The Expression 'Daí' in Curitiba Informants Speech: a study based on the Brazilian Linguistic Atlas data

Dayse de Souza Lourenço **SIMÕES***
Josyelle Bonfante **CURTI****
Mariana Spagnolo **MARTINS*****

* PhD (2019) in Language Studies at State University of Londrina (UEL). Contact: dayse.lourenco1990@gmail.com.

** Master e PhD student in Language Studies at State University of Londrina (UEL). Contact: joosy.curtii@gmail.com.

*** Master e PhD student in Language Studies at State University of Londrina (UEL). Contact: mariana.spagnolo@hotmail.com

Abstract:

This paper aims to identify the circumstances of use of the item *daí* (then) in the speech of informants from Curitiba and has as corpus data from interviews of the Projeto Atlas Linguístico do Brasil (ALiB) (Brazilian Linguistic Atlas Project) conducted in this city. This research is justified by the fact that in southern Brazilian localities the use of *daí* is frequent in contexts different from those established by grammars, that is to say, as an adverb of place or time. For the constitution of the theoretical foundation, we used studies on grammaticalization process and stratified the results based on the Linguistic Variation Theory, in which we observed the data according to the diasexual, diageneric and diastratic variables. Therefore, this research intents to: (i) verify, in the speech of Curitiba informants from ALiB corpus, the use of the expression *daí* as an adverb and as other grammatical or discursive classes; (ii) analyze each employment according to extralinguistic variables; and (iii) verify if there are linguistic factors intervening in the use of the expression *daí*. Thus, we found that *daí* proved to be a discursive mark in the speech of older women with higher education, which may indicate that the change did not consolidate, but the frequency of its usage remains at the level of diageneric and diastratic variation.

Keywords:
Useage of *daí*. Brazilian Linguistic Atlas. Curitiba.

The Expression 'Daí' in Curitiba Informants Speech: a study based on the Brazilian Linguistic Atlas data

Dayse de Souza Lourenço Simões; Josyelle Bonfante Curti; Mariana Spagnolo Martins

INTRODUCTION

The term *daí* (then) is the contraction of the preposition DE (from) with the adverb AÍ (there), and, according to Cunha (1994), dates from the 14th century, in the forms *dy*, *di* and *dj*. The interest in the theme is due to the use of *daí* in contexts other than those in which: (i) it indicates a place, as in "Get out of there" (from that place); or (ii) the passage of time, as in "Then (= after that) he left home". Although for morphological studies *daí* is the junction of the preposition *de+aí* = adverb of time or place, in the speech of some Curitibanos this expression performs other categories, having different functions from those established by grammars. This means grammaticalization, which, according to Tavares (2006, p. 266-267), represents

> an evolution from a state of greater autonomy and stability of meanings to a state of less autonomy and greater instability: as lexical items undergo grammaticalization, meanings of an abstract and generic nature emerge, more adaptable to diversified contexts of use.

According to the author, the grammaticalization process of *daí* is understood, in line with the Text Linguistics, as a sequential marker of the text that can be configured as a sequencer of retroaction (anaphoric use, resuming something already said) or of propulsion (leading the interlocutor's attention to what will still be said in the text), both important factors for textual cohesion.

Likewise, Risso (2006 apud SANTOS, 2014, p. 196) indicates that *daí* is configured as a sequential marker that presents the topic sequencing trace, performing in the textual microstructure and making the theme progress in a serial and articulated way or, as Koch says (2015), establishing, between the segments of the text, several types of semantic or pragmatic-discursive relationship. In other words, *daí*, as an essentially conversational marker, allows a topical continuity, supported by the insertion or in the resumption of ideas and arguments, acting on the textual progression and articulating the text and its ideas.

In this scenario of possibilities, this research aims to identify the use of the expression *daí* in the speech of Curitiba informants, seeking to ascertain in which circumstance such expression is used and which linguistic and extralinguistic factors can

act on the greater or lesser use of this item in contexts other than established by Normative Grammar.

