqui na cidade, mas []_{1P} arranjei_{p1} lá (STP)
‘because there that I met my boyfriend, after [I] got_{p3} children, [I] have three children there (...) [they] live here in the city, but [I] got there’, *because it was there that I met my boyfriend; afterwards, I had my children (...), I had three children there (...)*
they live here, in town, but I had them there
- d. tu_{2P} tem_{p3} que ir (...) para direcção da floresta (STP)
‘you has to go to (...) in direction of the forest’, *you have to go to the forest*
- e. nós_{4P} já vivia_{p3} até com um certo de medo: ah o que é que pode acontecer? (NIG-PB)
‘we already lived_{p3} even with a certain fear: ah what can happen?’, *we already lived with a kind of fear ...*
- f. nós_{4P} lá chama_{p3} o caparoto (ANGP)
‘we there call_{p3} the “caparoto”’, *there, we call it “caparoto”*
- g. muitos_{6P} vai_{p3} reagir]_{6P} (a)caba_{p3} morren:do ... (PB)
‘many_{6P} is going to react]_{6P} finish dying’, *many will react and finish dead*

(b) the syncretism is triggered by patterns of agreement characterized by the marking of the agreement features only on the left periphery (a strong position) of the NP domain and expanded to the subject-verb domain (the speakers would have a system of reduced morphosyntactic marking):

- (5) a. [os_{-SING} filho_{+SING}]_{-SING} ≡ 6P também *cultiva* P3 (ANGP)
 ‘the_{-SING} son_{+SING} also cultivate’, *the sons also cultivate the land*
- b. só [os_{-SING} criado_{+SING}]_{-SING} ≡ 6P do comandante *pode* P3 [usar]
 (STP)
 ‘only the_{-SING} servant_{+SING} of the commander can [use]’, *only the servants of the commander can [use it]*
- c. [as_{-SING} condições_{-SING} financeira_{+SING}]_{-SING} ≡ 6P não *permitia* P3
 prosseguir (STP)
 ‘the_{-SING} conditions_{-SING} financial_{+SING} NEG allowed to go further’,
the financial conditions didn’t allow (me) to go further
- d. [muitos_{-SING} filho_{+SING}]_{-SING} ≡ 6P que não *entende* P3 (BP)
 ‘many son who NEG understand_{+SING}’, *many sons who don’t understand*

(c) the syncretism is motivated by a particular matching between the morphological and the phonological form of the verbal cells.

We take this last possibility for the present analysis, emphasising that it can be crossed with the precedent one, in (b), in a great deal of productions of non-standard AP and BP speakers: respectively, those who speak Portuguese as a L2 and maintain their L1 in current use (a bantu language or a Portuguesebased Creole), and those who acquired their Portuguese as a L1 but under irregular transmission conditions and have a low degree of formal instruction (see, for instance, Lucchesi, Baxter & Ribeiro (2009) and Lopes & Baxter, 2011:44, Table 3). As far as the (a) hypothesis above is concerned, the results of the Project as well as other works show that it has to be considered with moderation, since we have evidence to assume that the same speakers who may produce *eu fia*, see Table 2, will tend to mark the 3rd plural preterit form (canonical *fiaram*), even if they do it in a reduced way (*fiar*[u], for instance, as attested in BP (Baxter (1997)) and in rural EP): the *-r* is a salient morphophonological marker of this PN, and it is maintained. The expression of this hypothesis is much more limited than that in (b), and the reason why we analyse the P6-P3 relationship independently is the following: each time speakers characterized by systematic agreement productions “fail” the operation, they do it by producing a P3-like form instead of a P6 full marked one. This kind of neutralization or contextual syncretism is the most expanded, when we consider all the varieties of Portuguese together (and it is also the most relevant in other languages, as the recent bibliography of agreement shows).

So, returning to our hypothesis in (c), and bearing in mind the description on Table 1, we next propose an explanation based on the assumption that speakers who produce non-canonical agreement have nevertheless the same morphological representation of P3 and P6 than those who mark the opposition between them (we include in the exemplification phonetic forms attested in EP rural areas):

Table 3 – The matching between the P6 morphological and phonological structure, and the /-(vowel)N/ phonetic resolution: type of matching I

P6 morphological structure		
↑ [+match] ↓		
P6 phonological form		
↗		↘
V-/N/ > V _{+nasal}		V-/N/ > V _{-nasal}
+canonical -identical to P3		-canonical +identical to P3
P6 phonetic output		
↓	↓	↓
[nasal diphthong]	[nasal vowel]	[oral vowel]
P6 ≠ P3		P6 = P3
<i>levam, dão, dizem, medem*</i>	lev[ẽ], d[ẽ], diz[ẽ], med[ẽ]/ med[ĩ]/med[ī]	<i>leva, dá, diz, mede</i>

* they {carry, measure, give, say}

In this description, we go on considering the morphological P6 marker as empty and the P3 marker as null.

The variable realization of the phonological features, in the phonetic output, leads to non-canonical agreement which corresponds to a covert agreement. Those variant realizations correspond to variant patterns which arise from variant interpretation of the phonological form of P6: different phonological knowledge or different phonological grammar implies different kinds of phonetic outputs. In other words, variability in the outputs of agreement may well correspond to phonological rule-governed different matches between phonological and phonetic forms; this phenomenon could explain by itself the P6-P3 syncretism (or the person-number neutralization, see Aalberse & Don (2011)). If this hypothesis is true, then the matching between the morphological and the phonological forms is a constant for all Portuguese speakers.

