CONFRONTING FACT AND THEORY IN THE DISCOVERY OF BRAZILIAN PORTUGUESE GRAMMAR¹

CONFRONTANDO FATO E TEORIA NA DESCOBERTA DA GRAMÁTICA DO PORTUGUÊS DO BRASIL

Mary Kato²

ABSTRACT

This paper shows interesting questions and results that one can obtain when we combine the variationist methodology with the generative theories in the analysis of data. We believe that, while the Principles and Parameters theory could provide us with predictions as to what kinds of correlation to find in real data, a well conducted empirical analysis could reveal new facts unaccounted for and correlated phenomena. Methodologically, formal grammars help the researcher to constrain the universe of formal factors to be considered and to formulate more explicit questions. The variationist methodology, in its turn, allows us to use not only syntactic variables, but also constraints of other modules of the language faculty. It can also distinguish results due to either innate principles or environmentally conditioned factors. This paper presents some of the interesting findings on Brazilian Portuguese regarding the Null Subject Parameter (NSP), a topic that has been dealt with from both diachronic and synchronic perspectives.

KEYWORDS: Brazilian Portuguese; Principles and Parameters; variation methodology; diachrony; spoken language; written language

RESUMO

Este trabalho mostra as interessantes questões e resultados que se obtém quando combinamos uma metodologia variacionista com a teoria gerativa na análise de dados linguísticos. Nossa convicção é de que, enquanto a teoria dos Princípios e Parâmetros podia nos fornecer teorias sobre que tipos de correlações encontrar nos dados, uma análise empírica bem conduzida poderia nos revelar fatos novos e correlações não preditas. Do ponto de vista metodológico, enquanto os formalistas ajudam o pesquisador a restringir o universo de fatores a considerar, e a formular questões de pesquisa mais explícitas, a metodologia variacionista nos permite usar, não apenas fatores morfossintáticos, mas também fatores de outra natureza, Esta combinação também permite distinguir resultados devidos a princípios inatos, mas também a fatores socioambientais. Este trabalho nos apresenta alguns dos resultados de um tópico de investigação, o Parâmetro do Sujeito Nulo, que perpassa tanto o projeto diacrônico, quanto os sincrônicos.

PALAVRAS-CHAVES: Português Brasileiro; Português Europeu; Princípios e Parâmetros; metodologia variacionista: diacronia; língua falada; língua escrita

² Professora do Programa de Pós-Graduação em Linguística. Universidade Estadual de Campinas (UNICAMP). Contato: mary.kato@gmail.com



Emeritus Miriam Lemle's career

¹ The content of a shorter version of this paper was presented at the Conference: *Linguistic Data Structures: on the relation between data and theory in Linguistics*, Tübingen, 2002. The present paper updates the previous one with more recent projects and results developed in the same spirit. I thank Marcello Marcelino for his usual help with content and form.

1. Introduction

Traditionally, most generative work dispenses with empirical data from E-language and most empiricist researchers, in turn, ignore formal theories of grammar, which have I-language as their object. The Brazilian linguistics developed at the University of Campinas has a unique tradition in combining the two approaches in several projects. The late Fernando Tarallo and myself have set in motion this line of research in Brazil, developing two projects: a diachronic study of Brazilian grammar and a description of spoken Brazilian Portuguese (SBP). After his premature death, I continued with this tradition with several colleagues and students.

I will first touch on the history of each project and then I will go on to discuss some of the important results obtained.

2. The projects

When Fernando Tarallo came to Brazil to collect data, during his Ph-D in Pennsylvania, he read an article of mine on the individual variation on relative clauses in BP, in which I posed an implicational hypothesis between the variation in the relative clauses and in the overall pronominal system. Tarallo decided to test this hypothesis empirically and diachronically and he found a statistical confirmation of its main points, namely, that a) BP lost the case-marked relative pronouns and developed the chopping relative along with the resumptive relative clause and b) that chopping relatives correlated with the increase of null objects. When he finished his doctorate, he was hired by my university (at the time, the Catholic University in São Paulo). Excited by our first endeavor, the first question he asked me was whether I had some new hypotheses to test. It so happened that I did have one question to do with the loss of the indefinite *se* construction in BP. We wrote a paper together, which was also an original enterprise combining formal hypothesis and empirical data³. However, when we first presented the paper to a Brazilian audience, we were attacked by both generativists and variationists, who claimed that we were heretic linguists. But the article ended up being published in the States in a volume edited by a generativist and a sociolinguist (cf. Jaeggli and Silva-Corvalán, 1986).

³ Cf Tarallo and Kato (1993).





A diachronic project was then planned⁴. UPenn was a good example to follow, since Anthony Kroch and collaborators were doing exactly what we intended to do, namely, the study of old *corpora* using generative hypotheses⁵. Generativists could not criticize us since we could not use the intuition of deceased people. Positive data were the only source to test grammaticality. No negative evidence could be used. We used theater plays, preferentially comedies, a genre closer to spoken language. According to Chomsky and other generative psycholinguists, use of positive data is also what happens when a child acquires his/her grammar (cf. more recently Yang 2002). If a child can discover his/her grammar being exposed just to positive data, why couldn't a linguist discover a grammar without drawing on negative evidence? Moreover, what was interesting to observe was that immersion in old *corpus* often made the researchers become as competent as native speakers of the language of the past periods. One starts having intuitions about the grammar of the period in question, and the *corpus* often showed that the researcher was right in his/her intuitions. We tried to base our study on genres that closely mirrored the vernacular use of the language: personal letters and plays.

The partial results of the diachronic project can be seen in a volume edited by Ian Roberts and myself (cf Roberts and Kato 1993; re-edited as Roberts and Kato 2018), a *Festschrift* for Fernando Tarallo, who we unfortunately lost at an incredibly early age⁶. The book is quite exhaustive in topics as it covers the realization of arguments, position of clitics, wh-constructions, agreement, possessive pronouns, and periphrastic verbal forms, besides two classical diachronic articles by Fernando Tarallo, in which he discusses the possible causes for the changes that gave birth to a new grammar: Brazilian Portuguese (BP)⁷.

While planning the ongoing diachronic project, Fernando and I also started a project on the grammar of synchronic oral *corpora* (the NURC⁸)⁹. The first problem we faced was: what to do with the discursive elements. Should we sanitize the *corpus* the way Fries, the structuralist, did in the establishment of the English sentence patterns? We decided not to, as most studies

⁹ The project was part of a larger Project on Spoken Brazilian Portuguese, coordinated by Ataliba de Castilho (UNICAMP).