For this purpose, we rely on data collected by the Brazilian Linguistic Atlas Project (ALiB), which involves 15 Brazilian university institutions and has its headquarters at the Federal University of Bahia. The project sights to describe the diatopic, diastratic, diageneric and diaphasic differences; to provide content to language scholars; to plot dialectal areas; to build a collection of data capable of assisting in dictionaries development and to "contribute to the understanding of the Portuguese language in Brazil as a diversified social instrument of communication, holding several norms of use, but endowed with a systemic unit" (COMITÊ NACIONAL…, 2001).

ALiB data includes 250 research points, fixed according to demographic density, including the state capitals – except Palmas and Brasília, due to the their recent foundation –, totaling 1,100 informants, as asserted by Cardoso (2006), stratified according to sex, age group (18 to 30 and 50 to 65 years) and school level. Thus, we are also based on the Linguistic Variation Theory, intending to collaborate with the description and the analysis of the Curitiba's speech characteristics, taking as an object of analysis the speech collected in eight interviews conducted in the capital of Paraná.

A Bit About Grammaticalization

From the observation of the use of *daí* in the corpus collected by ALiB in Curitiba, with functions different from those foreseen or established by the Normative Grammar, we make way for the study of grammaticalization,[1] which briefly is when a lexical item begins to assume, in some cases, a new grammatical position, or when a grammatical item becomes even more grammatical.

We refered to the studies of Vitral (1996), Castilho (1997a, 1997b), Tavares (2002, 2006) and Gonçalves, Lima-Hernandes and Casseb-Galvão (2007), and we invoke Vitral, Castilho and Gonçalves, Lima-Hernandes and Casseb-Galvão for the theoretical basis on grammaticalization. In Tavares (2002, 2006), which served as a basis for the description and the data analysis; in addition to concepts about this process, we observed, specifically, the item *daí* under the sociofunctionalist perspective of the language.

According to Castilho, the modular conception of the spoken language consists of semanticization, discoursivization and grammaticalization processes:

---

[1] [From the french *Grammaticalisation*] noun. In the process of linguistic change, it is the passage of a morpheme that works as a free form or as a derivational affix, the grammatical marker (FERREIRA, 2008).

Briefly, the processes intuited by so many oral linguists consisted, in my view, of activating properties already contained in the lexicon, which would result in the discursive (= discursivization process), grammatical (= grammaticalization process) and semantic (= semanticization process) categories. In this perspective, the Lexicon is taken as a primitive linguistic component, as it does not derive from other instances of the language (CASTILHO, 1997b, p. 109).

Such processes create sets of language modules, a kind of linguistic organization. So, again, persuant to the author:

each module will be characterized by having a set of properties. Thus, referencing, predication, deixis, forcefulness, etc., are properties of the semantic module. The classes, the relationships that are established between them, the functions that these classes assume in the utterance, as well as their phonological representation, come from grammatical properties. Topical organization, the arrangement of topics in discursive units and the operators of textual cohesion can be postulated as properties of speech (CASTILHO, 1997b, p. 108).

As the linguist exemplifies, the functions that the word classes perform are defined as belonging to the grammar, thus the change that a class may undergo corresponds to grammaticalization, that is, when a lexical item, which is already a grammatical item, becomes even more grammatical, assuming other categories besides those already established by the normative aspect of the language.

In addition, Castilho (1997a) states that grammaticalization is also the extension of the limits of a morpheme, whose grammatical charter advances from the lexicon to the grammar, which is the transformation from a derivative format to a flexional format. We understand that in the grammaticalization process a lexical item loses its meaning and assumes grammatical functions according to the use of the oral language.

Thus, grammaticalization is a process of reworking the language, being one of the most common ways of observing languages in general, after all, according to Gonçalves, Lima-Hernandes and Casseb-Galvão (2007), there is no grammar as a finished product, but a constant grammaticalization. Furthermore, just as the authors consider grammaticalization to be a constantly changing process, Tavares (2002) understands that grammaticalization, like grammar, is emerging, is always on the way to constitution, but never actually constitutes itself, since that grammatical rules do not correspond to the demand of the spoken language, whereas grammaticalization, in a simplified way, has greater acceptance over the oral modality of the language.

