

On Indefinite Subject in Brazilian Web-Mediated Specialized Discourse¹

Gian Luigi De Rosa¹ 

Martina Desantis² 

ABSTRACT:

The present research, based on a corpus of Web-mediated Talk Videos (Academic Talk Videos and TEDx Talks), illustrates the representation of subjects with indefinite reference in Brazilian Portuguese (BP). Recent studies have demonstrated a preference for nominative pronominal forms to express generic and arbitrary reference, particularly in finite sentences of both educated urban speech and popular speech. This study, grounded in the theory of language change proposed by Weinreich, Labov, and Herzog (2006), combined with the theory of Principles and Parameters (Chomsky 1981, 1995), aims to investigate subject indefiniteness strategies in specialized web-mediated speech, using two corpora: an Academic Talk Videos corpus (semi-popularizing talks) and a TEDx Talks corpus (popularizing talks). Our main hypothesis is that, depending on the corpus, we should observe a more conservative grammar, with a preference for standard strategies, in the Academic Talk Videos corpus [-popularizing], [+monitored] and [+formal], and a less conservative grammar, with a preference for innovative strategies, in the TEDx Talks corpus [+popularizing], [-monitored], and [-formal].

KEYWORDS: *Brazilian Portuguese, Indefinite subjects, educated urban speech, Academic Talk Videos, TEDx Talks.*

¹Sections 2 "Subject indefiniteness in BP"; 3 "Web-Mediated Specialized Discourse" (and its subsection); subsection 4.2 "The Academic Talk Videos"; 5 "Results" and its subsection 5.1 "Subject indefiniteness in the Academic Talk Videos" are authored by Gian Luigi De Rosa. Subsection 4.1 "The TEDx Talks" and 5.2 "Subject indefiniteness in TEDx Talks" were written by Martina Desantis"; section 4 "Our corpus", the Introduction and the Final considerations were written by the two authors.

Bethânia Mariani
Editora-chefe dos
Estudos de Linguagem

Recebido em: 14/11/2024
Aceito em: 10/12/2024

¹Università degli Studi Roma Tre, Roma, Italia. E-mail: gianluigi.derosa@uniroma3.it

²Universidade Federal Fluminense, Niterói, RJ, Brasil. E-mail: martina.desantis1995@gmail.com

How to cite:

DE ROSA, Gian Luigi; DESANTIS, Martina. On Indefinite Subject in Brazilian Web-Mediated Specialized Discourse. *Gragoatá*, Niterói, v. 30, n. 66, e65369, jan.-abr. 2025. Available from: <https://doi.org/10.22409/gragoata.v30i66.65369.en>

Introduction

Recent studies on the representation of subjects with indefinite reference in Brazilian Portuguese (BP) have demonstrated a preference for nominative pronominal forms to express generic² and arbitrary³ reference, particularly in finite sentences within both educated urban speech and popular speech. Furthermore, these innovative forms are becoming increasingly implemented in finite sentences in standard written language (Kato; Tarallo, 1986; Duarte, 1993, 1995, 2000, 2003, 2007, 2012, 2015, 2016, 2020, 2023; Tarallo, 1993; Cavalcante, 1999, 2006; Duarte; Lopes, 2002; Barbosa; Kato; Duarte, 2000; Couto, 2005; Cavalcante; Duarte, 2009; Vargas, 2010, 2012; Marins; Soares da Silva; Duarte, 2017).

The gradual decline in the use of standard strategies for expressing subject indefiniteness in educated urban varieties of BP (neo-standard BP)⁴ – specifically, the indefinite SE (used as a passive or impersonal marker) and 3PP with a null subject – alongside the emergence and establishment of alternative, innovative strategies – nominative forms (preferably overt): VOCÊ, A GENTE, NÓS, and ELE⁵ – should be considered a side effect, according to Weinreich, Labov, and Herzog (2006, p. 110), of the shift in the null subject parameter (Chomsky, 1981, 1995). As a result of this shift, which does not affect standard BP, neo-standard BP would transition from a pro-drop language to a partial pro-drop language (Kato, 1999; Barbosa; Kato; Duarte, 2000; Cyrino; Duarte; Kato, 2000; Duarte; Soares da Silva, 2016).⁶ In fact, while European Portuguese (EP) and standard BP can be considered positively marked languages for the pro-drop parameter, showing a strong preference for null subjects in both defined and indefinite references, neo-standard BP shows a preference for filling the subject position with overt pronominal forms.

This study, grounded in the theory of language change proposed by Labov (1994), Weinreich, Labov, and Herzog (2006) and the Principles and Parameters theory (Chomsky, 1981, 1995), aims to investigate strategies for subject indefiniteness in specialized web-mediated speech. It utilizes two corpora: an Academic Talk Videos corpus (semi-popularizing talks) and a TEDx Talks corpus (popularizing talks).

Our primary hypothesis is that, depending on the corpus, we will observe a more conservative grammar characterized by a preference for standard strategies in the Academic Talks corpus, which is [-popularizing], [+monitored] and [+formal], and a less conservative grammar with a preference for innovative strategies in the TEDx Talks corpus, which is [+popularizing], [-monitored], and [-formal].

Subject indefiniteness in BP

In Portuguese and Brazilian normative grammars, an indefinite subject refers to a subject when “o verbo não se refere a uma pessoa determinada, ou por se desconhecer quem executa a ação, ou por não haver interesse no seu conhecimento” (Cunha; Cintra, 1985, p. 128).

²We identify the generic reference subjects, which correspond to those whose reference cannot be determined by the text/discourse, being unlimited, in the sense that it can refer to an infinite set of entities (...) (Marins; Soares da Silva; Duarte, 2017, p. 151).

³By arbitrary reference, we mean subjects that cover a finite set of entities or an individual, whose identification cannot be individualized in the discourse (...) (Marins; Soares da Silva; Duarte, 2017, p. 152).

⁴Neo-standard BP is the variety of BP in common use (educated urban varieties), employed by educated urban Brazilian speakers and which can be considered a new standard in formation, whose most prominent constructions, forms and realisations are also recorded in the most monitored spoken and written textual genres (De Rosa, 2011, 2012).

⁵These strategies indicate the possibility of pronominal forms - which in principle have a definite reference - to represent subjects with an indefinite reference, receiving “uma interpretação figurada por um processo que poderíamos chamar de metáfora de pessoa” (Ilari; Franchi; Neves; Possenti, 1996, p. 95).

⁶Languages that allow the omission of the grammatical subject, such as Italian, are called null subject languages (or pro-drop languages), while languages that don't allow the omission of the subject, such as English, are called obligatory subject languages (or non-pro-drop languages).

In this context, the only strategies for representing subjects of indefinite reference in finite sentences are: 3PP with a null subject (1 and 2); the SE+3PS clitic with indirect transitive verbs (3), intransitive verbs (4) or direct transitive verbs with intransitive use (5) or the co-occurrence and the concurrence of the two strategies (6):

- (1) *Contaram-me*, quando eu era pequenina, a história duns naufragos, como nós.
- (2) *Reputavam-no* o maior comilão da cidade.
- (3) Precisa-se do carvalho; não *se precisa* do caniço.
- (4) Ainda *se vivia* num mundo de certezas.
- (5) *Comia-se* com a boca, com os olhos, com o nariz.
- (6) Na Casa *pisavam* sem sapatos, e *falava-se* baixo.⁷

⁷The examples are taken from Cunha and Cintra (2008, p. 128-129).

Indeed, the Portuguese Brazilian grammatical tradition (Cunha; Cintra, 2008, p. 128-129; Bechara, 1977, p. 200; Rocha Lima, 1972, p. 236; Mateus et al., 2003, p. 283) primarily identifies these two strategies. Bechara (1977, p. 200) includes the 3PS with a null subject, but this is limited to the verb “dizer,” as in “Diz que eles vão bem,” and Inês Duarte (Mateus et al., 2003, p. 283) notes that for EP, the 2 PSgen with a null subject is also applicable, as in “Ajudas sempre os amigos e apesar disso eles criticam-te”.

The classification of the clitic SE with direct transitive verbs (3PS and 3PP) differs from the standard norms of grammatical classification. Normative grammar categorizes this as a passivizing construction; therefore, since it is a synthetic passive construction, it does not include this strategy among the subject indefiniteness strategies. However, Bechara (2004) includes the SE clitic with direct transitive verbs, regardless of agreement, in the category of subject indefiniteness strategies, as illustrated by the following examples: “alugam-se casas/aluga-se casa”.

[O] se como índice de indeterminação de sujeito – primitivamente exclusivo em combinação com verbos não acompanhados por objeto direto –, estendeu seu papel aos transitivos diretos (onde a interpretação passiva passa a ter uma interpretação impessoal: Vendem-se casas = ‘alguém tem casa para vender’) e de ligação (É-se feliz). A passagem deste emprego da passiva à indeterminação levou o falante a não mais fazer concordância, pois o que era sujeito passou a ser entendido como objeto direto, função que não leva a exigir o acordo do verbo:

Vendem-se casas (= ‘casas são vendidas’) → Vendem-se casas (= ‘alguém tem casa para vender’) → Vende-se casas

In this context, Antenor Nascentes (1960 [1938]), Said Ali (1965), and Mattoso Câmara Jr. (1976) argue that constructions featuring the clitic SE with direct transitive verbs should not be classified as passive. Similar conclusions are drawn by Raposo and Uriagereka (1996, p. 800), who regard the synthetic passive as an indefinite SE, and by Scherre (2005, p.