⁴ We planned the Project during a graduate course I gave with Tarallo, and the same students worked in the Project during another diachronic course I taught with Ian Roberts. Several Ph/D. Dissertations were developed during these courses.

⁵ Cf. Kroch (1994).

⁶ All the authors mentioned here were part of the Projects mentioned above and below, but, for lack of space, we did not mention all the authors and their works in the project. Two examples are: a) de Oliveira (year?), who wrote a dissertation comparing Italian and BP, on their similarities and differences concerning the NS parameter, and b) Nicolau ((1996) who worked with the relation between NSs and VS order in BP.

⁷ A second diachronic volume is soon going to be published, organized again by myself and Ian Roberts, this time, with the collaboration of Charlotte Galves (also as an editor).

⁸ NURC= Norma Urbana Culta (educated urban norm)

on Brazilian Portuguese (BP) are methodologically limited in that they provide a partial and idealized transcription of the data, being based only on portions of the *corpus* that are relevant for a specific syntactic or discursive phenomenon. The project had, therefore, the daring aim of studying how the syntax of predication and complementation integrated syntactic adjuncts and how these elements interacted with discursive ones. Thus, while in the diachronic project, the emphasis was basically on the expression and position of arguments, or the syntax of the main constituents of the sentence, the first object of our spoken BP project was adjunction. The NURC Project, coordinated by Ataliba de Castilho, was split into several subgroups, which produced a collection of eight volumes containing the proceedings of each meeting, but each project also produced specialized volumes. Our group published a volume called *A Construção da Sentença* (cf. Kato and Nascimento 2009)¹⁰, containing from predication to adjunction, and wh-constructions, all chapters developed within the Principles and Parameters framework and using the NURC *corpus*¹¹.

The third project, which I coordinated with João Peres (University of Lisbon), a formal semanticist, was a bi-national comparative study of European (EP) and Brazilian Portuguese (BP)¹². We used a computerized corpus of written language, but trying to cover different "genres," like newspaper editorials, letters from readers, interviews, plays and novels¹³. We have revealed not only the differences in the writers' E-language, better revealed in informal texts, but also the written norms prevalent in each country, which are better captured in formal texts. But through the writers E-language we were able to reveal their I-language. In this project, besides formal syntacticians, we also had formal semanticists. Many of the topics have been analyzed in co-authorship by both Portuguese and Brazilian linguists. The results of the project were published, not only in the *Proceedings* of the meetings, but also as two special issues of *Journal of Portuguese Linguistics*, edited by João Peres and myself¹⁴.

The project that I am presently coordinating is called *Romania Nova*, whose aim is not only to study BP, a Romance variety spoken in America, as opposed to EP, but also Latin-American Spanish, which has also shown interesting changes when compared to European



 $^{^{\}rm 10}$ Several authors were sponsored by a CNPq research grant.

¹¹ A critical survey of the spoken BP can be read in Kato (2002a).

¹² The Project was named Português Europeu e Português Brasileiro: unidade e diversidade, which had the sponsorship of CNPQ and ICCTI.

¹³ This computerized data-base of Portuguese is available on the internet (www.cgi.portuguese.mct.pt), and it is invaluable for the data analysis since It contains data from both the Portuguese and Brazilian varieties.

¹⁴ See Journal of Portuguese Linguistics, vols (3,1) and (4,1).

Spanish. The first book with its results has been published in 2018 by Oxford U. Press, edited by myself and Francisco Ordoñez, who has coordinated the Project with me for eight years 15.

In addition to my projects, Charlotte Galves, from UNICAMP, also conducted a diachronic project, whose major aim was to construct a syntactically annotated corpus¹⁶.

3. Some of the topics studied in the Projects

All the Projects that I have coordinated approached well-known parameters: a) the null subject parameter (NSP) b) the *wh*-movement parameter (WH-P), and c) the null object parameter (NOP)¹⁷. The three topics were a constant through all the projects, but we have also covered many more topics during all these years. In the present article, however, I will limit the discussion of the results to the NSP only.¹⁸

3.1. The Null Subject Parameter

The Diachronic project started investigating the well-defined Null Subject Parameter, in its macro-parametric view as a cluster of properties¹⁹, including:

- [a] the possibility of null subjects;
- [b] free inversion / postposed subjects /Romance inversion
- [c] clitic climbing (Kayne, 1989)

The correlation between NS and inversion has been challenged by Safir (1982), who found out that Northern Italian dialects, while disallowing the former, license the latter. In our first diachronic project, we found out that BP has been exhibiting a change in progress in both properties ([a] (Duarte 1993, 1995) and in [b] (Andrade Berlinck 1995; Torres-Moraes 1993), a fact that can be used in favor of properties of a single parameter (cf. Roberts 1993). Duarte

¹⁹ We do not have here the classic cluster in Chomsky (1981) and Rizzi (1982), but we added a third one included by Kayne (1989). (que tal: We're not referring to the classic cluster in Chomsky...?



¹⁵ Cf. Kato and Ordoñez (eds) 2018. The Project Romania Nova is now being coordinated by three members: Francisco Ordoñez (UStonybrooke, NY) Andrés Saab (Universidade de Buenos Ayres), and Mary A. Kato (UNICAMP).

¹⁶ http://www.tycho.iel.unicamp.br/~tycho

¹⁷ Also called the Hot-Cool Parameter.

¹⁸ Regarding the NOP, the first comparative work between EP and BP can be read in Kato & Raposo (2001), and a thorough review can be read in Cyrino (2018). Regarding the WH-P, the diachronic analysis can be read in Lopes Rossi (1993), and the most recent results can be read in Kato (2017; 2018)

(1995/2000) shows that from a preference of null subjects in 80% of cases in the middle of the 19th century, the situation is found reversed at the end of the 20th century, namely a little more than 20% of null subjects, a clear indication of a change in progress. As for VS order, Torres Moraes (1993) finds that from the beginning of the 19th century to the present date, the XVS form in declaratives decreases from 20% to \emptyset %, while the VSX, which was already rare in the 19th century (6%), becomes totally absent in modern plays. Andrade Berlinck (2000) shows that the residual VS that was possible in the 20th century, was only with unaccusative verbs. But Kato (2000) shows that in the 20th century, BP reanalyzed unaccusative verbs as existential, with VS no longer exhibiting agreement.