The emphasized process refers to Givón's functional cycle, when he points to a pragmatic-discursive motivation among the stages of grammaticalization. Givón (apud LOPES, 2008) presents a regularizing scheme processes for the use of language, dealing with Saussure's concepts in a diachronic perspective: Discourse> Syntax>Morphology>Zero. In adapting to this cycle, Vitral (1996), when studying the

grammaticalization of the lexical item *you*, develops something similar to what Givón presented. Vitral (1996) points that the grammaticalization process can be exemplified as: item with lexical meaning>grammatical item>clitic>affix>flexional. According to the author, in the grammaticalization process the lexical item undergoes changes both in its structure and in its functionality.

Vitral's (1996) statement is explained in the terminology of Gonçalves, Lima-Hernandes and Casseb-Galvão (2007), by stratifying grammaticalization into three sub-processes:

> Grammaticalization is divided into three sub-processes: phonologization (changes in the phonic structure of words), morphologization (changes that affect the stem and affixes) and syntacticization (change that affects lexical categories, syntagmatic arrangements and the assignment of functions in the sentence) (GONÇALVES; LIMA-HERNANDES; CASSEB-GALVÃO, 2007, p. 16).

Besides the sub-processes, there are two forms of grammaticalization: either it is a *process*, when it is dedicated to the analysis of items that are already grammatical and become more grammatical; or it is a *paradigm*, when the intention is to focus the study to observe how grammatical constructions arise and how they are used. It can also be seen within the diachronic and synchronic perspectives, in the case of our object of study, *daí*, so we can understand it as a *paradigm*, from *diachronic* perspective.

According to Hopper (apud LOPES, 1991), one of the grammaticalization effects is persistence, that is, a lexical form or a lexical element can be retained for a long time. Due to its grammatical heirs, a word undergoes changes over time, or even in the way speakers use these words in everyday vocabulary. Grammaticalization is precisely concerned with this transformation undergone by the word, due to time, for example, as in the case of *you*, in which the ancients used the expression *Your mercy*, but, with adaptation or the popular use of the expression, *you* became more used. We have, so, a case of *paradigmatic* grammaticalization with a *diachronic* perspective, because the expression *you* arose from *Your mercy*, and the use by most of the speakers or even for a matter of adapting the language the old expression was extinguished and replaced with a new one.

Regarding the lexical item *daí* and its various contextual functions, our analysis turns to the grammaticalization of this item from the syntacticization point of view, considered in a panchronic perspective, since, for Tavares (2002), grammaticalization is an ongoing process, which makes it impossible to statically cut periods of time – they have indistinct boundaries and cannot be characterized as discreet and isolated slices (TAVARES, 2006).

Tavares (2006) makes her analyzes in a broad perspective in relation to the object under study, not only based on synchrony or on diachrony, but in both perspectives, what means, in a panchronic view. In this last perspective, the author analyzes texts

from the 13th to the 20th centuries, written in Portuguese, which are close to speech. The results obtained about the ancient texts were low, because only in one case (in Steinbeck's 1940 novel *As Vinhas da Ira*) there was a considerable percentage of occurrences of *daí*.

Instead, the analyzes made on the VARSUL data (Variação Linguística Urbana da Região Sul – Urban Linguistic Variation of the Southern Region) indicated an expressive number of grammatical cases of *daí* in the speech of the informants from Porto Alegre, Florianópolis and Curitiba.

Tavares (2006) believes that there are four mechanisms involved in the process of changing this item: the *metaphor* (process based on a relationship of inferred similarity between literal and figurative meaning); *metonymy* (a specific part is used to refer to the whole); *reanalysis* (reorganization and changes in linear rules); and *analogy* (similarity of particularities or functions). These concepts are used in grammaticalization and in the process of changes in each transformed grammatical form, whether through a transformation of a synchronic or diachronic perspective.