80), who states that “[a] forma verbal nas estruturas denominadas passivas sintéticas é variável e ocorre, segundo a tradição por “atração o por falsa concordância com o objeto direto” (...) em função do conhecimento da gramática normativa da língua portuguesa”. Therefore, in this analysis, we will also consider and categorize strategies involving the clitic SE (with or without agreement) as strategies of subject indefiniteness.

Contrary to the prescriptions of normative grammars for standard BP, the subject indefiniteness found in neo-standard BP reflect an ongoing change in the categorization of the null subject parameter. As previously mentioned, and supported by the cited research, the emergence of nominative pronominal forms for indefinite reference – whether generic or arbitrary – tends to favor overt expressions. This phenomenon can be seen as a side effect of the implementation of overt definite reference subjects. In fact, in neo-standard BP, both oral varieties [+/-monitored] and written [+/-monitored] and transmitted [+/-monitored] varieties, there is an implementation of overt subjects of definite and indefinite reference.

However, among the innovative strategies for expressing subject indefiniteness in neo-standard BP – in addition to the nominative pronominal forms (preferably overt): VOCÊ, A GENTE, NÓS and ELES – we note the use of an indefinite null subject Ø+3PS (Galves, 1987; Nunes, 1990, 1991; Martins; Nunes, 2016). This approach contrasts with the implementation of overt definite and indefinite subjects. In this context, Kato (2020, p. 83) connects this strategy to the loss of 3P clitics.

Com a perda dos pronomes fracos clíticos, um deles o se (...), o PB passa a codificar sentenças indefinidas e genéricas sem esse clítico, em sentenças de sujeito nulo, mas antes sofre a mudança de ênclise para próclise.

- a. Ø_{gen} conserta-se Iphone.
- b. Ø_{gen} se- conserta Iphone.
- c. Ø_{gen} conserta IPHONE.

Regarding this construction, Galves (1998, p. 87) states that in Brazilian Portuguese (BP), the null subject is quite productive in “generic clauses,” as in “Não Ø usa mais freio” and that this type of construction yields an indefinite interpretation only in BP, whereas in European Portuguese (EP), the clitic SE would be required to achieve an indefinite interpretation.

Duarte (2023, p. 32) connects this phenomenon to parametric change, while also noting that “o surgimento de um sujeito nulo arbitrário ou genérico, de certa forma, contradiz o curso de uma mudança em direção a sujeitos expressos”. However, regarding its origin, she argues that these strategies “podem ser reduções de formas na 3ª pessoa do plural (“Na televisão falam muito isso”) ou variantes de pronomes genéricos (“você/a gente está precisando de políticos sérios”/“pra beber, você/a gente tem que ter noção”).

In our analysis, following the work of Egerland (2003), Marins, Soares da Silva, and Duarte (2017), as well as Holmberg and Phimsawat (2017), we categorize neo-standard BP subject indefiniteness strategies into three distinct groups.

- I. 3PP_{arb/excl} SE+3PP and the pronoun ELES (preferably overt);
(excludes the speaker)
- II. 3PS_{+/-gen +/-incl} SE+3PS, Ø+3PS and the pronoun VOCÊ (preferably overt);
(may or may not include the speaker and the addressee)
- III. 1P_{gen/semi-incl} the pronouns EU, NÓS and A GENTE (with A GENTE considerably more frequent than NÓS);
(includes the speaker necessarily, but may or may not include the addressee)

In this tripartition, we considered arbitrary reference, which excludes the speaker, and generic reference, which includes the speaker (Egerland, 2003). We also examined the inclusion, semi-inclusion, or exclusion of the addressee (Holmberg, 2005, 2010; Holmberg; Phimsawat, 2017).

Web-Mediated Specialized Discourse

When we discuss Specialized Discourse, we consider all the aspects and various relationships that develop between ‘matter-language-body-action-dialogue-actors’ in diverse interactions. Specialized Discourse is defined as ‘language-in-texts-in-situations-in-cultural-contexts’, representing an integration of different components that must be taken into account when analysing specialized linguistic manifestations (Cavagnoli, 2007, p. 17).

The term “Specialized Discourse” is particularly fitting, as it encompasses the concept of three continua. The first two continua enable us to visualise a variety of text types and cybergenres - written, spoken, and transmitted - that circulate within specialized domains (Sabatini, 1999, 2016; Sobrero, 2006). The third continuum is typological, ranging from [+formal] to [-formal]. This continuum can be applied to web-mediated text genres based on genre typology, taking into account the degree of stylistic monitoring and control (Marcuschi, 2008; Bortoni-Ricardo, 2005).

The first continuum clearly highlights the heterogeneity that characterises Specialized Discourse, which encompasses three levels: a scientific level, consisting of more “specialized” texts aimed exclusively at experts (expert/expert); a semi-popularizing level, which includes texts that can be classified as scientific-didactic and are intended for educating the public (didactic sphere); and a popularizing level, comprising inherently informative texts aimed at the general public (expert/non-expert) (cf. Gotti, 1991; Sabatini, 1999, 2016; Sobrero, 2006; Dardano, 2008; Gualdo; Telve, 2011).

To analyse web-mediated semi-popularizing and popularizing texts, we will employ Sabatini's (1999, 2016) model for classifying text types, specifically the second continuum. This model categorizes text types based on the degree of interpretative bond that the sender establishes with the receiver, which reflects the levels of 'rigidity' introduced into the communicative pact between them. In this theoretical framework, the receiver, as the interpreter of the text, plays a crucial role in defining the text's nature through the pact established by the sender. This pact can determine the deep and superficial structure of the text. The primary parameter guiding the sender's behaviour is their intention to regulate and convey - in a more or less rigid (explicit) way - the receiver's interpretative activity. If the sender aims for the text to be interpreted as unambiguously as possible, they will design it to minimize opacity. The continuum between the extremes of 'rigidity' (maximum bond) and 'elasticity' (minimum bond) identifies three broad categories of text typologies: Rigid Texts, Semi-Rigid Texts, and Elastic Texts. These categories are distinguished by the presence or absence of specific linguistic features related to factors such as: 1) the overall structure of the text; 2) logical coherence; 3) the system of textual cohesion connectors (morphosyntactic, semantic, prosodic, and phonetic); 4) the use of various types of sentence constructions; 5) punctuation; and 6) the graphic aspect of the text (Sabatini 1999, p. 637).

Regarding the other two continua, Marcuschi (2008) states that text genres are numerous and defined by their socio-communicative functions. They are arranged along a typological continuum that ranges from [+formal] to [-formal], irrespective of whether they are oral, written, or transmitted. This continuum can be effectively applied to hybrid text genres, such as Academic Talk Videos and TEDx Talks, by considering the degree of stylistic monitoring and control exercised by the speaker, which varies according to genre typology. This perspective aligns Marcuschi's (2008) typological continuum of text genres with Bortoni-Ricardo's (2005) proposal, which posits that all linguistic choices are influenced by the variable degree of stylistic monitoring (Figure 1 and Figure 2).

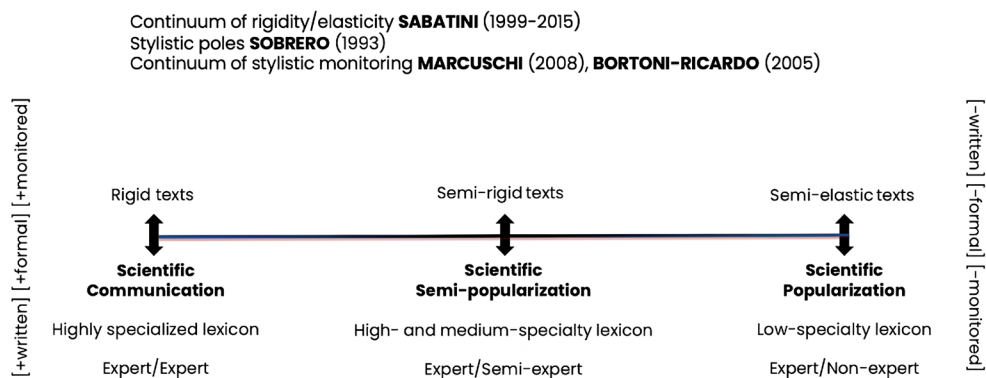


Figure 1. Web-Mediated text types in three continua (Sabatini, 1999, 2016; Sobrero, 2006; Marcuschi, 2008; Bortoni-Ricardo, 2005).

Source: own elaboration.

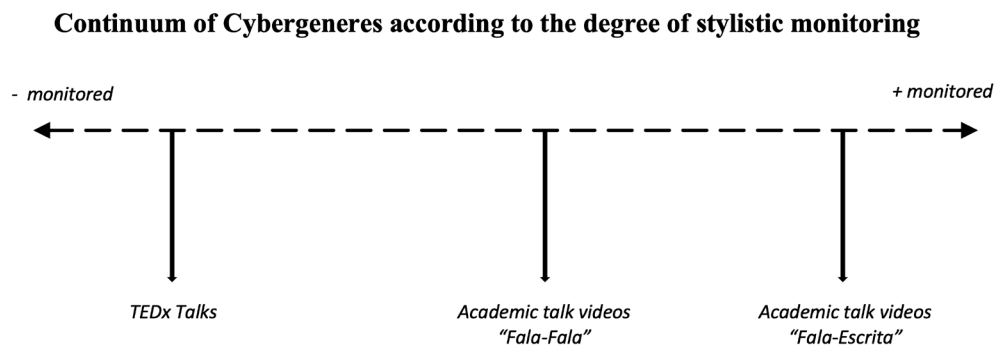


Figure 2. Cybergenres and stylistic monitoring.