Now, let's consider that some speakers (a) have the same representation of the P6 morphological structure but there is no matching between that structure and its phonological representation, and (b) they don't have the same representation of the P6 morphological structure, and they match the morphological and phonological structures:

Table 4 – The matching between the P6 morphological and phonological structure, and the /(vowel)N/ phonetic resolution: type of matching II

(a)	(b)
P6 morphological structure	P6 morphological structure
P6 marker: empty vs. P3 marker: null	P6 and P3 markers: null
↑	↓
[-match]	
P6 phonological form	
V _{-nasal}	
P6 phonetic output	
[oral vowel]	
<i>P6 = P3 leva, dá, diz, mede</i>	

It is patent that the result of (a) and (b) is the same, but the (b) scheme may correspond to the grammar of those speakers who reduce most of the person-number inflectional suffixes, and may produce *tu*_{PRON 2ND (+SING)} {*leva, dá, diz, mede*}_{P6=P3}, *nós*_{PRON 1ST (-SING)} {*leva, dá, diz, mede*}_{P6=P3} and even *eu*_{PRON 1ST (+SING)} {*leva, dá, diz, mede*}_{P6=P3} (see example (4)).

In this case, the /N/ segment is not present in the phonological representation of the P6, because the morphological structure has a null marker, identically to P3, and, in discourse, there is no apparent compatibility of features and non apparent co-variation between the controller and the target. To conclude, the “leva-type” realizations for P6 don’t have the same origin in all the varieties of Portuguese: on the one hand we have a /N/ interpretation question (/N/ is not anchored), and on other and extreme cases, a combination of the same phenomenon plus a tendency for syncretism.

2.1. The [+/-matching] schemes: some illustrations

Rural data from EP (and from BP, as presented in Brazilian publications) are useful to show the different realization of -/N/ since the speakers are less exposed to the standard variety. The typically urban speakers with medium/high instruction don’t show the same variation so clearly (note that examples from Lisbon, Oeiras, Cacém, and Sintra, for instance, are extremely rare). But, depending on the regions, the urban speakers may produce 3rd pl. forms with nasal unstressed vowels, -[ẽ] (in the South) and -[ẽ̃] (in the North), but not with unstressed -[õ], -[ũ], -[ĩ] nasal vowels and -[u] oral vowel, typical to non instructed old people from conservative rural regions.

In the next Table, we present the distribution of 245 verbal forms in a P6 syntactic context. We have created two different cells for the paradigms of the verb (1st column) according to the fact that the preterit paradigm has a particular internal morphological structure, with specific PN markers (P1-*i*, P2 -*ste*, P3 -*u*, and P6 -*ram*). Assuming that morphological or phonological complexity could carry less canonical markers, we also detail some types of opposition, which present some kind of particularity concerning (a) the stress (*dão* (they give) is a monosyllable, thus the PN information is located in a stressed syllable, but its structure is identical to other P6 regular forms, /daN/); concerning (b) the last segment of the base: when it is a vowel, (i) there is a possible phonetic insertion of a glide if the vowels are not identical (*saem* (they live)), (ii) the vowel of the base is identical to the theme vowel and they stay in situ (*vêm* (they see)); finally, when the base ends in /z/, the adjunction of -/N/ to P6 triggers a resyllabification (*faz*_{P3} vs. *fa.zem*_{P6} (he does, they do); *fa.zer*_{P3} vs. *fa.ze.rem*_{P6} (to do, personal infinitive)), and the verb is lexically marked [+irregular]. The empty cases correspond to the absence of corresponding forms, but we inscribe “zero” on the last column to underline the absence of 3rd plural forms identical to 3rd singular ones. Notice that we didn’t consider *têm/tem* ((they) have, has), *vêm/vem* ((they) come/comes) because they present only a lengthening of the 3rd sing. diphthong, and almost all the speakers neutralize this difference:

Table 5 – Results from an EP sample: 7 small rural (semi)urban zones (*ALEPG corpus*)

	1	2	3	4	5
	Nasal diphthong	Nasal vowel	Oral vowel ≠ P3	Oral vowel = P3	Clearly = P3
Present, Imperfect	121	23 8 -[ẽ] 7 -[ê] 6 -[õ] 2 -[ũ]	15 -[u]	10 4 -[i] 6 -[e]	0
Preterit (irregular verbs included)	2	4 -[rũ]	8 -[ru]		0
<i>dão</i> / <i>dá</i> type	13				0
<i>sa[j].em</i> / <i>sai</i> type	6				0
<i>vê.em</i> / <i>vê</i> type	3				1 <i>vê</i>
<i>fa.zem</i> / <i>faz</i> type	18				1 <i>faz</i>
<i>fa.ze.rem</i> / <i>fa.zer</i> type	9	2		7 -[r] = P3	0
Tot.	172	29	23	17	2
		85,7%		16,3%	0,08%

Note that the 7 occurrences of *fazer* included in this table belong to the personal infinitive and are present in SP complements, like *para fazer(em)*. This verbal paradigm is commonly used by European speakers from all sub-varieties, but its selection is optional in some contexts (e.g. *eles_i foram à praia para []_i fazer(em) surf* (lit., ‘they went to the beach for [] to do).