The use of *daí* in the speech of Curitibanos is frequent within different contexts. For morphology, *daí* is the junction of the preposition *de* + *aí* = adverb of time, but, in the speech, this expression gains other categories, that is, it starts to be used with divergent functions that grammars stipulate. Thus, grammaticalization occurs, which can be considered in the shape of a paradigm or in the shape of a process, able to be based on Saussurian dichotomy, in synchronic and diachronic perspective, and in panchronic perspective – as we saw in Tavares (2006).

Based on the occurrences of *daí* in the speech of Curitiba informants, according to ALiB database, we explore the various functions that this lexical item assumes in everyday speech and observe how grammaticalization happens in words (lexical items) that take on several shapes and meanings.

The understanding of the item *daí* can be approached by different areas, thus, we observed that in this one, in particular, both studies of linguistic variation and studies of textual articulators can present coherent subsidies for the use of *daí* in the speech of the informants analyzed in this research, since the oral text becomes a written text and then enables the study to be carried out.

Therefore, we can interpret the lexical item that undergoes the grammaticalization process, being a textual sequencer, as we saw earlier, as well as in the perspective of linguistic variation, as we will see below.

LINGUISTIC VARIATION THEORY

The Variationist Sociolinguistics considers that speaker characteristics are capable of influencing his way of speaking, and that influence was attested already in 1963, by William Labov, in his study carried out on Martha's Vineyard island, Massachusetts coast.

However, the term *Sociolinguistics* appeared only the following year, at a conference organized by linguist William Bright, at the University of California, Los Angeles (UCLA). In this conference, respectable models were present, such as John Gumperz, William Labov, John Fisher and others.

In 1966, in the text *The dimensions of sociolinguistics*, Bright defines the new area of study whose objective is "to show the systematic covariation of linguistic and social variations and, perhaps, even demonstrate a causal relationship in either direction" (BRIGHT, 1974, p. 17). Thus, for Sociolinguistics, diversity, or linguistic variation, is systematic and motivated by structural and social factors (MOLLICA, 2008). So, we emphasize that Sociolinguistics sees linguistic variation as "a constitutive quality of the linguistic phenomenon" (ALKMIN, 2005, p. 33), and not a defect.

Thus, we highlight that it is concerned with the "study of the spoken language, observed, described and analyzed in its social context, that is, in real situations of use" (p. 31), since

> Language and society are unquestionably linked to each other. More than that, we can say that this relationship is the basis of the constitution of the human being. The history of humanity is the history of beings organized in societies and possessing an oral communication system, that is, a language (p. 21).

We emphasize, therefore, that linguistic communities have a multifaceted character (SILVA-CORVALÁN, 1989) and it is precisely the language heterogeneity the object of study in Sociolinguistics (BRIGHT, 1974; SILVA-CORVALÁN, 1989; MORENO-FERNÁNDEZ, 1998; ALKMIN , 2005; CAMACHO, 2005; MOLLICA, 2008).

Linguistic variation is not random, but systematic and motivated by linguistic and non-linguistic factors. Among the non-linguistic, we highlight diachronic, diatopic, diassexual, diagerational and diastratic variations.

Diachronic or historical variation refers to changes occurred over time, which can happen at different linguistic levels, as explained by Ilari and Basso (2014). As an example, we have grammaticalization process and the inverse process, lexicalization.

Diatopic, or geographic variation, refers to the differences distributed in physical space, Alkmin (2005) attests. Like diachronic variation, diatopic variation occurs at different linguistic levels, such as lexical, phonological and morphosyntactic variations.