Source: own elaboration.

Considering our sample, we position TEDx Talks on the far left of the [-monitored] continuum, Academic Talk Videos from the FALA-FALA Subcorpus in an intermediate [+/-monitored] position, and Academic Talk Videos from the FALA-ESCRITA Subcorpus on the far right [+monitored]. This arrangement reflects a greater emphasis on standard norms and their associated rules, as these videos represent a [+monitored] text typology, with the speaker utilizing a written text to support their speech.

In this paper, we use both the term Specialized Discourse and the term Language for Specific Purposes (LSP), understanding Specialized Discourse as a pragmatic concept, focusing on its production conditions, the speaker's intentions, and motivations. In contrast, LSP is viewed as an essentially linguistic concept. According to Gotti (1991, p. 8), an LSP is a functional variety of a natural language that relies on the presence of three elements:

- specialized users (experts);
- specific spheres of activity/realities/domains and fields of knowledge being referred to;
- specialized use of language.

Indeed, it is in scientific communication among specialized users (expert to specialized fields of knowledge and activities) that a Language for Specific Purposes (LSP) is utilized to its fullest extent.

Regarding the dimensions of variation, LSPs are articulated across five dimensions of variation. The two primary dimensions are the horizontal, or more precisely, diathematic dimension, which is based on content and highlights various disciplinary fields and subfields; and the vertical, or diachronic dimension, which is defined by registers. This vertical dimension emphasizes different levels determined by the communicative situation, the purposes and functions of the message, the context in which linguistic exchange occurs, and text typology (Cortelazzo, 1994; Gualdo, 2019, p. 57).

The other three dimensions that contribute to define the physiognomy of the LSP are:

- Diachronic variation (evolution over time).
- Diatopic variation refers to the differences in language use across various linguistic regions, which is particularly relevant when analysing a pluricentric language such as Portuguese.
- Diamesic variation, which highlights the medium and the transmission channel used for Specialized Discourse.

For the popularizing and semi-popularizing talk videos, it is important to note that these are web-mediated planned communicative events. In these contexts, speakers can use multimodal resources and an LSP that includes medium to high-specialty vocabulary (as seen in TEDx Talks) and high-specialty vocabulary (as found in Academic Talk Videos). Additionally, these modalities are conducive to the creation of neologisms and structuring texts (cf. Sobrero, 2006, p. 239).

Indeed, semi-popularizing and popularizing talk videos can be classified as hybrid genres due to their text typology and textual functions. They are characterized by the use of supposedly monoreferential language, in which univocal concepts are intended. The text is crafted to convey only one meaning per term, avoiding synonyms to eliminate ambiguities and ensure absolute clarity of content. This approach achieves clarity of concepts, allowing for direct and immediate comprehension.

A Language for Specific Purposes (LSP)

The specialized language observed in our sample of web-mediated talk videos, classified as a Language for Specific Purposes (LSP), is primarily characterized by the diathematic dimension of variation, which pertains to content, and the diaphasic dimension of variation, which is associated with the communicative context and the relationship between the speaker and audience. These dimensions significantly influence the selection of a specific linguistic register along the continuum [+formal] to [-formal], as well as particular lexical and syntactic features. Furthermore, given that this is a web-mediated specialized speech that is [+/-monitored] and [+/-formal], the diamesic dimension of variation – encompassing the medium, transmission channel, and language modality – emerges as a critical factor in the analysis of this linguistic variety.

As an LSP, it is presupposed – based on genre, typology, and the textual functions in which it is utilized – that web-mediated specialized speech, [+/-monitored] and [+/-formal], should exhibit the following characteristics:

- a high level of stylistic monitoring and control in discourse articulation. Consequently, characteristic features of spontaneous speech, including repetitions, false starts, reformulations, and discourse markers, should be absent.

- a formal or semi-formal register;
- extensive use of highly specialized technical-scientific lexicon;
- organised textual statements, structured as fixed phrases, with all verbal valences appropriately satisfied.

The analysis of our sample indicates that only the second characteristic (formal or semi-formal register) and the third characteristic (extensive use of highly specialized technical-scientific lexicon) are consistently present in all semi-popularizing talk videos. In contrast, popularizing talk videos tend to employ a lexicon of medium speciality, and the register is [-formal] and [-monitored] (De Rosa, 2020a, 2020b). This observation supports the argument that LSPs, which were predominantly written and highly standardized until a few decades ago, are now being shaped by the medium or transmission channel, thereby adopting features of neo-standard grammar. Gualdo (2019, p. 61) posits that the selection of a less “rigid” style for audiovisual genres and textual typologies of a popularizing level is influenced by at least three distinct factors:

- a. The entertainment function, which is predominant over the didactic function in audiovisual genres.
- b. The impact of social structures on the utilization of new media.
- c. The effectiveness of the Anglo-American communicative teaching model is increasingly supplementing the descriptive-prescriptive model in various educational contexts.

This phenomenon can be elucidated by the presence of various syntactic constructions and phenomena characteristic of the neo-standard BP within Web-Mediated Specialized Discourse, both semi-popularizing and popularizing. These include innovative strategies for rendering the subject indefinite.

Our Corpus

The quantitative and qualitative analysis presented herein is based on a sample of semi-popularizing web-mediated specialized speech (Academic Talk Videos) and popularizing speech (TEDx Talks) in neo-standard BP. For the examination of semi-popularizing specialized speech, we employed a selection of 27 Brazilian Academic Talk Videos sourced from the Monitored Academic Speech Corpus of Brazilian Portuguese (PB-FAM). This selection is categorized as follows: 12 Academic Talk Videos from the “Videoverbetes” series, the official channel of Encidis (Mariani, 2018, 2020), recorded between 2019 and 2024; 7 Academic Talk Videos from the “Linguística em minutos” series on the Linguistics and Teaching Project channel, recorded in 2022; and 8 Academic Talk Videos from the “O que é isso?” series on the Ciência na Rua channel, recorded in 2018. Additionally, for the popularizing specialized speech, we utilized 9 Brazilian TEDx Talks, recorded between 2019 and 2022.

In the Monitored Academic Speech Corpus of Brazilian Portuguese (PB-FAM), the Academic Talk Videos are categorized into two subcorpora: the FALA-FALA subcorpus and the FALA-ESCRITA subcorpus. However, to ensure comparability with data from TEDx Talks, only the Academic Talk Videos from the FALA-FALA subcorpus were selected for this phase of the research.⁸

TEDx Talks

TEDx Talks are fundamentally characterized as popularizing oral communication. These events are organized locally and independently, mirroring the communicative model created by the TED ONG: TED Talks.

Each Talk is presented by a speaker recognized as a specialist in a particular subject or area, which may not be directly related to an academic field. The speaker develops an oral presentation intended for a lay audience with diverse backgrounds and may utilize multimodal supports. The audience has the option to attend the Talk in two ways: live or online. Indeed, each presentation is recorded and subsequently published on the official TEDx YouTube channel, where various playlists are available, categorized by specific events, organizing entities, or event locations.

In linguistic and textual terms, the TED Talk model reproduced in TEDx Talks engenders hybrid forms of communication. This phenomenon emerges from the employment of diverse linguistic and communicative codes, the various channels through which messages are transmitted, and the processes of re-contextualization and re-conceptualization that are inherent to the primary aim of scientific popularization.

[T]heir discursive hybridity stems from the fact that the talks lie at the intersection between university lectures, scientific communication, newspaper articles, conference presentations and TV science programmes. (Caliendo, 2012, p. 113).

The relationship between the speaker and the target audience, whether in-person or online, is characterized by the speaker's primary objective of fostering interaction. This interaction is essential as it acknowledges the audience as active participants in the communicative event.

According to Caliendo (2012, p. 113), in general terms, it can be stated that TEDx Talks are "an interesting example of the increasing difference between written and spoken monologic and dialogic." This observation suggests the employment of different communicative strategies designed to foster engagement, including both direct and indirect references to the audience, as well as a range of textual resources aimed at making the Talk as understandable and effective as possible.

⁸The examples in the PB-FAM corpus will be identified by the following parameters: the corpus (PBFAM), the subcorpus (FF), the series (VV, LM and OQI) and the identification number within the subcorpus and series.

Regarding text composition strategies, it is essential to consider the multiple resources employed to build the speaker's credibility and to emphasize their status as an expert. This approach serves to ensure reliability and truthfulness in relation to the Talk's contents. Indeed, we can note references to the speaker's educational background and professional experiences, as well as affiliations with professional or specialized communities, are significant indicators of their expertise.

Simultaneously, the speaker frequently employs personal anecdotes, which help reduce the natural interactional distance between an expert and a non-expert. They can also include personal opinions and references to emotional states, that emphasize the human element/component of the communication, an essential aid in achieving the highest efficiency in the popularization process.

Academic Talk Videos

Academic Talk Videos are planned communicative events that predominantly exhibit a monological format. In these presentations, the speaker may employ multimodal resources, although such resources are limited to visual elements. Additionally, the discourse is characterized by a web-mediated specialized language, or Language for Specific Purposes (LSP), which oscillates between a [+/-formal] register (De Rosa, 2020a, 2020b). This diamesic variety can be considered an LSP due to its significant level of specialization, which includes a specific lexicon and methodologies for the creation of neologisms and the organization of texts (cf. Sobrero, 2006, p. 239).