The theory of parameters provides us with another correlation which has not been much explored, namely the fact that NS languages have clitic climbing (Kayne 1989). In our diachronic project, Pagotto (1993) and Cyrino (1993) studying the position of clitics in the history of BP, find that, like present day EP, the clitic in classic BP used to be enclitic to the verb, while in Modern BP it is proclitic to the main verb. Nunes (1993) attributes this difference to different directions of cliticization: while EP cliticizes to the left, BP cliticizes to the right (cf. (1)).

```
a. Maria pode me encontrar hoje.
(1)
         Maria can-clitic find
                                   today
                                           (enclisis to the Modal) (EP) (Classic BP)
               1
                                           (proclisis to the main verb) (Modern PB)
       b1. Maria não me pode encontrar hoje.
                                           (enclisis to negation) EP
                                                                       (Classic BP)
                 Γ
                         1
                      ſ
                                           (proclisis to the main verb) (Modern BP)
       (Lit.: Maria cannot find me today)
```

In other words, BP corroborates Kayne's theory, which defends that clitic climbing is a property of the NS property, since the loss of enclisis in BP (and clitic climbing) also involved the loss of NS properties.

Though the empirical results corroborate the progressive loss of the three properties (null subjects, VS order and clitic climbing), the facts regarding Brazilian Portuguese also present challenging problems for the theory of parameters. Thus, though Duarte (1995 and 2000) draws on corpus to show a significant decrease in the use of referential null subjects (ex. (2), she also shows that the change affected mainly the first and second persons, leaving a still



reasonably stable use of null subjects in the third person (exs (3), especially the non-referential ones.

(2) Eu/\emptyset já te contei sobre a foto que eu vi uma vez? already you told about the photo that I saw once "Have I told you about the photo that I saw once?" b. Você/∅ podia estudar com o Nilson,... you could study with the Nilson "You could study with Nilson" Ele/*Ø fala bem inglês. 20 c. he speak well English "He speaks English well." (3) Ø Está chovendo. a. is raining "It is raining" Antigamente Ø punha a mesa pra tomar lanche. b. formerly laid the table to have snack "Formerly, one laid the table to have a snack" Ninguémi sabe o que Øi quer. c. nobody knows what wants "Nobody knows what he wants"

Challenged by this problem, Brazilian researchers start to work on it. At the time there were two theoretical lines to interpret the problem; a) the one that treated the present state as an intermediate diachronic stage, with the prediction that non-referential subjects would also end up filled by a pronoun, like in French (Cyrino, Duarte and Kato 2000)²¹, and b) the one that treats some of the existing null subject as a variable (ex.(4a) (Negrão & Muller 1996; Figueiredo e Silva 2000, Modesto 2000) or a trace (Ferreira 2000, Nunes 2008; Martins & Nunes, 2009) (exs. (4ab b).²²

²² Most lines of research can be found in the articles in Kato & Negrão (eds) (2000).



²⁰ This sentence is possible with a Null Subject as a short elliptical answer for the question "Que língua ele fala bem?" (What language does he speak?")

²¹ Cyrino, Duarte and Kato (2000) propose a referential hyerarchy, which predicts the route of the change from more referential categories to less referential ones, a hierarchy that also predicts types and sub-types of NSs.

(4) a. Ninguémi acha que Øi é burro.nobdy thinks that __ is stupid "Nobody thinks that he is stupid"

 $b. \ Eles_i \ parecem \ que \ t_i \quad estão \ cansados.$

they seem -3pPl that tare tired "They seem to be tired"

The first line claims that parameters are morphologically dependent, and that the non-uniform behavior of null subjects derives from the fact that morphology is not regular, or that speakers may present variation across generations (Duarte 1995). The second line presupposes a more radical explanation: Brazilian Portuguese has completely lost the null **pronouns**²³. Their argument is mainly based on the fact that null subjects in BP cannot occur in islands. A third line was due to acquisition studies (Magalhães 2000; Kato 2005; Martins and Nunes 2009²⁴). Based on Magalhães' (2000) acquisition findings, Kato (2005, 2011) claims that as null referential subjects are not found before schooling, they are not in BP core grammars, and are learned through instruction and not selection, constituting a second grammar. A child in preschool years would produce the following contrasting sentences:

- (5) a. Ninguémi acha que elei é burro.
 - b. Elesi parecem que elesi estão cansados.

The Project in which we compared BP and EP supported the last hypothesis. According to Barbosa, Duarte and Kato (2005), BP prescriptive grammar tries to preserve old forms through stylistic rules learned during schooling.

The starting point of the EP vs BP was Kato's (1999), a theory of the NS parameter in the Minimalist framework. She proposes that the grammatical subject of the NS languages, like EP, is the agreement affix itself and the apparent lexical subject is in a dislocated position in a Clitic Doubling construction (6a. and b). BP is like French (7a), as agreement is no longer pronominal, and the weak pronoun is the subject, which can be doubled by the strong one:

(6) a. Ele comeu a pizza.

he ate the pizza (="Him, he ate the pizza")

b. [ELEi TP come-ui VP ti a pizza]]] EP

²⁴ For Martins and Nunes, in their core grammar children have pronouns, but not NSs. Later they discover that T can have incomplete phi-features, which makes the complement subject move to get case in the upper clause.



²³ See Kato (2009) for a thorough critical review of these authors.

(7) a. Lui, il a mangé la pizza. FR
b. [ELE [ele comeu a pizza.]] BP

BP created a paradigm of weak pronouns quasi-homophonous with the strong ones, and doubling today sounds like repetition:

(8) a. Eu, eu não como pizza. (=Me, I don't eat pizza")

b. Você, cê gosta de pizza. (=You, you like pizza)

Kato's theory predicted that:

- a) EP rules out such doublings; in BP doubling is extremely common;
- b) when the third person pronoun is expressed in EP, it is [+human] because it is a strong pronoun; when the third person is expressed in BP it does not have to be [+human] because it can be either a strong pronoun or a weak one.

One case of third person null subject studied in the EP vs BP Project were the constructions like (9c), first noticed by Galves (1989). In our Project we show that, while in EP the arbitrary subjects are generally coded through the clitic se with agreement as in (6a), in BP, with the on-going loss of third person clitics, such a construction is replaced by several other possibilities: *se* without agreement; null clitic and personal pronouns *cê* and *a gente*.