The analysis of diassexual variation should consider the social organization of the speech context, as this variable illustrates a very interesting scenario, since it points to "a greater female consciousness of the social status of linguistic forms" (PAIVA, 2015, p. 35). In this sense, we underline the relationship between sex, linguistic variation and the prestige attributed by the community to linguistic variants, as women tend to be at the forefront of linguistic change processes. However, it is essential to consider the social value of the innovative variant, since women tend to implement socially prestigious

forms in the language, while, alternatively, assume a conservative stance with regard to discredited variants. We emphasize, therefore, that women tend to prefer the most prestigious varieties, because

> when it comes to implementing a socially prestigious form in the language [...], women tend to lead in the change. Conversely, when it comes to implementing a socially discredited form, women take a conservative attitude and men take lead in the process (PAIVA, 2015, p. 36).

This scenario reflects the socio-cultural organization in which women are expected to use a more appropriate language em relation to standard.

In turn, diagerational variation may have language changing evidence. In this sense, this work is based on the study of change in apparent time, supported by the theory that an adult speaker reflects the state of the acquired language when he was approximately 15 years old (NARO, 2004). Thus, the speech of a person aged 60 represents the language of 45 years before, while the speech of a person aged 40 reveals the language of 25 years before.

Diastratic variation, on the other hand, refers to the sociocultural organization of the speakers, that is, it is linked to a set of factors related to the identity of the speakers and the sociocultural organization of the speech community.

The social class can be measured by several biases (MOLLICA, 2004), nonetheless, in this study, we consider school level as the marker of social class, since the school is able to inspire models and standards, effectively acting in the process of language change or maintenance. The school acts in the preservation of forms of prestige before changing trends and is characterized as a source for the knowledge of aesthetic and moral standards according to the conformity of saying and writing, Votre (2015) says.

This variable acts in different instances:

> The first is between a form of social prestige and a relatively neutral form. The second distinction is between the socially stigmatized phenomenon and the phenomenon immune to stigmatization. The third deals with the phenomena that are the object of school education and those that escape the normative attention of the school (VOTRE, 2015, p. 51).

The first distinction is made in the economic status and in the social prestige attributed to speakers of a given variant. Social prestige is opposed to discredit directed towards those who do not enjoy social and economic prestige. The socially prestigious way of speaking is established from the rules of the official literature and of the normative grammars. The second division is based on stigmatized speech and prestigious speech, thus stigmatized speech is the target of rejection. And yet, the third division focuses on the phenomena controlled by the school, that is, the imposition of normative grammar exercised by school environment, Votre (2015) explains.

Considering, therefore, the types of variables investigated in the linguistic variation, we will highlight diasexual, which presented greater evidence of *daí* use by women, diagenational and diastratic variables. For this study, diatopic analysis does not matter, since the data cover only one capital. We emphasize, however, that the conditioning factors do not act in isolation, but together and are interrelatedly.

RESEARCH LOCUS

The research relies on data collected in Paraná capital, Curitiba, which currently consists of eight urban districts: Portão, Cidade Industrial, Pinheirinho, Bairro Novo, Cajuru, Boqueirão, Boa Vista and Santa Felicidade. Its bordering cities are Almirante Tamandaré, Colombo, Pinhais, São José dos Pinhais, Fazenda Rio Grande, Araucária, Campo Largo and Campo Magro. According to IBGE (2010), Curitiba has 1,751,907 inhabitants, its territorial unit is 435,036 km² and its Municipal Human Development Index (MHDI) is 0.823.

The municipality, the state's economic center, is considered one of the country's best cities for business. Highly technological and industrialized, it does not depend on agro-industrial activities to move its economy. It has a wide network of preschools, elementary and high schools, as well as important higher education institutions such as the Federal University of Paraná (UFPR), the Federal Technological University of Paraná (UTFPR) and the Pontifical Catholic University of Paraná (PUCPR), which characterize the capital as a pole of higher education.

METHODOLOGICAL PROCEDURES

This paper relies on data from the Brazilian Linguistic Atlas Project (ALiB). The initiative for the elaboration of these took place at the Federal University of Bahia, in 1996, during "Caminhos e Perspectivas para a Geolinguística no Brasil Symposium" (Symposium on Paths and Perspectives for Geolinguistics in Brazil). ALiB National Committee is formed by the union of 16 Brazilian universities, subdivided into committees responsible for the development activities of the Brazilian Linguistic Atlas. The Project counted on the participation of 1,100 informants, spread over the 8,515,767 kilometers of Brazilian lands covered by ALiB Project research teams.