Consequently, Academic Talk Videos cannot be strictly classified as 'popularizing' in the same manner as TEDx Talks, although they aim to reach a broad audience of scholars and students (didactic context) and to popularize scientific topics for a wider audience beyond just specialists. However, unlike TEDx Talks, there is no live audience⁹ during the recordings, and the talk can only be accessed online.

In the Academic Talk Videos, the speaker is consistently an academic, while in TEDx Talks is always an expert. However, the communicative purposes and the expectations of the audience differ significantly between the two contexts under analysis, despite their shared characteristics of specialized and academic discourse (Caliendo; Compagnone, 2014; Gotti, 1991, 2003, 2005; Gotti; Šarčević, 2006; Gotti; Giannoni, 2006; Kermas; Christiansen, 2013).

Academic Talk Videos exhibit a well-defined and coherent information structure, which is further enhanced by a highly specialized technical lexicon specific to the sector. A notable distinction from popularizing talks, which typically employ a [-formal] register, is that the diamesic variety utilized by the speaker predominantly represents a neo-standard variety of Brazilian Portuguese. This variety fluctuates between a [+formal] and [-formal] register, depending on the stylistic characteristics of the speaker and the degree of stylistic monitoring (whether it is a FALA-FALA talk [-monitored] or a FALA-ESCRITA talk [+monitored]).

⁹In Academic Talk Videos there is almost exclusively an audience made up of web users, while in TED(x) Talks there is an audience "at two levels: a group of co-present participants attending the TED conference and web-users at home" (Caliendo; Compagnone, 2014).

Ultimately, the textual functions discernible in Academic Talk Videos include cognitive, expository, and argumentative dimensions. Each presentation elucidates a scientific topic, aiming to clarify it for an audience that may be unfamiliar with the subject matter, while also requiring that the audience possesses the requisite cognitive skills to interpret the information effectively.

Results

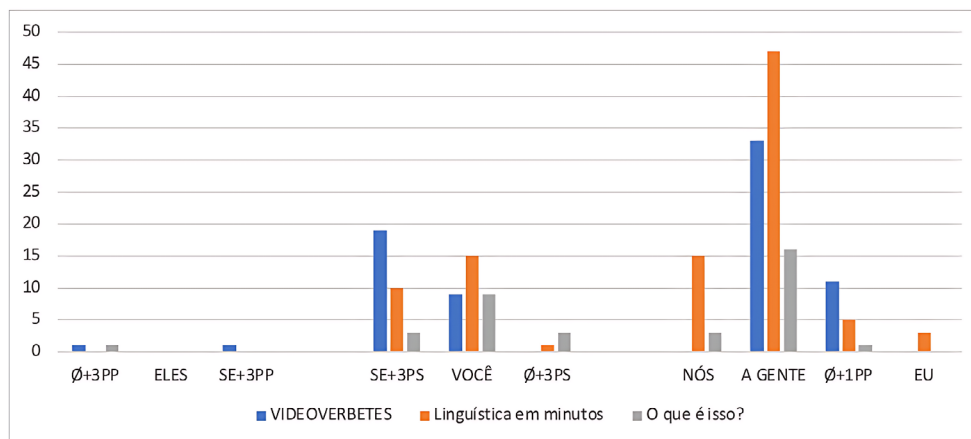
A total of 428 occurrences of subject indefiniteness strategies were found in finite sentences, with a distribution across both popularizing and semi-popularizing talks. Specifically, there were 222 occurrences in TEDx Talks and 206 occurrences in Academic Talk Videos.

Subject Indefiniteness in Academic Talk Videos

To present the results, we created a graph illustrating the distribution of the subject indefiniteness strategies analysed in our study (Graph 1): (ELES), SE+3PP, SE+3PS, (VOCÊ), Ø+3PS, (NÓS), (A GENTE), (EU),¹⁰ categorised by type of Academic talk video in finite sentences.

The graph demonstrates that both conservative and innovative grammars coexist in semi-popularizing specialized discourse, with standard strategies (SE+3P and ELES), accounting for 34 out of 206 occurrences, which represents approximately 17% of the total.

¹⁰ The nominative forms have been placed in brackets to account for both the overt forms and the null forms.



Graph 1. Subject Indefiniteness Strategies in Semi-Popularizing Specialized Discourse.

Source: own elaboration.

In this context, Duarte (2007, p. 105-106) asserts that in standard writing, the analysis of indefinite reference subjects “mostra que a escrita dos jornais brasileiros, nos gêneros focalizados, não se aproxima da fala; ao contrário, privilegia (...) o clítico se e as formas de primeira e 3ª pessoas do plural com o pronome preferencialmente nulo”. This observation is also reflected in our sample, where these three strategies – SE+3P, ELES, and 1PPmos – collectively account for 33.4% of occurrences (69 out of 206). The following examples serve to illustrate this point:

- (7) (Esse gesto analítico que **se**_{gen} faz ao compreender uma formação discursiva...
(PBFAM-FF-VV8)
- (8) Seja o Estado, pode-**se**_{arb} propor qualquer objetivo, mas o que o distingue de outras comunidades humanas é que ele faz uso legítimo da violência física, coisa que nenhuma outra comunidade humana poderia fazer.
(PBFAM-FF-OQI1)
- (9) Mas, do ponto de vista eh... da... de como me **Ø**_{arb} perguntam muitas vezes...
(PBFAM-FF-OQI7)

In a [+monitored] discursive context, such as semi-popularizing specialized discourse, traces of the co-occurrence between standard strategies ELES and SE+3PS can still be observed (Vargas, 2012, p. 57). However, it is important to note that these strategies are not concurrent or interchangeable, as they do not signify the same type of indefiniteness (Lunguinho; Medeiros Junior, 2009).

- (10) ...como **Ø**_{arb} não disseram o que é que **se**_{gen} ganhou, quem é que ganhou...
(PBFAM-FF-VV1)

In this context, it is essential to emphasize that the clitic SE can serve both generic and arbitrary reference, with the latter “em variação com o verbo na 3a pessoa do plural ou com o pronome nominativo ‘eles’ expresso e ainda com o verbo na 3a pessoa do singular sem qualquer marca” (Duarte *et al.*, 2021, p. 158-159). Nevertheless, within our sample, only 8 out of 33 constructions utilizing the clitic SE have arbitrary reference, accounting for 24% of the occurrences of this strategy. The following examples illustrate this point:

- (11) ...a sintaxe passa a ter uma relevância no procedimento de análise que não tinha na “Análise automática” de 1969, onde **se**_{arb} produzia um processo de desintagmatização para a construção do objeto discursivo.
(PBFAM-FF-VV8)
- (12) Para começar a situar a sociolinguística, a gente precisa voltar um pouco no início do que **se**_{arb} convencionou chamar de linguística moderna.
(PBFAM-FF-LM6)
- (13) ...há duas, basicamente duas, interpretações fonológicas para as vogais do português, uma mais antiga, em que **se**_{arb} defende que o português apresenta 12 fonemas vocálicos...
(PBFAM-FF-LM4)

Our data corroborate the findings of Lunguinho and Medeiros Jr. (2009) and Marins, Soares da Silva, and Duarte (2017), indicating that the two standard strategies are no longer in variation or concurrent, but are complementary. These strategies have become specialized to represent two opposing references: the generic reference, represented by the clitic SE, and the arbitrary reference, represented by ELES.

[T]he strategy that bears the features [+3rd person/+plural] has specialized to represent arbitrary reference whereas the clitic *se*, bearing the features [+3rd person/+singular], clearly specialises for the generic reference; the rare uses of *se_{arb}* confirm that they are residual (Marins; Soares da Silva; Duarte, 2017, p. 158).

Although it is not traditionally recognized by standard grammar as a strictly conventional strategy for indicating subject indefiniteness – similar to the clitic SE and ELES, 1PPmos – whose use in spontaneous speech has been declining, as previously noted alongside the two standard strategies (Duarte, 1993, 1995), it is observed in all three types of academic discourse. This strategy appears in both overt (18 occurrences) and null forms (17 occurrences). The data suggest that the utilization of this strategy remains productive in [+monitored] written typologies and text genres, including letters and journalistic texts (Cavalcante, 1999; Duarte; Lopes, 2002; Duarte, 2007).

- (14) So Então, como, na verdade, **Ø sustentamos** as... as instituições políticas existentes...
(PBFAM-FF-OQI2)
- (15) (...**ai Ø precisamos** compreender essa posição como a causa real da ilusão subjetiva que afeta o locutor na sua prática de enunciação.
(PBFAM-FF-VV8)
- (16) **nós** cada dia mais **precisamos** de aparatos mais complexos para alcançar ah... maiores energias.
(PBFAM-FF-OQI8)
- (17) Costuma-se acreditar que... **nós estamos** veiculando significados diferentes também.
(PBFAM-FF-LM7)

As previously noted, the pronominal forms of the third group (1P +/-plural) share generic reference (necessarily including the speaker) and the semi-inclusive feature (optionally including the addressee). However, if the 1PPmos, despite being a pronominal form, is incorporated into a [+monitored] discursive context, the pronominal form A GENTE, innovative like VOCÊ, is still rarely used in [+monitored] written discursive genres, despite its increasing use.