The results obtained from the same sample concerning the *são/é* opposition (are/is) and the *vão/vai* opposition (go/goes) are: 39 *são*, 5 [sẽ]/[sõ], and 3 *é*, the last one used in responses to the inquiring of definitions. Concerning the *vão/vai* opposition: we have attested 17 *vão*, 1 [võ], 0 *vai*. These results fully confirm that the most different forms are the less affected (see the Saliency Hierarchy by Lemle & Naro (1977) and successive works), as well as the fact that the word final stress is relevant: the robustness of these P6 forms may come from the fact that stressed vowels have a stronger degree of nasality than non stressed vowels, and that a nasal vowel is longer than an oral one (see Moraes & Wetzels (1992) BP analysis). Furthermore, according to recent research on the structure of the mental lexicon, this kind of oppositions would be stored in the lexical memory of the speakers (as it happens with the allomorphic verbal roots or stems). In the productions of less instructed L2 speakers/L1 speakers having acquired Portuguese through irregular transmission, who show more non-canonical agreement, it’s very possible that they don’t have the same internalized forms in their lexicon, and it’s obvious that there are some cases of P6 *vai, dá* in the corpus of São Tomé as well as in BP (see Brandão & Vieira (2012)).

We must distinguish the forms in column 5 from those in column 4: oral vowel forms or *-r* personal infinitive forms identical to 3rd sing. are best analysed as cases where /N/ doesn’t anchor to the syllabic rhyme. The rareness of those forms allows us to sustain that, in EP, clear syncretism 3rd sing./pl. is not relevant, but the non-anchored /N/ scheme is present, even if at a small scale (17 final oral vowel tokens/246 tokens; see column 4 of Table 5). In the indicative preterit paradigm (including irregular verbs), the zero occurrences of P6 identical to P3 may result from the fact that there is no modification of the stem, the 3rd pl. presents the sequence /vowelN/, like the present, the imperfect, and so on, even though the opposition supposes a more complex morphological analysis of the singular vs. plural structure (stem]-u marker vs. stem]-ra/N/ marker) than in the present or the imperfect paradigms («the two forms [fez/fizeram] are very different both in terms of root vowel and total number of syllables» (Scherre & Naro 2010:165). Nevertheless, in a corpus of 35 informants of Cape Verdean Portuguese and Luanda Portuguese (educated speakers), we have found only two examples of non-canonical *deu/deram* saliency type agreement and none in the sample of dialectal EP, as noted above; Vieira 2011 found three cases in the Lisbon region, syntactically determined (presence of unaccusative verbs and VS order); see Baxter, 1997:269, for the P6 preterit paradigm “levar” forms in Helvécia, identical to EP rural forms in Table 5.

Moreover, the observation of forms in column 4 of Table 5 favours the conclusion that an oral vowel, *per se*, is not a manifestation of syncretism. This aspect seems central to an objective interpretation of agreement – if this assumption is correct, the presence of a plural subject combined with a final vowel verbal form may well not be equivalent to the absence of agreement.

Taking into account the data of our Table 5 sample corresponding to the 1c level *faz/fazem* opposition of the Saliency Hierarchy, and considering also the *vêem/vê* and the *saem/sai* oppositions, we find that only 2/55 occurrences present P6 identical to P3: *faz* and *vê*.

We have noticed that stress, specially combined with nasal segments, favours the canonical morphology-phonology matching. Now, as far as the resyllabification of P6 forms is concerned: in *faz/fazem* or *fazer/fazerem*, the number of syllables is bigger in the plural form, but this phenomenon is also present in other types of forms in the Table 5 and nevertheless only 9/47 forms (19%) of the *sa[j].em/sai* type, *vê.em/vê* type, *fa.zem/faz* type, and *fa.ze.rem/fa.zer* type are identical to 3rd sing. Thus, the number of syllables and the resyllabification process are not, in essence, a blocking factor for distinguishing P6 from P3 forms, the same being true for the presence of an extra syllable when the preceding segment is a vowel.

To sum up, the saliency scale, in the way that is has been conceived for BP by Lemle & Naro 1977, doesn't apply totally to EP data (remember that rural speakers from whom we have taken the data in Table 5 are those who show more variant realizations of P6 -/N/) and, partially, to AP data.

We are aware that in BP, in AP and, in a lesser average, in EP, the level of formal education is a prominent factor conditioning the absence of person-number oppositions. Brandão as well as Vieira & Bazenga (in this volume) have confirmed this fact: in Brazilian and in São Tomé Portuguese this variable is quite relevant, and in Funchal (EP) there is some evidence of its significance. The same holds true for the type of acquisition, as we noticed before, in particular in some BP and AP varieties:

«The comments in Emmerich (1984, p. 208) about the reality observed in Kamayurá Portuguese data are instructive. She notes that the speakers' mother tongue phonology may condition the perception of the saliency. In fact, in L2 varieties or in L1 varieties undergoing strong influence of L2 data, the perception of the saliency may well not correspond to the one proposed by the linguist when he/she describes the internal factors of standard Portuguese system.» (Lopes & Baxter 2011:45)²

In the next two sections, we consider the nominal and pronominal subject elements (the controller of the agreement) inasmuch as their morphological

² The quotes from Brazilian and Portuguese authors were translated into English by the author of this paper.

and phonological features and structure may be compared with those present in the verb; through this comparison, our goal is to contribute to the understanding of both the lack of overt marks on the verb and on the nominal subject for non syntactic reasons.