Out of curiosity, we highlight the Brazilian scenario, with regard to the publication of issues 1 and 2 of the Brazilian Linguistic Atlas, held in 2014 at the State University of Londrina. Upon completing the first publication of the 18-year work of the National Committee and the regional teams, the Brazilian Linguistic Atlas presents, in principle, issue 1 (CARDOSO et al., 2014a) with its introduction, its history, its methodology, its points network, its questionnaires and its informants, and issue 2 (CARDOSO et al.,

2014b), with the results of the 25 Brazilian capitals on linguistic maps with phonetic, morphosyntactic and semantic-lexical data. The publication of the third and fourth issues of ALiB is expected, which will present the broadest profile of the results, covering the entire national network.

Our corpus is concentrated in eight informants from the capital of the state of Paraná, Curitiba, whose profile obeys the requirements of ALiB project, being equally distributed among two age groups – 18 to 30 years old and 50 to 65 years old –, between literate men and women , who attended elementary school; in the capitals, four informants with a higher education level are added.

The Chart 1, below, presents ALiB informantes profile.

Chart 1 – Informants profile

Informant	School level	Age range	Gender
01	Elementary school	I (18-30)	Male
02	Elementary school	I (18-30)	Female
03	Elementary school	I (18-30)	Male
04	Elementary school	I (18-30)	Female
05	Higher education	II (50-65)	Male
06	Higher education	II (50-65)	Female
07	Higher education	II (50-65)	Male
08	Higher education	II (50-65)	Female

Source: Cardoso et al. (2014a).

For the studied phenomenon analysis, we observed its realization in response to the Phonetic-Phonological (QFF), Semantic-Lexical (QSL) and Morphosyntactic (QMS) Questionnaires and to the Themes for semi-directed speeches of ALiB project questionnaire (COMITÊ NACIONAL…, 2001).

Regarding the issues related to the Themes for semi-addressed speeches presented in the 2001 Questionnaires (COMITÊ NACIONAL…, 2001), we have:

Chart 2 – Semi-adressed questions

	Subject	Question
1	Personal report	Report an important event in your life (marriage, dating ...).
2	Comment	What television shows do you like the most? Why?
3	Description	What do you work on? Tell us a little about your work.
4	Non-personal report	Tell us a case/fact that you know (that you heard about, that happened to a friend, ad so on).

Source: ALiB data (2001).

Constituted from all the occurrences of *daí* as an answer given to the Phonetic-Phonological (QFF), Semantic-Lexical (QSL) and Morphosyntactic (QMS) Questionnaires and to the Themes for semi-directed speeches, this study takes into account different levels of speech monitoring, we mean, from a less monitored speech in the semi-directed speech to the most monitored in the responses to the QFF.

The data were quantified according to the Linguistic Variation Theory, that is, we consider the linguistic data of the speech of men and women, stratified in two age groups (Chart 1) and with different levels of education. The corpus corresponds to the data of ALiB Project, in which we proceeded with the following steps: i) first we performed the data collection through the audios provided by the Project, using Winamp player to run the recordings; ii) next, we check the audios with the transcriptions and collect our object of study, the lexical item *daí*; iii) we used Microsoft Excel 2013 spreadsheet as a way to organize the data, in which we selected the variants according to the informant's profile – an important procedure for counting the data and better visualizing the results; and finally, after the collection and organization of data, we moved on to iv) production of graphs according to the variables worked on, taking into account the sex, the school level and the age of the informants.

We emphasize that, according to the analysis, the use of *daí* is more frequent in the speech of women of both age groups and schooling, therefore we chose to present graphs 2 and 3 addressing this preference. As a consequence, we elaborated a graph for the diasexual variable, a graph for the diagerational variable and one for the diastratic variable, with the exposure of the sexes, in order to reinforce the greater preference of the variant in female speech at all levels of variables.