As estratégias com **você** e **a gente**, ausentes na escrita do PE, já começam a se implementar na escrita do PB, mas muito lentamente, e preferencialmente nas crônicas, um texto que é também opinativo, mas de natureza mais leve que os textos de opinião (Duarte, 2007, p. 107).

This gradual and incremental adoption of the form A GENTE in the [+monitored] written varieties of BP is substantiated by our analysis of a semi-popularizing specialized speech sample, in which it appears with notable frequency – 96 occurrences out of 206, representing nearly 47%. Nevertheless, despite this high rate, it is not used exclusively; indeed, for every three occurrences of A GENTE, there is one occurrence of 1PPmos (null or overt), as illustrated in Graph 1. This observation suggests that,

within a [+monitored] discursive context, both in standard writing and semi-popularizing specialized speech, 1PPmos remains a viable option for indefiniteness and, on the other hand, that, as the secondary option for 1PP, “*nós* constitui a escolha para uma indicação mais definida, enquanto a gente pode efetuar uma referência mais indeterminada, mesmo que essa expressão continue sendo usada, claramente, em referência à primeira pessoa” (Ilari; Franchi; Neves; Possenti, 1996, p. 93). However, these results are not reflected in spontaneous speech data, where, beyond the lack of co-occurrence and concurrence between A GENTE and NÓS, 1PPmos is declining among speakers under 30 (Duarte, 1993, 1995). (Duarte, 1993, 1995). Examples include:

- (18) ela tem elementos que são fixos e aqui um local, né?, que **a gente** chama de slot.
(PBFAM-FF-LM2)
- (19) **A gente** pode dizer que ela é, na verdade, o que sustenta a teoria.
(PBFAM-FF-VV10)
- (20) Então, **a gente** começa a perceber que o léxico tem uma relação de restrição linguística pelos contextos da própria estrutura linguística.
(PBFAM-FF-LM1)
- (21) ...com as características que **nós achamos** condizentes com os valores que a gente prega.
(PBFAM-FF-OQI1)
- (22) Quando **a gente** se espanta diante do cotidiano e Ø percebe que as explicações
(PBFAM-FF-OQI4)

Regarding the use of the 1PS, EU, which also belongs to the third group, the limited number of occurrences – only three – precludes the possibility of drawing any generalizations. Notably, two of these occurrences are found within the same talk and may be interpreted primarily as a stylistic choice. The following examples will illustrate this point:

- (23) “X-mente”, **eu** tenho “frequentemente”, “totalmente”, “claramente”, então no lugar do X aparece, pode aparecer qualquer adjetivo.
(PBFAM-FF-LM2)
- (24) Então **eu** posso, através da linguagem, discutir leitura escrita, por exemplo, como **eu** faço os processamentos de textos no caso, né?
(PBFAM-FF-LM1)

Despite the limited occurrence of only three instances, we agree with Marins, Soares da Silva, and Duarte (2017, p. 164) on the assertion that the use of EU, 1PS, enhances the sense of reality. This linguistic choice enables the speaker “to offer an example relevant to the situation addressed in the discourse, (...) indicating it could be experienced by anyone, including the speaker, thus reducing the scope of generic reference.”

In addition to the clitic SE+3PS, the second group includes the nominative form VOCÊ and the indeterminate null subject Ø+3PS. Similar to the clitic SE, this latter strategy can convey both arbitrary and generic reference. However, within our sample, all four instances demonstrate a generic reference. The following examples illustrate this point:

- (25) para ser linguista, Ø_{gen} **tem** que fazer uma graduação, o mestrado, o doutorado, então, né? precisamos de uma área um pouco, de uma formação consolidada.
(PBFAM-FF-LM3)
- (26) Acho que Ø_{gen} **não pode** pensar à política somente a partir do que alguns interpretam como a ciência política.
(PBFAM-FF-OQI8)

The final nominative form analysed is VOCÊ. This strategy for expressing subject indefiniteness – initially an allocutive form characterized by inclusive and generic reference – appears as the preferred strategy among educated urban speakers in Brazil, with a frequency of 44%, as well as among popular speakers, whose frequencies range from 45% to 49% (Duarte, 2007, p. 103). However, in Academic Talk Videos, primarily due to the monologic, speaker-centred nature of this textual genre, we found only 33 occurrences out of 206, resulting in a frequency of 16%. Below are some examples:

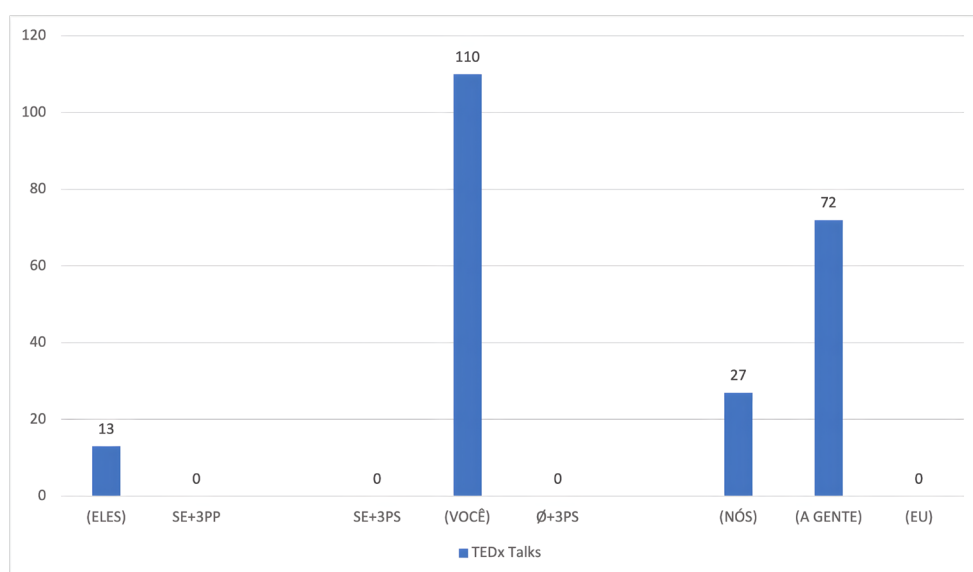
- (27) o que está sendo desenvolvido nessa tese é que **você** tem ao mesmo tempo na contemporaneidade o cinismo grego, quer dizer, aquele cinismo que parodia o poder, que desnuda o poder, e o cinismo funcionando do lado do poder.
(PBFAM-FF-VV5)
- (28) Então, a partir de agora, se alguém perguntar pra **você** se a sociolinguística é uma metodologia ou uma teoria, você já sabe dizer que ela é uma teoria e que tem, como qualquer outra teoria, uma metodologia própria pra fundamentar os seus princípios e as suas hipóteses.
(PBFAM-FF-LM6)
- (29) Então qualquer gesto que **você** faça, qualquer ação que **você** faça, pode significar ou não.
(PBFAM-FF-VV9)

However, if we add up the data for the two most innovative nominative forms, A GENTE (with 96 occurrences) and VOCÊ (with 33 occurrences), we arrive at 129 out of 206 occurrences representing a rate of 62%.

Subject indefiniteness in TEDx Talks

In summarizing the comprehensive findings from the analysis of the 9 TEDx Talks, the following graph can be presented (Graph 2).

The initial point to highlight is the preference for innovative strategies concerning subject indefiniteness. In support of the ongoing parametric change, we recorded 193 occurrences of overt indeterminate subjects, representing an 87% rate, in contrast to only 29 occurrences of null indeterminate subjects, which account for 13%. Additionally, it is noteworthy that there is a scarcity of occurrences of standard indeterminate strategies utilizing the clitic SE+3P, a strategy that, as previously indicated, is no longer in use in the spontaneous speech of individuals under the age of 30. This observation may lead to the conclusion that, within the context of specialized [-monitored] popularizing speech, the influence of spontaneous speech and neo-standard grammar is greater compared to the [+monitored] discursive contexts discussed in the preceding section.



Graph 2. Subject Indefiniteness Strategies in Popularizing Specialized Discourse.

Source: own elaboration.

Nevertheless, we recorded 13 occurrences of the standard 3PP strategy, of which 7 occurred with a null subject, Ø+3PP, and 6 with an overt subject. See the examples:

- (30) Ø+3PP_{arb} falaram aqui da minha formação, eu sou advogada, tenho um mestrado em direito e eu sou professora, professora de processo penal.
(TEDx 5)
- (31) aquela mesma escola onde Ø+3PP_{arb} me disseram que eu era um viadinho que eu era uma pessoa suja, que eu não devia tar ali... nessa mesma escola eu voltei.
(TEDx 3)
- (32) Quando Ø+3PP_{arb} me chamaram pra palestrar aqui né? No TEDx São Paulo [...]
(TEDx 8)

- (33) E-e aí meu primeiro trabalho fora de Belo Horizonte, **eles**_{arb} me convidaram pra um desfile em São Paulo.
(TEDx 4)
- (34) [...] alguém já percebeu isso? Agora **eles**_{arb} inseriram mais dois blocos de propaganda.
(TEDx 1)

It is important to highlight that the 3PP strategies, with either a null or overt subject ($\emptyset$ +3PP or ELES), constitute the only occurrences of indeterminate subjects with arbitrary interpretation within the sample of popularizing talks. This phenomenon accounts for a rate of 5.8%, in contrast to a 94.2% rate for subjects with generic reference.

As indicated in the preceding section, the 1PPmos strategy can be included in the standard BP strategies. This structure may exhibit a null subject, $\emptyset$ +1PPmos, or an overt subject, NÓS. We recorded a total of 27 occurrences: 12 occurrences with null subject (with a 5% rate) and 15 occurrences with overt subject, corroborating the trend previously illustrated.