(9) a. Não **se** usam mais chapéus. EP BP not *se* wear-3PP1 any-longer hats

"You don't wear hats any-longer."

b. Não **se** usa mais chapéu(s) %EP BP not *se* wear-3PSG any-longer hat(s)

c. Não \varnothing_i usa mais chapéu. *EP BP not wear any-longer hat

d. Cê não usa mais chapeu. "You don't wear hat(s)any longer you not wear any-longer hat "EP BP

e. **A gente** não usa mais chapéu" "We don't wear hat(s) any longer. we-folks not wear any-longer hats ?EP BP



BP makes use of bare nouns to avoid agreement with se.

(10) a. Admite-(se) estrangeiro.
admitted se foreigners "Foreigners admitted"
b. Vende-(se) livro usado. "Used books for sale sell se book used

This shows again that constructions with se+ agreement (or passive se) are preserved in written language and formal genres in BP as opposed to vernacular EP, which retains this construction uniformly and productively (Nunes 1990). Sometimes one can find variation of these forms even in the discourse of single individuals (cf. Kato, Duarte & Barbosa 2003)

```
(11) Parte-se um ovo; () põe a gema; () espa/ ah ah; () derrama em cima e () põe breaks-se an egg; () put the yolk; () spre/ah ah; () spill on top and () put [.......] () retira os dois e serve-se.
[......] () take-away the two and serv-se (NURC-D2 POA)
```

Notice the variation between *se* and null [], and also the overt *se*, always enclitic, in the beginning and end of the discourse, where speakers are more sensitive to the written style.

The conclusions in KDB' (2003) study are the following:

- a) the loss of the passive *se* in BP was caused by the change in the NSP, since only Romance languages with NSs exhibit such an agreement;
- b) the possibility of indefinite NSs without *se* can be attributed to the loss of 3rd person clitics in BP (Tarallo 1983, Nunes 1990, and Cyrino 1994); The empty clitic *se* is present, however, in LF, and erased at PF (Kato 1993). This variant is favored when it co-occurs with bare nouns;
- c) BP has only two types of NSs: a) the null expletive (example (4c), b) the null 3rd person argument, which is subsumed under the empty clitic *se*, considered an anaphor in the sense of (Reinhardt and Reuland 1993)²⁵.

During the EP vs BP bi-national Project, Romance inversion also deserved attention (cf. Britto (2000). Since in EP word order was used to distinguish types of predication, the author asks herself how such a distinction is codified after BP lost the order VXS. Inspired by Kato (1988) and Martins (1994), she makes use of Kuroda's (1972) distinction beween *thetic* and



²⁵ See Kato (2009) for a detailed discussion of their conception.

categorical types of judgments in natural languages. Working with Romance, Kato and Martins had proposed that VXS and SVX in Romance were used to codify thetic and categorical judgements, respectively. Britto discovers that, except with unaccusative verbs, which retain the contrast VS and SV, BP starts to have the latter for thetic sentences, while Left Dislocation (LD) was the preference structure for categorical judgments ²⁶.

(12)	a. Telefonou a Maria.	EP	•	(VS) (Thetic)
	b. A Maria telefonou.	EP		(SV)(Categorical)
(13)	a. A Helena telefonou.	BP		(SV)(Thetic)
	b. A Helena, ela telefono	ı. BP		(LD-SV)(Categorical

Formally, if we consider that the preverbal subject, in EP SV sentences, occupies not Spec, TP but a higher Σ P projection, as proposed in Martins (1994), the difference between the *categorical* sentences in EP and in BP can be claimed to be the result of the null or pronominal element in Spec, TP. In both cases the subject would be higher, or dislocated (Kato 1999, 2000).

(14) a.
$$[\Sigma P \ A \ Maria_i [TP \ \emptyset_i \ [T-u \ telefon-v [vP \ t_i \ tv]]]$$
 (Categorical) EP b. $[\Sigma P \ A \ Maria_i [TP \ ela_i \ [T-u \ telefon-v [vP \ t_i \ telefon-]]]$ (Categorical) BP

We could propose, as a consequence, that if we derive VXS from SVX in BP what we get as the result is not the VXS construction but the right dislocation one:

```
(15) a. [ [TP Ø<sub>i</sub> [T-u telefon-v [vP t<sub>i</sub> tv ]]] [ΣP A Maria<sub>i</sub> ]]
a' Telefonou a Maria.
b. [TP ela<sub>i</sub> [T-u telefon-v [vP t<sub>i</sub> tv]]] [ΣP A Maria<sub>i</sub> ]]
b' Ela telefonou a Maria.
```

Here I am proposing that the Romance inversion construction can be derived if we move the whole TP to a sentence initial position. Notice that what we get in BP is not the Romance inversion construction, but a right dislocation construction. This would explain why BP lost the Romance inversion.

While in the EP vs BP Project, we were attracted by the left dislocation construction codifying categorical sentences in BP, another construction that called our attention in BP were the so-called *topic-subjects*, in which subjects are definite, topic-like elements, which appear through raising of some sentential element to the position of subjects triggering agreement

Emeritus Miriam Lemle's career



²⁶ See also Kato & Martins (2009) for a BP&EP comparative work on thetic and categorical patterns.

(Duarte 2004; Kato & Duarte 2005). For many linguists, the constructions are attributed to languages of the Topic Prominent type, which do not dispose of expletives ²⁷.