The description and analysis of data is divided into two stages: the data according to the microdomains and, then, according to the social variables.

Data from the Microdomains View

In this paper, the data under the microdomains view is presented with a more descriptive character, and for that we adopt the concepts of Tavares (2006) with regard to the five microdomains: (i) textual sequencing: cohesive strategy that marks the order in which the connected units succeed each other throughout the discursive time; (ii) temporal sequencing: information introduced that succeeds temporally in relation to the information already given; (iii) introduction of effect: information introduced that represents a consequence or a conclusion in relation to what was previously said; (iv) resumption: movement to recover the previous thematic flow, interrupted by a digression; and (v) finalization: signaling the end of the topic/subject in progress. We will exemplify such concepts with corpus data to facilitate the understanding of each one:

(1) INQ.- Nossa! Vocês nasceram pertinho uns do outro, né?
INF.- É. (risos). Vinte sete, meu irmão tem vinte e quato, eu tenho vinte e um, *DAÍ* vem, tem oto cum vinte, dezenove, *DAÍ* o Val, Valdinei com dezesseis, *DAÍ* o Ednei cum doze, e o Júnior cum nove. (Curitiba, informante 1).

(2) INF.- (risos) Ah, é simples. Ah, torta de maçã tem que fazê aquela massa, né, tipo a de empadão. *DAÍ*... faiz... faiz... tipo assim na massa, da, da, pra torta, né, eu uso trigo, duas colheres de margarina, meia de... de... meia xícara de leite, uma pitada, uma meia, uma meia colher de açúcar e bate bem ela, tipo assim, mexo bastante, sabe? dois ovos, até ela ficá aquela massa consiten... consistente, que *DAÍ* vai espalhá. E *DAÍ*, eu faço... o recheio, né? Que é maçã. Coloco a maçã... primeiro eu bato a maçã, né? Que... depois, depois coloca a maçã, aqueles pedaços que a gente vê, né. (Curitiba, informante 1)

In (1), the studied element can be considered within the microdomain of (a) textual sequencing: in the first moment, *daí* precedes the verb *come* in the indicative present (*comes*). Thus, we can associate the item with the adverb function, as it is "directing" the verb, however we cannot say that the position in which *daí* is found is that of the adverb, as it is not modifying the verb. The next two occurrences of *daí* can be interpreted as connected units that succeed each other throughout the discursive time.

In the three occurrences of *daí* in example (2), we find that the first indicates the text sequence, in which the informant is explaining/teaching how to make a recipe, therefore the first time the lexical unit appears is indicating the order of the units

connected. In a second moment, the term appears with an effect introducing meaning – the information entered represents a consequence or a conclusion in relation to what was previously said. Let's see: "até ela ficá aquela massa consiten... consistente, que daí vai espalhá." The consequence of the spreading of the dough is due to its consistency. We note that the use of *daí* really introduced an effect to the informant's speech, and in the third case *daí* works as a textual sequencer.

The use of *daí* in the resumption and finalization functions can be observed in the discursive sequences 3, 4, 5 and 6:

(3) INQ.- Senac?
INF.- Num é. Num é, num tem nada a vê de sési. É, é um... é, tipo assim, uma casa mesmo, sabe, que ensina. Então, isso *DAÍ* foi a... foi o meu patrão que me colocô lá, que eu falei pra ele que eu, que eu quero... num nasci pá sê... pá trabalhá em escritório. (Curitiba, informante 1).

(4) INQ.- Quando uma criança está dormindo e não se quer acordar, a gente diz: "Fale baixo, não faça...'"?
INF.- Não faça barulho, que se não vai acordá a criança, né. *DAÍ* acorda reinando, né. (Curitiba, informante 3).