2.2. Person-number morphophonological markers: heavy vs. light syllables

The (person-)number markers occupy the right edge of the words implicated in the agreement domains (for pronouns, only *ele(s)* show this pattern, the other ones being independent lexical units). This peripheral position is morphologically determined and, in principle, weak.

In the verb, each paradigm have an inherent, lexically and morphologically determined position for main stress – the syllable containing the theme vowel (e.g. for.ma.mos (we form)), the root (e.g. for.m]a (he/she forms)) or the syllable including the TAM suffixes (for.ma.rá] (e.g. he/she will form)).

Parallel to nouns, most of the verbal word-forms are paroxyton; however, the adjunction of person-number suffixes to the base may induce proparoxyton words (e.g. subjunctive imperfect, P4 suffix *-mos*: for.ma.sse.mos) and oxyton ones (indicative future, e.g. P3 for.ma.rá).

On the other hand, this same operation gives rise to a high number of CVC syllables in the majority of the paradigms of most Portuguese verbs. Again, the comparison with nouns is useful: CVC is the most frequent syllable structure in plural nominal word-forms, as described in section 2.3, and in plural³ person verbal word-forms, as well as in the majority of 2nd person singular forms (for.mas, for.ma.rás, for.ma.vas, for.mes, for.ma.ses, for.ma.res).

The person-number suffix has the property of never receiving the word main stress, its post-stressed position reducing its robustness; this property may be important to explain the non realization of person-number (PN) markers, in reduced agreement patterns, as noticed above (see also Cysouw (2008), for other languages description). The suffix under analysis may nevertheless be included in a stressed syllable, in the 3rd pl. indicative future: in for.ma.r[ẽw̃] ('they will form'), the phonological structure of the last syllable is /raN/, where /ra/ is the TAM constituent and /N/ is the person-number suffix. This situation may be consistent with syllable robustness, growing the probability of 3rd pl. overt agreement. All the other 3rd plural forms present a -N/, the result of this process being a phonetic unstressed nasal diphthong (the glide is absent in the phonological form): for.m[ẽw̃]

³ We use “plural” in an informal way, since the “plural” Nominative pronouns don't behave like true plural equivalents of “singular” ones. Cysouw (2008) proposes the concept of group pronouns, instead of plural pronouns. We don't consider, in this paper, the 5th cell of the paradigms, corresponding to the subject pronominal *polite form vós*, which has a very restricted use, contemporarily.

(‘they form’), for.ma.r[ẽw̃] (‘they formed’), for.m[ẽj̃] (subjunctive present), etc. correspond to final unstressed C[V_{+NASAL}glide_{+NASAL}]NUCLEUS syllables.

The parallelism with some of the nouns described in the next section seems evident – the adjunction of *-S/* and *-N/* is common to a great amount of nouns and verbal forms (despite the fact that those segments realize different categories), and it is responsible for the same syllabic structure in both lexical classes (cf. CVC # column).

Table 6 – Adjunction of PN markers and their effect in the resyllabification

	<i>-S/</i>		<i>-N/</i>
CVC #	Nouns: ca.dei. <u>ras</u> , ma. <u>res</u> , paí. <u>ses</u>	C nasal diphthong C #	Nouns: ir. <u>mãos</u> , con.fu. <u>sões</u> , re. <u>fêns</u>
	Verbs: for. <u>mas</u> , for.ma. <u>vas</u> , for.ma. <u>rás</u> , ...	C nasal diphthong #	Verbs: for. <u>mam</u> , for. <u>mem</u> , for.ma. <u>ram</u> , ...
C oral diphthong C #	Nouns: ca. <u>nais</u> , an. <u>zóis</u> , tú. <u>neis</u>		

Specific to verb paradigms, the systematic 1st pl. CVC suffix *-mos* is itself a CVC unstressed syllable: for.má.va.mos, for.ma.mos, for.másse.mos, etc. In the preterit paradigm, the singular cells present *-i* (1st), *-u* (3rd), and *-ste* (2nd); the consonants of 2nd PN split, *-s* occupying the pre-final coda and *-t* anchoring to the onset position of the final syllable: *formaste* /for.máf . t e/⁴, after resyllabification. In the present paradigm, the 1st sing. PN is *-o* (/u/). They behave in the following way, respectively (the transcription corresponds to EP):

⁴ The orthographic *-<s>* correspond to [ʃ], if followed by unvoiced consonants (within the word or in sandhi contexts) or by a pause; in some non-standard varieties, it may correspond to [s], in the same contexts. The allophones [z] and [ʒ] of /S/ occur in other contexts (see Mateus & Andrade (2000), for instance), but, to simplify the present description, we will consider only the first realization.

Table 7 – Preterit singular paradigm and Present P1: resyllabification and phonetic outputs

	Singular cells		EP (standard)	EP (other varieties)
Preter.	1 st : for.mei	forma] _{STEM i}]PN	[fuɾ.méj]	[fuɾ.mé]
	3 rd : for.mou	forma] _{STEM u}]PN	[fuɾ.mó]	[fuɾ.mów],[fuɾ.méw]
	2 nd : for.mas.te	forma] _{STEM Ste}]PN	[fuɾ.máftʰ]	[fuɾ.máf.ti]
Pres.	1 st : for.mo	form] _{ROOT Ø}]STEM o]PN	[fɔɾ.mu]	[fɔɾ.mu]

Elsewhere, Ø is the person-number suffix. In 7 of 10 paradigms, Ø is present in the 1st and the 3rd singular cells, conveying structural syncretism⁵.