- (35) Então falar o “menos emoção e mais razão” é uma ode à nossa liberdade, é pra que **nós tenhamos** mais autonomia, pra que nós não sejamos escravos das nossas emoções.
(TEDx 5)
- (36) Mas acontece que **nós não somos feitos** em pedaços, **nós somos** por inteiro e o que afeta uma parte afeta o todo.
(TEDx 6)
- (37) O que $\emptyset$ **vimos** no passado recente, nos permite refletir sobre um novo tempo [...]
(TEDx 7)
- (38) A minha missão aqui é realmente criar novos imaginários, gritar pro mundo que $\emptyset$ **precisamos** de novas histórias, precisamos de sua história e da sua história completa.
(TEDx 8)

Additionally, confirming what was previously stated in the preceding section regarding the 1P +/- plural strategies, the corpus of popularizing talks demonstrates a trend favouring the use of the inclusive A GENTE – 72 occurrences, accounting for 32% of the total. The following examples illustrate this trend:

- (39) Só que **a gente** não precisa dar um peso muito grande pra eles. **A gente** pode simplesmente escolher como é que **a gente** quer encarar esses problemas, eles não precisam ter um peso enorme.
(TEDx 4)
- (40) E eu acredito que **a gente** pode ser racional, se **a gente** não pudesse a gente não tinha descoberto o papel que a emoção tem, o papel que desempenham os vieses cognitivos.
(TEDx 5)

Nevertheless, VOCÊ is the preferred nominative pronominal form identified within the corpus of popularizing talks, as evidenced by the following excerpts:

- (41) O que faz você um herói, uma heroína, são as circunstâncias que **você** passa na sua vida. As circunstâncias que acontece(m) na sua vida (eh) que retiram de você coisas que **você** nem sabia que existia.
(TEDx 2)
- (42) Mas, é difícil você esquecer aquela vivência escrota, sabe? aquele momento que de alguma forma **você** já foi preconceituoso, reproduziu alguma fobia, algo do tipo que te incomoda hoje. Isso não vai sair da sua cabeça, isso tá aqui, ainda que você queira apagar [...]
(TEDx 3)

In terms of the use of this pronominal form, as previously demonstrated in the data from the semi-popularizing sample (TEDx Talks corpus), we recorded 100 occurrences of the nominative form VOCÊ as an overt subject, resulting in a frequency rate of 45%.

In conclusion, it is important to highlight that one of the primary objectives of the speaker(s) in TEDx Talks is to present content effectively, capture the audience's attention, and foster an expert/non expert interaction. Consequently, along an inclusive/semi-inclusive/exclusive continuum, we found inclusive uses of both the nominative form VOCÊ and the nominative forms of 1PP, which seem to function as meta-references (Scotto di Carlo, 2014a, 2014b, 2014c).

Final considerations

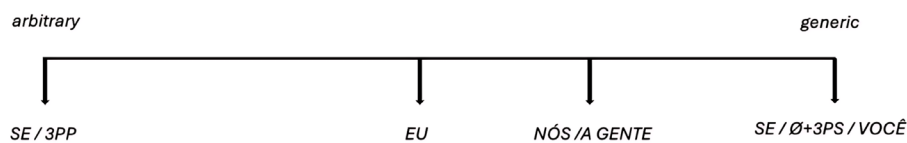
The analysis of the data presented herein substantiates our initial hypothesis: web-mediated popularizing and semi-popularizing genres, due to their [+/-formal] register and a varying degree of stylistic monitoring, exhibit differing uses of subject indefiniteness strategies. Specifically, these strategies are more conservative in Academic Talk Videos (as per our initial hypothesis) and less conservative in TEDx Talks. Indeed, within the context of web-mediated specialized popularizing speech, strategies employed align more closely with educated urban speech, using only 1PPmos (overt or null) as the most standardised nominative pronoun, which constitutes only 12% of the sample. In contrast, semi-popularizing web-mediated specialized speech, both standard strategies – SE+3P (33 occurrences) and ELES (1 occurrence) – are employed, forms that are rarely seen in educated urban speech and are nearly extinct in popular speech (Marins; Soares da Silva; Duarte, 2017).

Our data corroborate the findings of Duarte (2007, p. 108) regarding the existence of two distinct grammatical systems among educated urban speakers, particularly in standard written language: “é grande a distância que separa a gramática da fala e a da norma culta escrita prescrita no Brasil e é notória a falta de estudos comparativos entre o conhecimento que a criança leva para a escola e o conhecimento do indivíduo letrado”.

In the context of our analysis, it is pertinent to note that web-mediated specialized speech situates itself on the FALA-ESCRITA diamesic variation continuum between educated urban speech and standard writing, between a neo-standard grammar and a standard grammar. It is crucial to emphasize that diaphasic variables related to [+/-formal] register and [+/-monitored] stylistic monitoring reveal a distinct differentiation in the strategies of indefiniteness employed in web-mediated popularizing discourse compared to semi-popularizing discourse. Specifically, in TEDx Talks, there is a polarised use of indefiniteness along an arbitrary/generic continuum, which maintains 3PP (either overt or null) and incorporates both innovative pronominal forms (also either overt or null) while reintroducing an outdated indeterminate form like 1PPmos as generic referential strategy.

In Academic Talk Videos, the specialized semi-popularizing speech revives outdated indefinite forms characteristic of educated urban speech, such as the SE+3P clitic (with or without agreement) and NÓS. It incorporates innovative forms such as VOCÊ, A GENTE, and Ø+3PS, while maintaining ELES as a strategy for arbitrary indefiniteness. The figure 3 illustrates the varying uses along two continua, which reflect the the popularizing or semi-popularizing typology of the talks.

Subject indefiniteness Forms in web-mediated specialised semi-popularising speech



Subject indefiniteness Forms in web-mediated specialised popularising speech

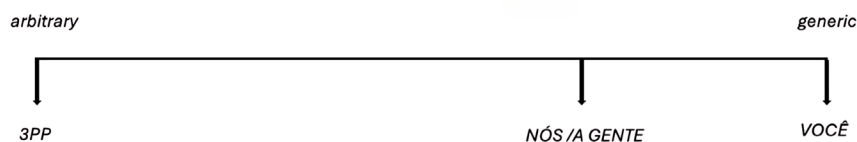


Figure 3. Subject Indefiniteness Forms in Web-mediated Specialized Discourse.

Source: own elaboration.

It becomes evident that:

- a. in popularizing talks, the restricted range of subject indefiniteness forms presents a scenario that closely resembles educated urban speech. This phenomenon is characterized by innovative pronominal forms, which can be distilled into four strategies as outlined by Marins, Soares da Silva, and Duarte (2017, p. 165);
- b. in semi-popularizing talks, we observe a complex and articulated array of linguistic forms, both conservative forms that are outdated in educated urban speech, and innovative forms. This array of strategies lends support to the assertion that “o brasileiro letrado se comporta como um bilíngue fazendo uso da gramática falada atual/inovadora e da gramática da escrita de épocas anteriores, em um processo de code-switching” (Kato, 2019, p. 56).

It can be argued, reaffirming the insights of Kato (2005, 2019), Duarte (2007, 2023), and Duarte and Serra (2015), among others, that with regard to strategies for subject indefiniteness in web-mediated specialized discourse, the grammar of the educated urban Brazilian speaker does not correspond to a standard grammar or to the grammar of a literate speaker from the past. Indeed, what emerges among contemporary educated urban Brazilian speakers is a competition of grammars, due to the possibility of indirect access to “uma periferia marcada onde valores paramétricos opostos ao da gramática nuclear podem estar presentes, (...), valores esses que podem assumir um valor competitivo, durante a escolarização” (Kato, 2005, p. 131).

In conclusion, the educated urban Brazilian speaker, within written or transmitted contexts characterized as [+monitored] and [+formal], such as in semi-popularizing specialized discourse, exhibits a phenomenon of code-switching and competition between two types of grammar: a more conservative, standardized grammar acquired through formal education, which is never used in spoken language, and an innovative grammar that serves as a reference in educated urban speech, often referred to as neo-standard Brazilian grammar.

REFERENCES

BARBOSA, Pilar; KATO, Mary Aizawa; DUARTE, M. Eugenia. Null Subjects in European and Brazilian Portuguese. *Journal of Portuguese Linguistics*, v. 4, n. 1, p. 11-52, 2005.

BECHARA, Evanildo. *Moderna gramática portuguesa*. São Paulo: Editora Nacional, 1977.

BECHARA, Evanildo. *Moderna gramática portuguesa*. 37. ed. Rio de Janeiro: Lucerna, 2004.

BORTONI-RICARDO, Stella Maris. *Nós cheguei na escola, e agora?* Sociolinguística & Educação. São Paulo: Parábola, 2005.

CALIENDO, Giuditta. The popularization of science in web-based genres. In: CALIENDO, Giuditta; BONGO, Giancarmine (org.). *The language of popularization: Theoretical and descriptive models*. Bern: Peter Lang, 2012. p. 101-132.

CALIENDO, Giuditta; COMPAGNONE, Antonio. Expressing epistemic stance in university lectures and TED Talks: a contrastive corpus-based analysis. *Lingue e Linguaggi*, v. 11, p. 105-122, 2014.

CAVAGNOLI, Stefania. *La comunicazione specialistica*. Roma: Carocci, 2007.