(16) a. Chove	nuito nestas	cidades.	BP	EP
rain3PSG a lot	in-these cit	ies		
b. Estas cidades ch	novem muito		BP	*EP
these cities rai	n3PPL a lot			
(17) a. Furou o pne			BP	EP
went-flat the tyr	e of my car			
b. O carro furou	o pneu.		BP	*EP
the car went-flat	the tyre			
(18) a. Faltou se	orte aos m	eus times.	BP	EP
lacked-3PSG lu	ick to-the my	y teams		
b. Os meus times	faltaram	sorte.	BP	*EP
the my teams	lacked-3PP	luck		

According to Duarte (2004), the tendency in BP was not only to fill referential null subjects with pronouns, but also to replace null expletives with raised elements, since Topic Prominent languages do not dispose of expletives. The Portuguese linguist Costa ((2010) finds out that the data he collected, mostly from the Portuguese internet, can be considered the same phenomenon, namely instances of topic-subjects. Kato & Duarte (2017) show, however, that raising in EP is to an A'-position, triggering no agreement, while in BP raising is to an A-position, with agreement. In the singular, Dislocation and Topic-subjects look the same:

(19) a. Estas florestas # chove muito no verão.	EP ?BP
these forests rainPSG a lot in-the summer	
b. Estas florestas chovem muito no verão.	*EP BP
these forests rain3pPl a lot in-the summer	
c. Esta floresta chove muito no verão.	EP BP
this forest rains a lot in the summer	

It seems, however, that EP has also adopted the existential verb *ter* (besides *haver*), a verb that accepts personalization not only in BP, but also in EP (cf. Costa, 2010).

²⁷ The idea was first inspired by Pontes (1987), a functionalist, and Duarte, a sociolinguist, but formal syntaticians like Galves (1998), Negrão &Viotti (2000), and Modesto (2008), among many others, also adhered to this idea.



(20) a. Tem muita corrupção no país.

has much corruption in-the country

"There is much corruption in this country

b. Nós temos muita corrupção no país.

we have much corruption in-this country

"We have much corruption in this country."

More recently this sort of construction with topic subjects deserved the attention from linguists working with language contact, which is the case of Avelar and Galves (2011 and 2014), who propose that these topic-subject constructions were due to the contact with Bantu languages.

As Caribbean Spanish also lost the null subject, Kato & Ordoñez (forth.) went looking for such constructions in the *Romania Nova* Project, with the hypothesis that similar constructions might be found in the variety that maintained the null expletive, like in Dominican Spanish (DS)²⁸. No topic-subject constructions were found. What we found instead were Clitic Left Dislocation (CLLD) constructions. The conclusion was that in the context where BP has topic-subjects, Dominican Spanish has CLLD.

(21) a. Les faltó suerte a los equipos. DS

3PCl lacked luck to the teams

b. A los equipos les faltó suerte. CLLD DS

c. * Los equipos faltaram suerte. Topic-subjects DS

Kato and Ordoñez raised the following hypothesis concerning the appearance of topic-subjects in BP. As was shown by Cyrino (1993, 2018) and Pagotto (1993, 2018) the clitic paradigm in BP underwent the loss of third person clitics, but preserved the first and second person ones, a fact that was shown in the diachronic Project. With the loss of third person clitics, BP developed the topic-subject constructions for the third person to replace the *categorical* function of CLLD. Independent evidence is presented in Kato and Ordoñez for first and second persons: topic-subject constructions are not possible, and what appears instead are CLLD constructions. The logic of this is that first and second persons did not lose clitics in BP, and, therefore, continue to exhibit CLLD, dispensing the creation of topic-subjects.

(22) a. Pra mim me faltou sorte. CLLD BP



²⁸ A corpus was not used, as in other projects, but questionnaires with informants.

	to me	CL-1p lack	ked-3ps luck		
b. ³	*Eu	faltei	sorte.	Topic subject	*BP
	I	lacked-1psg	luck		
(23) a.	Para v	você te ²⁹	faltou sorte.	CLLD	BP
	for y	ou Cl-2nd la	cked luck		
b.	* Você	faltou sorte.		Topic subject	*BP
	you	lacked luck			

Considering the change from a functional view, we realize that both dative CLLD construction in Spanish and the Topic Subject construction in BP are *categorical* sentences, while the impersonal constructions with null or lexical expletives are thetic sentences in Kuroda's (1972) sense. Thus, DS maintained the categorical CLLD construction while acquiring lexical expletives for thetic impersonal sentences, BP changed the CLLD constructions to Topic-Subject sentences, while also maintaining the null expletive for thetic sentences.

We formalize the typological phenomenon of *topic*-prominence *vs subject*- prominence using Miyagawa's (2010) proposal, known as the "feature inheritance approach." The author follows Chomsky (2005), according to which φ-features are not merged in T, but are merged in C and percolate to T during the derivation. With this formulation, he proposes to formally translate Li and Thompson's functional typology of Topic prominent languages vs Subject-prominent languages. According to him, both agreement features and discursive features can be inherited by T from C. He also raises the possibility of the existence of mixed languages that inherit both types of features, like Bantu, and we would add BP.

(24) a. C,
$$\phi \to T\phi$$
 ... (agreement-prominent language - e.g. English)
b. C, $\delta \to T\delta$... (discourse-prominent language - e.g. Japanese)
c. C, δ , $\phi \to T\delta$, ϕ (mixed type of language -e.g.Bantu, BP)

In agreement Prominent language, like English, what takes a constituent to Spec, T is what agrees with the verb. The discursive role of the subject does not matter:

(25) a. Hamilton won-φ the race.	(Thetic or Categorical)
b. Hamilton had his tyre flat.	(Categorical)
b. A German won- φ the race.	(Thetic)

²⁹ Brazilian Portuguese uses the original second person clitic te with você (Você acha que eu te amo?)



In a Discourse Prominent language, like Japanese, it is an argument with a discursive role, marked morphologically, that determines what obeys the EPP. If the sentence is categorical the subject is marked -wa (=Topic) while if the sentence is thetic, the subject is marked -ga, in the nominative case.

(26) a. Hamilton-wa kiosso-o kata (Categorical)

Hamilton-top race-accus won

b. Hamilton-wa tiya-ga pankushita. (Categorical)

Hamilton-top tyre bursted

c. Aru doitsugin-ga kiosso-o kata. (Thetic)

a German-nom race-accus won

Topic and subject prominent languages, on the other side, can percolate both discursive and ϕ -features to T.

(27) a. Hamilton ganhou a corrida. (Categorical or Thetic)

Hamilton won the race

b. O Hamilton furou o pneu. (Categorical)

the Hamilton burst the tyre "H. had his tyre flat."

c. Furou o pneu do Hamilton. (Thetic)

burst the tyre of-the Hamilton "H.'s tyre went flat."

c. Um alemão ganhou a corrida. (Thetic)

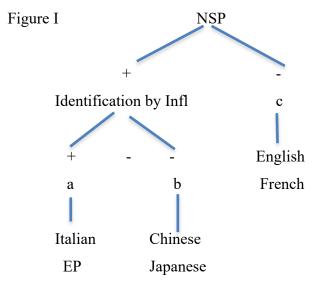
a German won the race "A German won the race."