2.3. Number morphophonological markers: heavy vs. light syllables

First of all, let's remember that in EP the plural marker is acquired very soon (before two years of age), independently of the morphophonological complexity of the nouns:

«patterns of variation show that the child elaborates an architecture of [phonological] features taking into account the syllabic structure of the target [the sibilant]. (...) The emergence of the point of articulation [-anterior] in the coda position (...) follows the order:

- (13) Word-final coda (without sandhi) (1;9) >> Word-final coda (with sandhi) (1;10) >> medial coda (3;0)

We know that, in EP, the final fricative coda carries very often a morphosyntactic information since the plural marker occurs in this phonological position. Differently, medial codas are exclusively lexical» (Almeida et al., 2010:164)

(See two examples from the same paper, p. 160: (9a) *bonecas* ('dollies') /bunékeʃ/ → [mɨnéteʃ] (1;9); (9c) *meninos* ('boys') /mɨnúnuʃ/ → [mɨnúnuʃ] (1;9)).

In the noun, as far as the word stress and its prosodic relevance are concerned, this position is in fact weak. The great majority of the Portuguese

⁵ The verbal theme vowel and the nominal thematic class vowel are both lexically determined and central to the well-formedness of inflectional forms. However, we must remember that some nouns are athematic and nevertheless undergo inflection, and that two of the verbal paradigms lack the thematic vowel (indicative present, P1; subjunctive present: all the cells).

nouns⁶ are thematic, and the thematic class vowels *-a*, *-o*, *-e* are inherently unstressed. As a result, the Portuguese stressed syllable is widely the penultimate⁷ (e.g. tra.ve.ssa (‘plate’), e.le.fan.te (‘elephant’), ca.dei.ra (‘chair’)). According to *Dicionário Inverso do Português* (Andrade (1993)), 74.5% of the 82.700 Portuguese non verbal words considered have open syllables in their singular form.

The adjunction of *-S/* plural (with three contextual-driven allophones [ʃ], [ʒ], [z]) does imply, however, the occurrence of a heavier syllable: a CVC_{-SING} as opposed to a CV_{+SING} syllable, as illustrated by te.lha.dos (‘roofs’), e.le.fan.tes, ca.dei.ras – in Portuguese, the majority of roots end in consonant (telhad]os, elefant]es, cadeir]as) which implies a CVC final syllable, in the plural word-forms; examples like óle]os or fême]as (‘oils’, ‘females’), with a root presenting final CV syllable, are in a few number.

In addition, the adjunction of *-S/* plural in nouns ending in *-r*, *-s*, and in a few ending in *-l* triggers the resyllabification of the word through the insertion of a default vowel (a supporting vowel, as Bisol (2010) also calls it; see also Mateus & Andrade (2000)), and the insertion gives rise to a new final heavy CVC syllable: e.g. mar_{+SING}, ma.res.SING (‘sea(s)’), mal_{+SING}, ma.les_{-SING} (‘disease(s)’), pa.ís_{+SING}, pa.í.ses.SING (‘country/-ies’). In nouns like *canal* (‘canal’), *anzol* (‘fish hook’) or *túnel* (‘tunnel’), other phenomena occur that lead to even more heavy syllables, in the plural, with a complex nucleus: CVglideC: ca.nais, an.zóis, tú.neis.

Most of the words with syllabically heavy finals correspond to nouns ending in stressed *-ão*_{SING}. They may receive *-S/* plural in the simplest way, if there is a nominal thematic class vowel in their morphological structure, which surfaces as the glide [w̃] (irma/N/]ROOT O]CLASS VOWEL /S/]PL *irmãos* (brothers). The final *-ão* nouns may undergo more complex modifications, if athematic, the nasal diphthong being not lexically conditioned. Independently of the origin of the diphthong (see Bisol (1998) for details), the result is that the plural triggers a heavier final syllable structure: C[V_{+NASAL}glide_{+NASAL}]NUCLEUSC, as in ir.m[ẽw̃]s, con.fu.s[õj]s (‘confusions’) – see Table 8. Moreover, those syllables typically receive the word stress, a supplementary factor of complexity.

Now, if we relate these properties with the probability of the overt realization of number and its result on agreement, we may predict them to be robust enough to be preserved by the speakers. We may similarly expect that

⁶ The same holds true to *ele(s)*, which has a number inflection identical to the thematic nouns. In this paper, we don’t extend this comment to the pre-nuclear or to other elements occurring within the NP agreement domain or in other syntactic contexts. See Brandão, Miguel & Mendes, in this volume.

⁷ Note that typical athematic nouns ending in oral or nasal vowel/diphthong are oxyton – *café* (coffee), *galã* (handsome man, gentleman), *herói* (hero), *refém* (hostage), as well as those ending in *-S/* or *-l/*, as *pais* (country), *capuz* (hood) or *canal* (canal).

the prominent allophones of *-S/* favour its overt realization, in particular the palatal [ʃ], present in many of the contexts of our samples from EP, AP and Rio de Janeiro BP.