CAVALCANTE, Silvia. *A indeterminação do sujeito na escrita padrão: a imprensa carioca dos séculos XIX e XX*. 1999. Dissertação (Mestrado em Letras) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1999.

CAVALCANTE, Silvia. *O uso de SE com infinitivo na história do português: do português clássico ao português europeu e brasileiro modernos*. 2006. Tese (Doutorado em Linguística) – Universidade Estadual de Campinas, Campinas, 2006.

CAVALCANTE, Silvia. O sujeito nulo de referência indeterminada na fala culta carioca. *Diadorim, Revista de Estudos Linguísticos e Literários*, v. 2, p. 63-82, 2007.

CAVALCANTE, Silvia; DUARTE, Maria Eugênia Lammoglia. Sujeitos de referência arbitrária em sentenças infinitivas do português e o parâmetro do sujeito nulo. In: TSIPLAKOU, Stavroula; KARYOLEMOU, Marilena; PAVLOU, Pavlos. (org.). *Language variation – European perspectives II: selected papers from the 4th international conference on language variation in Europe*. Amsterdam: John Benjamins, 2009. p. 47-58.

CHOMSKY, Noam. *Lectures on government and binding*. Dordrecht: Foris, 1981.

CHOMSKY, Noam. *The minimalist program*. Cambridge: MIT Press, 1995.

CORTELAZZO, Michele. *Lingue speciali*. Padova: Unipress, 1994.

COUTO, Marco Aurelio de Souza. *Estratégias pronominais de indeterminação do agente*. 2005. Dissertação (Mestrado em Letras Vernáculas) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005.

CUNHA, Celso; CINTRA, Lindley. *Nova gramática do português contemporâneo*. 5. ed. Rio de Janeiro: Lexicon, 2008.

CYRINO, Sonia; DUARTE, Maria Eugênia Lammoglia; KATO, Mary A. Visible subjects and invisible clitics in Brazilian Portuguese. In: KATO, Mary Aizawa; NEGRÃO, Esmeralda Vailati (org.). *Brazilian Portuguese and the Null Subject Parameter*. Madri: Iberoamericana; Frankfurt: Vervuert, 2000. p. 55-73.

DARDANO, Maurizio. Capire la lingua della scienza. In: DARDANO, Maurizio; FRENGUELLI, Gianluca. (org.). *L'italiano di oggi*. Roma: Aracne, 2008. p. 173-188.

DE ROSA, Gian Luigi. Reflexos do processo de restandardização do PB no falado fílmico brasileiro contemporâneo. In: SILVA, Augusto Soares; TORRES, Amadeu; GONÇALVES, Miguel (org.). *Línguas pluricêntricas: variação linguística e dimensões sociocognitivas / pluricentric languages: Linguistic variation and sociocognitive dimensions*. Braga: Aletheia, 2011. p. 377-392.

DE ROSA, Gian Luigi. *Mondi Doppiati*: tradurre l'audiovisivo dal portoghese tra variazione linguistica e problematiche traduttive. Milano: Franco Angeli, 2012.

DE ROSA, Gian Luigi. Características da fala acadêmica monitorada no Brasil: os videoverbetes da ENCIDIS entre PB técnico-científico e PB neo-standard. *Cultura Latinoamericana: Revista de Estudos Interculturais*, v. 32, n. 2, p. 238-261, 2020a.

DE ROSA, Gian Luigi. O discurso científico mediado pela web: legendar videoverbetes entre tipologias textuais, línguas especiais e problemáticas tradutórias. In: DE LAURENTIIS, Antonella; DE ROSA, Gian Luigi (org.). *Discorso specialistico e multimedialità: caratteristiche linguistiche e problematiche traduttive*. Roma: Edizioni Nuova Cultura, 2020b. p. 29-45.

DUARTE, Maria Eugênia Lammoglia. Do pronome nulo ao pronome pleno: a trajetória do sujeito nulo no português do Brasil. In: ROBERTS, Ian; KATO, Mary Aizawa (org.). *Português brasileiro: uma viagem diacrônica*. Campinas: Editora da Unicamp, 1993. p. 107-128.

DUARTE, Maria Eugênia Lammoglia. *A perda do princípio Evite Pronome no português brasileiro*. 1995. Tese (Doutorado em Linguística) – Universidade Estadual de Campinas, Campinas, 1995.

DUARTE, Maria Eugênia Lammoglia. A evolução na representação do sujeito pronominal em dois tempos. In: PAIVA, Maria da Conceição; DUARTE, Maria Eugênia Lammoglia (org.). *Mudança linguística em tempo real*. Rio de Janeiro: ContraCapa; FAPERJ, 2003. p. 115-128.

DUARTE, Maria Eugênia Lammoglia. Sujeitos de referência definida e arbitrária: aspectos conservadores e inovadores na escrita padrão. *Revista Linguística*, v. 3, n. 1, p. 83-115, 2007a.

DUARTE, Maria Eugênia Lammoglia. A sintaxe do português do Brasil: entre a fala e escrita padrão. In: SALOMÃO NETTO, Sônia. (org.). *Temas da Língua Portuguesa: do pluricentrismo à didática*. Roma: Edizioni Nuona Cutura, 2020. p. 131-152.

DUARTE, Maria Eugênia Lammoglia. Avanços no estudo da mudança sintática associando a Teoria da Variação e Mudança e a Teoria de Princípios e Parâmetros. *Cadernos de Estudos Linguísticos*, v. 57, n. 1, p. 85-111, 2015.

DUARTE, Maria Eugênia Lammoglia. Sociolinguística ‘Paramétrica’. In: MOLLICA, Maria Cecillia; FERRAREZI JR., Celso. (ed.). *Sociolinguística, sociolinguísticas – uma introdução*. São Paulo: Contexto, 2016. p. 33-44.

DUARTE, Maria Eugênia Lammoglia. (ed.). *O sujeito em peças de teatro (1833-1992): estudos diacrônicos*. São Paulo: Parábola, 2012.

DUARTE, Maria Eugênia Lammoglia. Sobre algumas consequências do enfraquecimento da categoria concordância no português brasileiro. *Lingue e Linguaggi*, v. 57, p. 31-56, 2023.

DUARTE, Maria Eugênia Lammoglia; MARQUES DE SOUSA, Antonio Anderson; FERNANDES, Ulli Santos Bispo; CARDOSO, Marianna Maroja Confalonieri. A redução no quadro de clíticos de 3ª pessoa no português brasileiro: um estudo diacrônico. *Revista Laborhistórico*, v. 4, n. esp., p. 154-187, 2021.

DUARTE, Maria Eugênia Lamoglia; LOPES, Célia Regina dos Santos. Realizaram, realizou-se ou realizamos...? As formas de indeterminação do sujeito em cartas de jornais no século XIX. In: DUARTE, Maria Eugênia Lamoglia; CALLOU, Dinah. *Para a História do Português Brasileiro*, Notícias de corpora e outros estudos. Rio de Janeiro: In-Fólio, 2002. v. IV, p. 155-165.

DUARTE, Maria Eugênia L.; SERRA, Carolina Ribeiro. Gramática(s), ensino de português e adequação linguística. *Matraga*, v. 22, p. 31-55, 2015.

DUARTE, Maria Eugênia Lammoglia.; SOARES da SILVA, Humberto. Microparametric variation in Spanish and Portuguese: The Null Subject Parameter and the role of the verb inflectional paradigm. In: KATO, Mary Aizawa; ORDÓÑEZ, F. (ed.). *The Morphosyntax of Portuguese and Spanish in Latin America*. 2016. p. 1-26.

EGERLAND, Verner. Impersonal pronouns in Scandinavian and Romance. *Working Papers in Scandinavian Syntax*, v. 71, p. 75-102, 2003.

GALVES, Charlotte. A sintaxe do português brasileiro. *Ensaio de Linguística*, v. 13, p. 31-49, 1987.

GALVES, Charlotte. A gramática do Português Brasileiro. In: CASTILHO, Ataliba Teixeira; BASÍLIO, Margarida (org.). *Gramática do português falado*. Campinas: Editora da Unicamp, 1998. p. 79-96.

GOTTI, Maurizio. *I Linguaggi Specialistici. Caratteristiche linguistiche e criteri pragmatici*. Firenze: La Nuova Italia, 1991.

GOTTI, Maurizio. Specialized Discourse. *Linguistic Features and Changing Conventions*. Bern: Peter Lang, 2003.

GOTTI, Maurizio. *Investigating Specialized Discourse*. Bern: Peter Lang, 2005.

GOTTI, Maurizio; GIANNONI, Davide S. (ed.). *New Trends in Specialized Discourse Analysis*. Bern: Peter Lang, 2006.

GOTTI, Maurizio; ŠARČEVIĆ, Susan (ed.). *Insights into Specialized Translation*. Bern: Peter Lang, 2006.

GUALDO, Riccardo. La variazione nella comunicazione specialistica. In: VISCONTI, Jacqueline (ed.). *Parole nostre*. Le diverse voci della comunicazione specialistica e settoriale. Bologna: Il Mulino, 2019. p. 57-75

GUALDO, Riccardo; TELVE, Stefano. *Linguaggi specialistici dell'italiano*. Roma: Carocci, 2011.

HOLMBERG, Anders. Is there a little pro? Evidence from Finnish. *Linguistic Inquiry*, v. 36, p. 533-564, 2005.