The Bantu languages, especially Kinande, have the same type of topic-subject sentences of BP. What is important to register is that these topic-subject constructions are described in Miyagawa's theory without resorting to the NS parameter.

3.2. The recent discussion on the typology of NS languages

In Kato and Negrão (2000) we represented the division found for the NSP in the literature until then. The Figure below represents the state of art until then:





But in 2005, Holmberg launches the concept of "partial NS language", and later Holmberg and Sheehan (2010) propose to include BP in this group, with other languages like Finnish. The list of properties of such "partial" NS languages would be: a) the null subject is non-thematic (28a): b) the NS is a generic subject of the *one* type in English (cf. ex.(28b)) and c) the referential NS is necessarily controlled by an argument in a higher sentence (cf. (28c)). If the main clause contains *dicendi* or similar verbs, we can designate such subjects as logophoric ³⁰.

Our first problem is that Japanese, normally classified with languages without inflection, in Fig I above, has exactly a parallel behavior with regard to the sentences below.

(28) a. Ø Está chovendo.
Ø is raining
"It is raining"
b. Ø_{gen} Não usa mais chapéu.
BP

b. \emptyset_{gen} Não usa mais chapéu. BP \emptyset_{gen} not wear any-longer hat "You don't wear hats any longer."

c. João_i disse que Ø_i fala bem espanhol. BP

John_i said that Ø_i speaks well Spanish

"John_i said that hei speaks Spanish well."

³⁰ "Logophoric subject" in Kuno's (1972) sense is the subject of a direct discourse. (i) *John said "I like Mary"*. -> *John said that he liked Mary*.



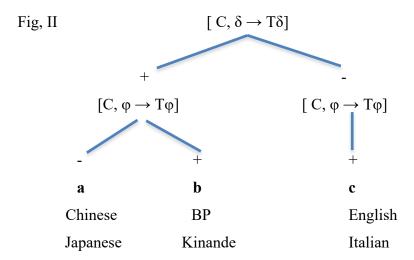
EP, a non-partial NS language, differs from Finnish and BP in that: a) the generic subject requires the clitic *se*, and b) the NS in the complement clause can have either the controlled reading by an argument in the upper clause or any other antecedent in discourse.

(29)	a. Ø Está chovendo.	EP
	b. \mathcal{O}_{gen} Não se usa mais chapéus.	EP
	c. Joãoi disse que Øi/j fala bem espanhol.	EP

What remains to be investigated, however, is whether Finnish behaves like BP with regard to *categorical* sentences. In other words, can Finnish have the type of raising that we saw for BP, namely topic-subjects (exs 30a)? Recall also that the generic NS in BP also allows pronominal variants with *você* or *a gente* (ex. 30b). Moreover, the controlled subject in (30c) can be a pronoun and not a NS:

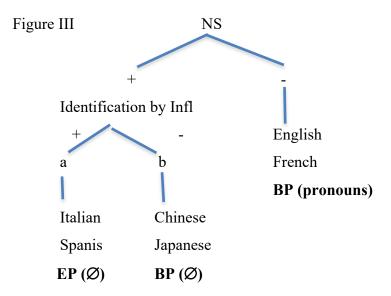
(30)	a. São Paulo está chovendo.		
	São Paulo is raining		
	b. Você/ a gente não usa mais chapéu.	BP	
	you/ we folks not wear any-longer hats		
	c. $João_i$ disse que ele_i fala bem espanhol.	BP	
	John said that he speaks well Spanish		

However, in the discussion on topic-subjects, many linguists emphasize the difference among languages regarding a different sort of parameter, [+ or – topic-prominence], which we can translate formally through an "inheritance" parameter, following Miyagawa (2010). Representing this typology in a Figure similar to the one above, we can have:





But notice that the ${\bf c}$ type does not tell English, a -NS language, from Italian, a prototypical +NS language. This shows that the NSP is independent from the Topic Prominence ([$C, \delta \to T\delta$]) parameter. But Kato and Duarte (2017) modifies Fig I to show the partial nature of BP regarding the NSP:



Empirically the NS in both Japanese and BP can be a: a) an expletive; b) a generic subject and c) a logophoric subject.

But what the last two Figures show is that BP is similar to Japanese and Chinese, regarding the inheritance parameter, but differ regarding the NS parameter:

- a) Japanese and BP are positive with regard to δ -feature inheritance;
- b) BP is also positive with regard to φ-feature inheritance, but Japanese is not.

Conclusions

The four projects, whose descriptions I briefly presented here, made it possible for us to learn a lot about BP, but many questions and problems still remain unanswered. The topic I selected for discussion, the NSP, is still under scrutiny in the linguistic literature, especially after the introduction of the concept "partial NS." Our Projects developed at UNICAMP may have contributed to understanding the concept, or may have added more fuel to the fire.

I would like to finish this article quoting Chomsky in the cover of Yang's (2002) book:



"Yang's carefully crafted and impressively argued study is a significant contribution to the exciting field of language acquisition, and cognitive growth more generally. He has shown how conceptions of "universal grammar" can be integrated with precise ideas about learning/growth mechanisms within a variationist model of language acquisition that yields sharp and well-confirmed predictions over quite an interesting range, with implications for language change as well. It is a substantial achievement, sure to become the basis for much fruitful inquiry and research.

We were thought to be heretic generativists, or variationists back in the day, but our master tells us that we are on the right track in embracing such perspective in studies of language change.

REFERENCES

ANDRADE-BERLINCK, Rosane (1995) La position du Sujet en Portugais. Université de Louvain: Ph.D. Dissertation.

AVELAR, J. & GALVES, C. (2011) Tópico e concordância em português brasileiro e português brasileiro. *In*: COSTA, A.; FALÉ, I. & BARBOSA, P. (eds). 49-65. *Textos Seleccionados, Actas do XXVI Encontro Nacional da Associação Portuguesa de Linguística*. Porto, APL, pp. 49-65.

_____ (2014). O papel das línguas africanas na emergência da gramática do português brasileiro. *Linguística – Revista da ALFAL* 30(2): 241-288.

BARBOSA, Pilar, M. Eugenia DUARTE & Mary A. KATO (2001) A distribuição do sujeito nulo no português europeu e no português brasileiro. *Actas do XVI Encontro Nacional da Associação Portuguesa de Lingüística* (APL 2001) Lisboa, 2001: 539-550.