Comparing the main lexical classes implicated in an agreement relation, we would like to remember some relevant facts. The controller and the target may present different degrees of specification of person and number; these categories may not have exactly the same properties, which results on no direct matching between them. For instance, person is (i) ordinal in pronouns and in verbs, and cardinal in nouns, (ii) it is lexical in nouns and pronouns but not in verbs; number and person are independent categories in nouns, in pronouns and in verbs, but the morphological and phonological features which realize them are amalgams in the verb, a *-S/* number suffix in the referential *ele*, which is lexically marked for person, the other Nominative pronouns being lexical independent words, marked for person and number on the lexicon.

The precedent condensed description allows us to conclude that the Portuguese nominal and verbal inflectional systems are extremely complex. We return to this aspect later on.

In a great number of nominal items and verb cells, the contrast between singular and plural morphologic, phonologic and phonetic forms is quite prominent. In Scherre & Naro, 2010:165, it is said that the saliency scale is «based on strictly phonetic criteria». Preponderance is given to the output forms, which are obviously important as far as the word-forms constitute the immediate input for language acquisition. Nevertheless, considering the results from recent research on acquisition, it is recognized that the child doesn't receive homogeneous phonetic inputs, in current social situations; he/she has to deal with variable and eventually contradicting stimuli (in shape and in frequency) and doesn't copy the input forms, but scan the evidences the inputs offer to "set the parameters" (Lightfoot, 1991) of his/her own grammar:

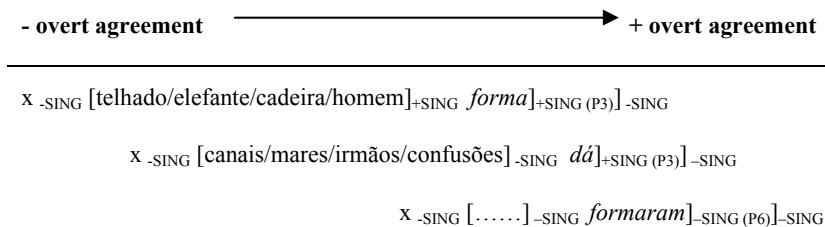
«The cue-based approach distinguishes external language, the mass of unanalysed utterances that a child might hear, from the internal languages or 'grammars,' the systems that grow in children on exposure to external language. (...) The core idea in the cue-based approach is that an internal language grows in a child in response to structures that are expressed in the ambient, external language that she hears» (Lightfoot, 2010:667)

Assuming this perspective, the observed variant patterns of agreement will result from the interpretation, on an abstract level, of the spoken material surrounding the child. Deliberately putting syntactic and other factors apart, if we concentrate on the intricate morphological structures of the lexical classes involved in agreement and on their phonological correspondents, we see as a strong possibility that the matching between morphological and phonological properties do intervene in and explain the existence of variant agreement patterns (see section 2). That is to say that the phonological/

prosodic phenomena interact with morphology in an obvious way. Furthermore, it is important to remember that «The nature of the human language capacity is linked inextricably with the way it is acquired» (Lightfoot, 2010: 678): as well explained by sociolinguists (see Lucchesi, Baxter & Ribeiro (2009), among others), the kind of transmission and the frequency of certain patterns in the linguistic environment of the speakers affect the grammar they grow, and, most probably, explain the maintenance, for long periods of time, of variant patterns in co-occurrence in society, as well as the changes observed, which mirror ongoing social changes, specially the increase of formal education.

In sum, in the subject-verb domain, agreement presupposes, beyond the strictly syntactic constraints, a complex interplay of relations, properties and features; its nature is morpho-phono-syntactic, sensitive to the lexical-semantic features and discourse properties of the controller. This complexity has to be considered to understand the multiple agreement patterns observed in Portuguese. In Brandão & Vieira, 2012: 22-23, the data from São Tomé Portuguese and Brazilian Portuguese show that the phonic saliency is one of the factors explaining the absence of the 3rd pl. marker, particularly in BP data. Their results correspond to those obtained by other researchers: the person-number opposition unstressed CV_{+SING} vs. unstressed Cnasal diphthong_{-SING} (see Table 6) corresponds to the highest rate of the relevant absence. In a previous version of that text (not published), the same authors presented the conclusion that, in the Project data from Copacabana and Nova Iguaçu, the nouns in the NP showing a final oral vowel (as illustrated by *telhado*, *elefante*, *cadeira*, above) or a diphthong resulting from /-eN/ singular ending (as illustrated by *homem*) favour the absence of an overt plural mark. The other types commented above (*mares*, *países*, *males*, *canais*, *irmãos*, *confusões*, see Table 6) tend to maintain the +sing/-sing oppositions. Crossing those results, we may conclude that, for Rio de Janeiro BP, a typical continuum scale of agreement would resemble to:

Figure 1 – Continuum of NP-V agreement in BP: a hypothesis



(“x” represents any pre-nominal element and is marked with [-sing] according to the fact that the first position in a NP is almost always marked for number)