HOLMBERG, Anders. The null generic pronoun in Finnish: a case of incorporation. In: BIBERAUER, Theresa et al. (ed.). *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge: Cambridge University Press, 2010. p. 200-230.

HOLMBERG, Anders; PHIMSAWAT, On-Usa. Truly Minimal Pronouns. *Diadorim, Revista de Estudos Linguísticos e Literários*, Rio de Janeiro, v. 19, n. esp., p. 11-36, 2017.

ILARI, Rodolfo; FRANCHI, Carlos; NEVES, Maria Helena Moura; POSSENTI, Sírio. Os pronomes pessoais do português falado: roteiro para a análise. In: CASTILHO, Ataliba Teixeira; BASÍLIO, Margarida (org.). *Gramática do português falado*. Campinas: Editora da Unicamp, 2002. p. 73-159

KATO, Mary Aizawa. Strong and weak pronouns in the null subject parameter. *Probus*, v. 11, n. 1, p. 1-38, 1999.

KATO, Mary Aizawa. A gramática do letrado: questões para a teoria gramatical. In: MARQUES, Maria Aldina; TEIXEIRA, José; LEMOS, Aida (org.). *Ciências da Linguagem: trinta anos de investigação e ensino*. Braga: CEHUM (U. do Minho), 2005. p. 131-145.

KATO, Mary Aizawa. Português brasileiro: Última flor do lácio, inculta e bela. *Revista da ABRALIN*, v. 17, n. 1, p. 52-80, 2019.

KATO, Mary Aizawa. As controvérsias sobre o sujeito nulo no português brasileiro. In: CASTAGNA, Vanessa; QUAREZEMIN, Sandra (ed.). *Da Linguística ao ensino: Travessias em Língua Portuguesa*. Venezia: Edizioni Ca' Foscari, 2020. p. 75-92.

KATO, Mary Aizawa; TARALLO, Fernando. Anything YOU can do in Brazilian Portuguese. In: JAEGGLI, Osvaldo; SILVA-CORVALAN, Carmen. (ed.). *Studies in Romance Linguistics*. Dordrecht: Foris, 1986. p. 343-358.

KERMAS, Susan; CHRISTIANSEN, Thomas (ed.). *The Popularization of the Specialized Discourse and Knowledge across Communities and Cultures*. Bari: Edipuglia, 2013.

LABOV, William. *Principles of Linguistic Change: Internal Factors*. Cambridge: Blackwell, 1994.

LUNGUINHO, Marcos Vinicius da Silva; MEDEIROS JUNIOR, Paulo. Inventou um novo tipo de sujeito: características sintáticas e semânticas de uma estratégia de indeterminação do sujeito no português brasileiro. *Revista de Estudos em Língua e Literatura*, v. 23, p. 147-191, 2009.

MARCUSCHI, Luiz Antônio. Oralidade e Letramento. *Da fala para a escrita: atividades de retextualização*. Cortez: São Paulo, 2008.

MARIANI, Bethania. Linguagem, conhecimento e tecnologia: a Enciclopédia Audiovisual da Análise do Discurso e áreas afins. *Linguagem & Ensino*, v. 21, n. esp., p. 359-393, 2018.

MARIANI, Bethania. La produzione e la circolazione del sapere su piattaforme digitali: lo status del portoghese brasiliano in un'enciclopedia digitale sottotitolata. In: DE LAURENTIIS, Antonella; DE ROSA, Gian Luigi (ed.). *Lingue e Linguaggi*, v. 35, p. 13-28, 2020.

MARINS, Juliana; SOARES DA SILVA, Humberto; DUARTE, Maria Eugênia Lammoglia. Revisiting Duarte (1995): For a Gradient Analysis of Indeterminate Subjects in Brazilian Portuguese. *Diadorim, Revista de Estudos Linguísticos e Literários*, Rio de Janeiro, n. 18, n. esp., p. 140-172, 2017.

MARTINS, Ana Maria; NUNES, Jairo. Passives and Se Constructions. In: WETZELS, Leo; MENUZZI, Sergio; COSTA, João (ed.). *The Handbook of Portuguese Linguistics*. Hoboken, NJ: Wiley-Blackwell, 2016. p. 318-337.

MATEUS, Maria Helena Mira et al. *Gramática da Língua Portuguesa*. Caminho: Lisboa, 2003.

MATTOSO CÂMARA JR., Joaquim. *Princípios de Linguística Geral*. Rio de Janeiro: Livraria Acadêmica, 1976.

NASCENTES, Antenor. *O idioma nacional*. Rio de Janeiro: Livraria Acadêmica, 1960 [1938].

NUNES, Jairo Morais. *O famigerado SE: uma análise sincrônica e diacrônica das construções com se apassivador e indeterminador*. 1990. Dissertação (Mestrado em Linguística) – Universidade Estadual de Campinas, Campinas, 1990.

NUNES, Jairo Morais. Se apassivador e se indeterminado: o percurso diacrônico no português brasileiro. *Cadernos de Estudos Linguísticos*, v. 20, p. 33-57, 1991.

RAPOSO Eduardo; URIAGEREKA Juan. Indefinite SE. *Natural Language and Linguistic Theory*, v. 14, p. 749-810, 1996

ROCHA LIMA, Carlos Henrique. *Gramática normativa da língua portuguesa*. Rio de Janeiro: José Olympio, 1972.

SAID ALI, Manuel. *Gramática Histórica da Língua Portuguesa*. São Paulo: Melhoramentos, 1965.

SABATINI, Francesco. Rigidità-esplicitzza vs elasticità-implicitzza: possibili parametri massimi per una tipologia dei testi. In: SKYTTE, Gunver; SABATINI, Fransesco (ed.). *Linguistica testuale comparativa*. Museum Tusculanum Press, 1999. p. 141-172.

SABATINI, Francesco. Lezione di italiano. *Grammatica, storia, buon uso*. Milano: Mondadori, 2016.

SCHERRE, Maria Marta Pereira. *Doa-se lindos filhotes de poodle: variação linguística, mídia e preconceito*. São Paulo: Parábola, 2005.

SCOTTO DI CARLO, Giuseppina. The role of proximity in online popularizations: The case of TED talks. *Discourse Studies*, v. 16, n. 5, p. 1-16, 2014a.

SCOTTO DI CARLO, Giuseppina. Figurative Language in Science Popularisation: Similes as an Explanatory Strategy in TED Talks. 3L: *The Southeast Asian Journal of English Language Studies*, v. 20, n. 3, p. 1-16, 2014b.

SCOTTO DI CARLO, Giuseppina. New trends in knowledge dissemination: TED Talks. *Acta Scientiarum: Language and Culture*, v. 36, n. 2, p. 121-130, 2014c.

SOBRERO, Alberto. *Lingue Speciali*. In: SOBRERO, Alberto (ed.). *Introduzione all'italiano contemporaneo. La variazione e gli usi*. Roma; Bari: Laterza, [1993] 2006. Vol. 2. p. 237-277.

TARALLO, Fernando. Diagnosticando uma gramática brasileira: o português d'aquém e d'além-mar ao final do século XIX. In: ROBERTS, Ian; KATO, Mary Aizawa (ed.). *Português Brasileiro. Uma viagem diacrônica*. Campinas: Editora da Unicamp, 1993. p. 69-105.

VARGAS, Amanda de Santana Campos. *Estratégias pronominais de indeterminação: um estudo diacrônico*. 2010. Dissertação (Mestrado em Letras Vernáculas) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2010.

VARGAS, Amanda de Santana Campos. A evolução na representação das estratégias pronominais de indeterminação. In: DUARTE, Maria Eugênia Lammoglia (ed.). *O sujeito nulo em peças de teatro (1843-1992): estudos diacrônicos*. São Paulo: Parábola Editorial, 2012. p. 45-68.

WEINREICH, Uriel; LABOV, William; HERZOG, Marvin. Empirical foundations for a theory of language change. In: LEHMAN, Winfred; MALKIEL, Yakov (ed.). *Directions for Historical Linguistics*. Austin: University of Texas, 2006 [1968]. p. 97-195.

Estratégias de indeterminação do sujeito no discurso especializado Web-Mediated

RESUMO:

O presente contributo, baseado numa amostra de Web-Mediated Talk Videos (Academic Talk Videos e TEDxTalks), propõe-se observar a representação de sujeitos com referência indeterminada no português brasileiro (PB). Estudos recentes têm demonstrado uma preferência por formas pronominais nominativas para expressar referência genérica e arbitrária, particularmente em sentenças finitas tanto da fala culta urbana quanto da fala popular. Este estudo, fundamentado na teoria da mudança linguística proposta por Weinreich, Labov e Herzog (2006), combinada com a teoria de Princípios e Parâmetros (Chomsky, 1981, 1995), tem como objetivo investigar as estratégias de indeterminação do sujeito no discurso especializado Web-mediated, usando dois corpora: um corpus de Academic Talk Videos corpus (Talks semi-divulgativos) e um corpus de TEDx Talks (Talks divulgativos). Nossa principal hipótese é que, dependendo do corpus, devemos observar uma gramática mais conservadora, com preferência por estratégias mais padronizadas, no corpus Academic Talk Videos [-divulgativo], [+monitorado] e [+formal], e uma gramática menos conservadora, com preferência por estratégias inovadoras, no corpus TEDx Talks [+divulgativo], [-monitorado] e [-formal].

PALAVRAS-CHAVE: português brasileiro, Sujeito indeterminado, fala culta urbana, Academic Talk Videos, TEDx Talks.