(2005) Null subjects in European and Brazilian Portuguese. In: João Peres & Mary A. Kato (eds) *Journal of Portuguese Linguistics*, (4,1), p.11-52.

BRITTO, Helena (2000) Syntactic codification of categorical and thetic judgments in Brazilian Portuguese. In: Kato & Negrão (2000), 195-222.

CHOMSKY, Noam. (1981) Lectures on Government and Binding. Dordrecht: Foris.

(2005) Three factors in language design. *Linguistic Inquiry* 36: 1-22.

COSTA. João (2010) PE e PB: orientação para odiscurso importa? *Estudos da Lingua(gem)*, v. 8, n. 1, pp.123-143

COSTA, A, I Falé & P. BARBOSA (2011) Textos Seleccionados, Actas do XXVI Encontro



Nacional da Associação Portuguesa de Linguística. Porto, APL.

CYRINO, Sonia M.L. (1993) Observações sobre a mudança diacrônica no português do Brasil: objeto nulo e clíticos. In Ian Roberts & Mary A. Kato (eds), 163-184.
(1997) O objeto nulo no português do Brasil - um estudo sintático diacrônico. Londrina: Editora da UEL.
(2018) The null object in <i>Romania Nova</i> . In: Kato & Ordoñez (eds) (2018).
CYRINO, S. M., M. E. DUARTE & M. A. KATO (2000) Visible subjects and invisible clitics in Brazilian Portuguese. In M. A. Kato & E.V. Negrão (eds.): 55-104.
De OLIVEIRA, Marilza (1996) <i>Respostas Assertivas e sua Variaçãoo nas Línguas Românicas.</i> UNICAMP: Ph.D. Dissertation.
DUARTE, M. Eugenia L. (1993) Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. In Roberts & Kato (eds), 107-162.
(1995) Evite Pronome"no Português Brasileiro. UNICAMP: Ph. D Dissertation.
(1998) Left –dislocated subjects and parametric change in BP. <i>Proceedings of the 16th International Congress of Linguistics. CD-rom. Syntax.</i> N. 219.
(2000) The loss of the "Avoid pronoun" principle in Brazilian Portuguese. In Kato & Negrão (eds), 17-36. M.
FERREIRA, Marcelo Barra (2000) <i>Argumentos Nulos em Português Brasileiro</i> . UNICAMP: Dissertação de Mestrado.
FIGUEIREDO SILVA, M. Cristina (2000) Main and embedded null subjects in Brazilian Portuguese. In Kato & Negrão (eds), 127-146.
GALVES, Charlotte (1984)) Pronomes e categorias vazias em Português do Brasil. <i>Cadernos de Estudos lingüísticos</i> ,7:107-136.
(1998) Tópicos, sujeitos, pronomes e concordância no português brasileiro. <i>Cadernos de Estudos Linguísticos</i> , 34, 1998, pp. 7-21.
HOLMBERG, Anders (2005) Is there a little <i>pro</i> ? Evidence from Finnish. <i>Linguistic Inquiry</i> , 36: 533-564
& Michelle SHEEHAN (2010) Control into finite clauses in partial null-subject languages. In: T. Biberauer, A. Holmberg, I. Roberts & M. Sheehan (eds) <i>Parametric Variation: null subjects in Minimalist theory.</i> Cambridge: Cambridge U. Press, 125-152.

JAEGGLI, Oswald and Carmen SILVA-CORVALÁN (eds) (1986) *Studies in Romance Linguistics*. Dordrecht: Foris Publications.



KATO, Mary A (1989) Sujeito e Topico: duas categorias em sintaxe? Cadernos de Estudos Linguísticos, 17, pp. 109-132.
(1993) The distribution of null and pronominal objects in Brazilian Portuguese. Linguistic Perspectives on the Romance Languages: Selected Papers from the XXI Linguistic Symposium on Romance Languages (Currents Issues in Linguistic Theory Series), 225-235, Amsterdam: John Benjamins.
(1999) Strong and weak pronominals and the null subject parameter <i>PROBUS</i> , 11,1: 1-38.
(2000) The partial pro-drop nature and the restricted VS order in Brazilian Portuguese. In Kato & Negrão (2000) pp, 223-258.
(2002a) La gramática del portugués hablado: reflexiones sobre el uso de la lengua. En Curcó, C., Colín, M., Groult, N. y Herrera, L.(eds) CONTRIBUCIONES A LA LINGÜÍSTICA APLICADA EN AMÉRICA LATINA. México:CELE-UNAM,. pp. 363-378.
(2002b) The reanalysis of unaccusative constructions as existentials. <i>Revista do GEL</i> . 157-186.
(2003) Desvendando a gramática do português brasileiro. In: E. Albano, M. I. H. Coudry, S. Possenti, T. Alkmin (orgs) <i>Saudades da Língua</i> . Campinas: Mercado de Letras, p. 261-272.
(2005b) A gramática do letrado: questões para a teoria gramatical. In: M. A. Marques: E. Koller: J. Teixeira: S. A. Lemos (eds) <i>Ciências da Linguagem: trinta anos de investigação e ensino</i> . Braga: CEHUM (U, do Minho) 131-145.
(2009) O sujeito nulo revisitado no português brasileiro. In: M. A. Torres-Morais & M.L. O. Andrade (eds) <i>História do Português Paulista: Série Estudo</i> , Vol II. Campinas: Ed. Da UNICAMP.
(2011) Acquisition in the context of language change: the case of Brazilian Portuguese null subjects. In: Esther Rinke & Tanja Kupisch (eds) <i>The Development of Grammar: Language Acquisition and Diachronic Change</i> . Hamburg Series on Multilingualism (HSM 11), 309-330. Amsterdam/ Philadelphia: John Benjamins.
(2017) A mudança na posição estrutural de foco no português brasileiro. <i>Revista Linguistica</i> , 13,2, pp158-173.
(2018) Mophological doublets in Brazilian Portuguese wh-constructions. In: Lori Repetti& Francisco Ordoñez (eds) <i>Romance Languages and Linguistic Theory 14. Selected Papers from the 46th Linguistic Symposium on Romance Languages. Benjamins, Stony Brook, NY.</i> Amsterdam/Philadelphia. John Benjamins.135-152