Keeping in mind that the more complex nouns are those which maintain the +sing/-sing oppositions, let's look at the internal morphologic properties of the -/vowelN/ nouns (see Table 8, below), those which surface as [ĩ(S)] (*marfim/-ns*, 'ivory'), [ũ(S)] (*atum/-ns* 'tuna fish(es)'), [õ(S)] (*tom/-ns* 'tones')⁸. They are athematic, and that is an additional complexity, since the majority of the Portuguese nouns are thematic. However, they receive -/S/ plural directly (see column a. of the next Table), like thematic nouns in general (e. g. *telhados, cadeiras, elefantes*), the same holding true for the more phonologically complex thematic type illustrated by - [ẽw̃] nouns (*irmãos* 'brothers'; see column c. of the next Table). On the other hand, the athematic -ão singular nouns, like *pão* ('bread') and *limão* ('limon') (see column b. of Table 8), with a low vowel (p/ã/) or a mid vowel (lim/õ/) in the syllable which receives the plural mark, present different plural endings from the thematic - [ẽw̃] nouns (another complication): they have [ẽjS], -[õj] plural forms, cf. [pẽj], [li'mõj] (in medieval Portuguese texts, we find *pã, pan, pam* and *leõ, leon* or *condiçõ, doaçom*, and in some present day EP varieties, the -[ẽ], -[õ] correspondent pronunciations). The question may be summarized as in the following Table, according to my analysis of the plural formation of these complex nouns⁹:

Table 8 – Morphological and phonological structure: -/vowelN/ nouns

a. athematic -/root vowelN/	b. athematic -/root vowelN/	c. thematic -/root vowelN/ + class marker /u/
marfi/N/] _{ROOT S}] _{SING} [mer'fij]	pa/N/] _{ROOT S}] _{SING} - INSERTION OF A DEFAULT V, LIKE IN -R, -S NOUNS p/ã/] _{ROOT es}] _{SING} - DEFAULT V → HOMORGANIC GLIDE, NASALISED [pẽj]	irima/N/] _{ROOT O}] _{CLASS MARKER S}] _{SING} iriñ/ãu/] _{STEM S}] _{SING} [ir'mẽw̃j]

⁸ In BP, words presenting -[õ] singular form may be produced with a phonetic glide in the plural form: [tõj], *tons*.

⁹ On behalf of the organization of this paper and considering its scope, we exclude from the discussion nouns and other non verbal words ending in stressed -/aN/, like *divã* ('divan'), *irmã* ('sister') or *vã* ('vaine'), and we don't develop the analysis of the -/eN/ final stressed words. Their consideration would oblige us to expand the analysis to other questions.

Note that there is a difference between oxyton and non oxyton athematic nouns: in the Figure 1, *homem* figures along with thematic nouns because it is paroxyton (the same for *virgem* ('virgin') or *margem* ('edge'), for instance), which is atypical for athematic nouns. In the Southern European varieties, the speakers will produce nasal vowels in verbal forms (see Table 5) and also in other lexical class words bearing a final *-/eN/* sequence (stressed (*al[ẽ]*, *além* ('there'); *por[ẽ]*, *porém* ('nevertheless')) or unstressed (*hom[ẽ]*, *homem*, etc.). On the other hand, the EP, BP and AP unlettered speakers may produce oral vowels in final unstressed words – *hom[ɨ]*, *marg[ɨ]* –, but both the Southern and the illiterate speakers never do it in final syllable stressed words: e. g. *além* never surfaces like **al[e/ɨ]*, the same holding true to the other oxyton words.

The fact that, as tentatively shown in Figure 1, the speakers don't prefer the most simple solutions (they have the best agreement results for b. and c., the most complex morphological, phonological and phonetic structures) highlights the conclusion that structural simplicity is not relevant in the grammar of the speakers in general; consequently, the most productive plural forms for agreement are the most complex ones, at a morphophonological level.

3. Final remarks

Speakers have different grammars. Different-level extra-linguistic and linguistic factors may concur to that fact, displayed by the agreement patterns variety. We have tried to show that overt and in particular covert agreement is best understood if we look carefully to the morphology-phonology interface. We sustain that there is a common morphological structure to all the speakers of Portuguese, and that there is always agreement, in an abstract level. However, the matching between morphological structures and phonological forms is not uniform, in Portuguese. In fact, this morphology-phonology frontier is ultimately the primary locus of the differentiation between overt vs. covert marking of the agreement features. If the phonological form of a verbal word is interpreted differently from the way the standard "full matching/agreeing" speakers do, its phonetic output will be also different. For instance, if the auto-segment *-/N/* is interpreted and nasalises the precedent vowel, some speakers may stop the process there, and produce *lav[ẽ]*-type P6 forms (the standard ones will proceed with a phonetic diphthongization); if speakers don't interpret the auto-segment, an oral vowel will surface and the P6 form will be identical to the correspondent P3.

We have advanced a short comparison between verbal and nominal complex formation of plural person-number and number forms, respectively, aiming at verifying the relationship between morphophonological complexity in verbs and nouns and the overt vs. covert agreement patterns. The fact that

“simplicity” is not relevant for the speakers is an interesting conclusion, in our view.

The propositions presented throughout this paper are to be crossed with the results obtained by the members of the Project and presented in this volume, so that their pertinence can be verified (namely those handling with subject-verb agreement, agreement within the NP, and the relevance of the subject type).

References

- Aalberse, S. & Don, J. (2011) Person and number syncretisms in Dutch. *Morphology*, 21, 327–350.
- Almeida, Letícia; Costa, Teresa & Freitas, M. João (2010) *Estas portas* [ˈɛtəsˈpɔtəs] e *janelas* [n:ɐˈnɛs]: o caso das sibilantes na aquisição do Português Europeu. In *Textos Seleccionados do XXV Encontro Nacional da Associação Portuguesa de Linguística* (Ana M. Brito et al, orgs.), pp. 153-167. Porto: APL.
- Andrade, E. d' (1993) *Dicionário Inverso do Português*. Lisboa: Cosmos.
- Baxter, A. N. (1997) Creole-like features in the verb system of an Afro-Brazilian variety of Portuguese. In *The Structure and Status of Pidgins and Creoles* (A. K. Spears & D. Winford, eds.), pp. 265-288. Amsterdam: John Benjamins.
- Baxter, A. N. (2004) The development of variable NP plural agreement in a restructured African variety of Portuguese. In *Creoles, contact, and language change: linguistic and social implications* (Geneviève Escuré & Armin Schwegler, eds.), Vol. 27, pp. 97-126. Amsterdam: John Benjamins.
- Bisol, L. (1998) A Nasalidade, um Velho Tema. *DELTA* [online], vol.14, n. spe.
- Bisol, L. (2010) O diminutivo e as suas demandas. *DELTA*, vol. 26.1, 59-85.
- Brandão, S. F. & Vieira, S. (2012) A concordância nominal e verbal no Português do Brasil e no Português de São Tomé: uma abordagem sociolinguística. *PAPIA*, 22(1), 7-39.
- Corbett, G. G. (1998a) Agreement in Slavic. (Workshop on Comparative Slavic Morphosyntax, Indiana) <http://www.indiana.edu/~slavconf/linguistics/index.html> (access: June 2013).
- Corbett, G. G. (1998b) Morphology and agreement. In *The Handbook of Morphology* (Andrew Spencer & Arnold Zwicky, eds.), pp. 191-205. Oxford: Blackwell.
- Corbett, G. G. (2006) *Agreement*. Cambridge: Cambridge University Press.
- Cysouw, M. (2008) *The Paradigmatic Structure of Person Marking*. Cambridge: Cambridge University Press.
- Duarte, M. E. L. (1993) Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. In *Português Brasileiro: uma viagem diacrônica* (Ian Roberts & Mary Kato, orgs.), pp. 107-128. Campinas: Editora da Universidade Estadual de Campinas, UNICAMP.
- Lemle, M. & Naro, A. J. (1977) *Competências básicas do português*. Relatório final de pesquisa apresentado às instituições patrocinadoras Fundação Movimento Brasileiro de Alfabetização (MOBRAL) e Fundação Ford. Rio de Janeiro. Fundação Movimento Brasileiro de Alfabetização.

- Lightfoot, D. (1991) *How to set parameters: Arguments from language change*. Cambridge, Ma: MIT Press.
- Lightfoot, D. (2010) Language acquisition and language change. *Advanced Review*, Vol. 1, 677-684. http://www.socsci.uci.edu/~lpearl/colareadinggroup/readings/Lightfoot2010_LangAcqLangChange.pdf (access: June 2013)
- Lopes, A. M. M. *As línguas de Cabo-Verde: uma radiografia sociolinguística*. PhD Thesis, manuscript. 2011. University of Lisbon. (Available at the University of Lisbon, Faculty of Letters Library)
- Lopes, N. & Baxter, A. (2011) A concordância verbal variável no português dos tongas. *PAPIA*, 21(1), 40-50.
- Lucchesi, D. (2003) O conceito de transmissão lingüística irregular e o processo de formação do português do Brasil. In *Português brasileiro: contato lingüístico, heterogeneidade e história* (Cláudia Roncarati & Jussara Abraçado, orgs.), pp. 272-84. Rio de Janeiro: 7 Letras.
- Lucchesi, D. & Baxter, A. (2009) A transmissão lingüística irregular. In *O Português Afro-Brasileiro* (Dante Lucchesi, Alan Baxter & Ilza Ribeiro, orgs.), pp. 101-24. Salvador: Edufba.
- Mateus, M. H. & d'Andrade, E. (2000) *Phonology of Portuguese*. Oxford: Oxford University Press.
- Moraes, J. & Wetzels, L. (1992) Sobre a duração dos segmentos vocálicos nasais e nasalizados em Português. Um exercício de Fonologia Experimental. In *Fonologia do Português* (Wetzels, Leo & M. Bernadete Abaurre, eds.), pp. 153-166. *Cadernos de Estudos Linguístico*, Nº Especial: IEL/UNICAMP.
- Mota, M. A.; Miguel, M. & Mendes, A. (2012) A concordância de P6 em português falado. Os traços pronominais e os traços de concordância. *PAPIA*, 22(1), 161-187.
- Scherre, M. M. P. & Naro, A. J. (2010) Perceptual vs. Grammatical Constraints and Social Factors in Subject-Verb Agreement in Brazilian Portuguese. *U. Penn Working Papers in Linguistics*, Vol.16.2, 165-170.
- Stump, G. (1998) Inflection. In *The Handbook of Morphology* (Andrew Spencer & Arnold Zwicky, eds.), pp. 13-43. Oxford: Blackwell.

Maria Antónia Mota
Faculdade de Letras da Universidade de Lisboa
Centro de Linguística da Univ. de Lisboa
motasaojoao@gmail.com