<u></u> & Ana M. MARTINS (2016) European Portuguese and Brazilian Portuguese: an Overview on Word Order. In: Leo Wetzels, Sergio Menuzzi & João Costa (eds) The Handbook of Portuguese Linguistics. Malde, Mass: WILEY Blackwell, 15-40.
& M. Eugenia DUARTE (2014) A variação entre construções finitas pessoais e impessoais no Português Brasileiro. <i>WEB-REVISTA SOCIOLETO</i> , v.4 n.12, 153-177.
(2017). O sujeito no Português Brasileiro e sua tipologia. In: E. Pilati, H. Lima Salles e R. Naves (eds) <i>Novos Olhares para a Gramática do Português Brasileiro</i> . Campinas: Pontes. p.13-42.
M. Eugenia L. DUARTE & Pilar BARBOSA (2003). Sujeitos indefinidos em PB e PE. <i>Boletim da ABRALIN (26):</i> pp. 405-409.
& Esmeralda NEGRÃO (2000) (eds) Brazilian Portuguese and the Null Subject Parameter Frankfurt: Vervuert-IberoAmericana.
& Eduardo RAPOSO (2001) O objeto nulo definido no português europeu e no português brasileiro: convergências e divergências. <i>Actas do XVI Encontro Nacional da Associação Portuguesa de Lingüística</i> (APL 2001), Lisboa, 673-685
<u></u> & Fernando TARALLO (1986). Anything YOU can do in Brazilian Portuguese. In: O. Jaeggli e C. Silva-Corvalán (eds) <i>Studies in Romance Linguistics</i> , Dordrecht: Foris, 343-358.
& Francisco ORDOÑEZ (2018) <i>The Morphosyntax of Portuguese and Spanish in Latin America</i> . New York: Oxford University Press.
& Francisco ORDOÑEZ (forth.) Topic subjects in Brazilian Portuguese and dative clitic left dislocation in Caribbean Spanish. <i>Syntax</i> .
& João PERES (eds) (2004) Journal of Portuguese Linguistics, v. 3, n. 1.
& João PERES (eds) (2005) Journal of Portuguese Linguistics, v. 4, n. 1.
KAYNE, Richard (1989) Null subjects and clitic climbing. In O Jaeggli & K. Safir (eds) <i>The Null Subject Parameter</i> . Kluwer: Dordrecht. 239-262.
KROCH, Anthony. Morphosyntactic variation. In K. Beals, et al. (eds) <i>Papers from the 30th Regional meeting of the Chicago Linguistic Society. Vol 2: The Parassession on Variation in Linguistic Theory.</i>
KUNO, S. (1972) Pronominalization, reflexivization, and direct discourse. <i>Linguistic Inquiry</i>

KURODA, S-Y (1972). The categorical and the thetic judgment. *Fondation of Language* 9, 153-185.



3(2): 161-195.

LOPES-ROSSI, M. Aparecida (1993) Estudo diacrônico sobre as interrogativas do português do Brasil. In Roberts & Kato (eds): 307-342.

MAGALHÃES, Telma (2000) *Aprendendo o Sujeito Nulo na Escola*. UNICAMP: Dissertação de Mestrado.

MARTINS, ANA M. (1994) *Clíticos na História do Português*. University of Lisbon: Ph.D Dissertation.

_____ & Jairo NUNES (2005) Raising issues in Brazilian Portuguese. *Journal of Portuguese Linguistics*. Lisboa, v. 4, p.53-77.

MIYAGAWA, S. 2010. Why Agree? Why move? Unifying agreement-based and discourse configurational languages. Linguistic Inquiry Monograph 54. Cambridge: Mass: MIT Press.

MODESTO, Marcelo (2000) Null subject without "rich" agreement. In Kato & Negrão

(2008) Topic prominence and null subjects. In: Teresa Biberauer

(eds). 147-176.

(Ed.) The Limits of Syntactic Variation. Amsterdam: John Benjamins. 75-406.

NEGRÃO, Esmeralda (1999) Português brasileiro: uma língua voltada para o discurso.

São Paulo, 1999. Tese de Livre Docência – FFLCH, Universidade de São Paulo.

<u>_____</u> & Ana L. MULLER (1996) As mudanças do sistema pronominal do Português do Brasil: substituição ou especialização de formas? *DELTA* 12: 125-52.

NCOLAU, Eunice (1996) *As Propriedades do Sujeito Nulo e da Inversão no Português Brasileiro*. UNICAMP: Ph.D Dissertation.

NUNES, Jairo M. (1990) O Famigerado SE: uma análise sincrônica e diacrônica das construções com SE Apassivador a Indeterminador: UNICAMP: MA. THESIS.

____ (1993) Direção de cliticização, objeto nulo e pronome tônico na posição de objeto em português brasileiro. In Roberts & Kato (eds): 207-222.

_____(2008) Inherent case as a licensing condition for A-movement: the case of hiper-raising constructions in Brazilian Portuguese. *Journal of Portuguese Linguistics*, 7: 83-108.

_____ (2018) Subject and topic hiper-raising in Brazilian Portuguese: A case study on reference sets for economy computations. In: M. A. Kato & F. Ordoñez (eds), 107-134.

PAGOTTO, Emílio (1993) Clíticos, mudança e seleção natural. In Roberts & Kato (eds): 185-206.



REINHARDT, Tanya & Eric REULAND (1993) Reflexivity. Linguistic Inquiry, 24: 657-720.

RIZZI, Luigi (1982) Issues in Italian Syntax. Dordrecht: Foris.

ROBERTS, Ian (1993) Posfácio: O Português Brasileiro no contexto das línguas românicas. In Roberts & Kato (eds), 409-425.

_____ & Mary A. KATO (eds) (1993) *Português Brasileiro: uma viagem diacrônica*. Campinas: Editora da UNICAMP.

TARALLO, Fernando & Mary. A. KATO (1993) Filling in syntactic boundaries in Spoken Brazilian Portuguese. *Language Variation and Change*. 5, 1: 91-112.

TORRES MORAIS, M. Aparecida. (1993) Aspectos diacrônicos do movimento do verbo, estrutura da frase e caso nominativo no português do Brasil. In Roberts & Kato

(eds), 263-306.

YANG, Charles (2002) *Knowledge and Learning in Natural Language*. New York: Oxford University Press.