(5) INQ.- Tem alguma coisa aí né ?
INF.- Canoa cê... cê vai no... aliás, o bote, o bote vai no remo, né, e a canoa, motor. É chamado praticamente... geralmente ele é chamado de bote, né, mas muitos já falam... canoa, canoa no caso seria um bote maior, né, que um bote já é peque...é uma estrutura menor, é pequeno, né, que vai fácil no remo, e a canoa já é maior, então... o documento é o motor *DAÍ*, né. (Curitiba, informante 3).

(6) INF.- Médio, nem alto nem baixo.
INQ.- Porquê?
INF.- Ah, pra não incomodá os vizinho, né, alto e baxo a gente também num dá nem pra escutá *DAÍ* né. (Curitiba, informante 4).

In (3) and (4), we have *daí* as a resumption, performing a function of recovering what was said earlier: "É, é um... é, tipo assim, uma casa mesmo, sabe, que ensina. Então, isso daí foi a... foi o meu patrão que me colocô lá." The resumption that *daí* represents in the sentence is to reinforce that it was in a house, and not at SESI, that he learned cooking. When the informant says "Então, isso daí...", he is resuming what he said earlier, affirming and attesting what has already been said in order to comprise all the information already mentioned in the speech, that is, *daí*, in this case, encapsulates the data already said, as a kind of "box", where information is stored in a summarized form, allowing a preceding period of the text to be condensed, concentrated and taken

up through a single expression that becomes its referent: *isso daí*, in a way that the demonstrative pronoun *isso* (this) aids in this process.

In example (4), the expression can be understood as an effect introducing, that is, making noise will result in the child waking up.

The finalization is exemplified in the excerpts (5) and (6): both in one and in the other, *daí* precedes *né* (isn't it?). We could say that the two are speech resources, however, we interpret them as inserted in the final microdomain, as it is concluding the speaker's speech and affirming, or interacting with the receiver so he agrees or disagrees with his statement, which is reinforced by *né*.

Tavares' concepts (2002) are effective in analyzing expression. At first sight, the term *daí* could easily be called a connector, since, in most instances, the term connects sentences. Although this denomination has not yet been established by Normative Grammar, its frequency in speech will provide a gradual grammaticalization of *daí*.

After analyzing the 492 occurrences of *daí*, under the perspective of the five microdomains proposed by Tavares (2002), we verified the performance of the factors sex, age range and school level on the amount of occurrences, and not on the frequency of functions.

Data from the Extralinguistc Facts View

In addition to the analysis of *daí* under the macrodomains perspective, we turn to the observation of the phenomenon through social variables from the perspective of the Linguistic Variation Theory. In this sense, we analyze how the variables sex, age range and education level can influence the productivity of discursive sequences with *daí*.

Altogether, we counted 492 occurrences of *daí* and started our analysis according to the gender variable, which is shown in Graph 1.
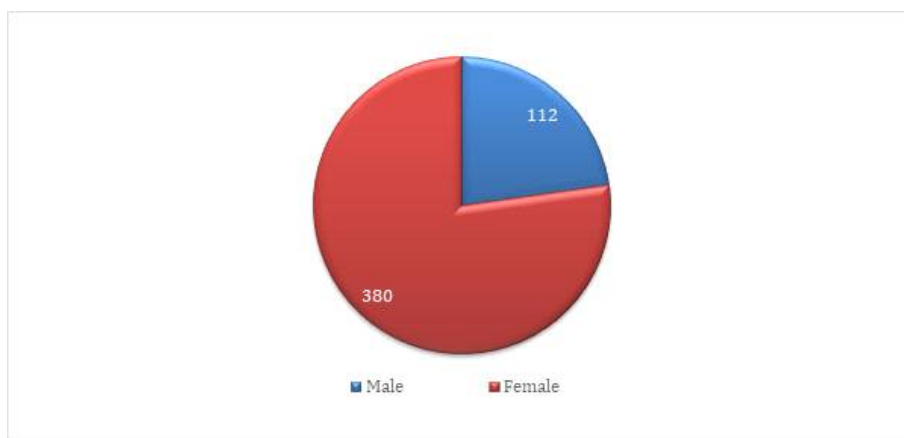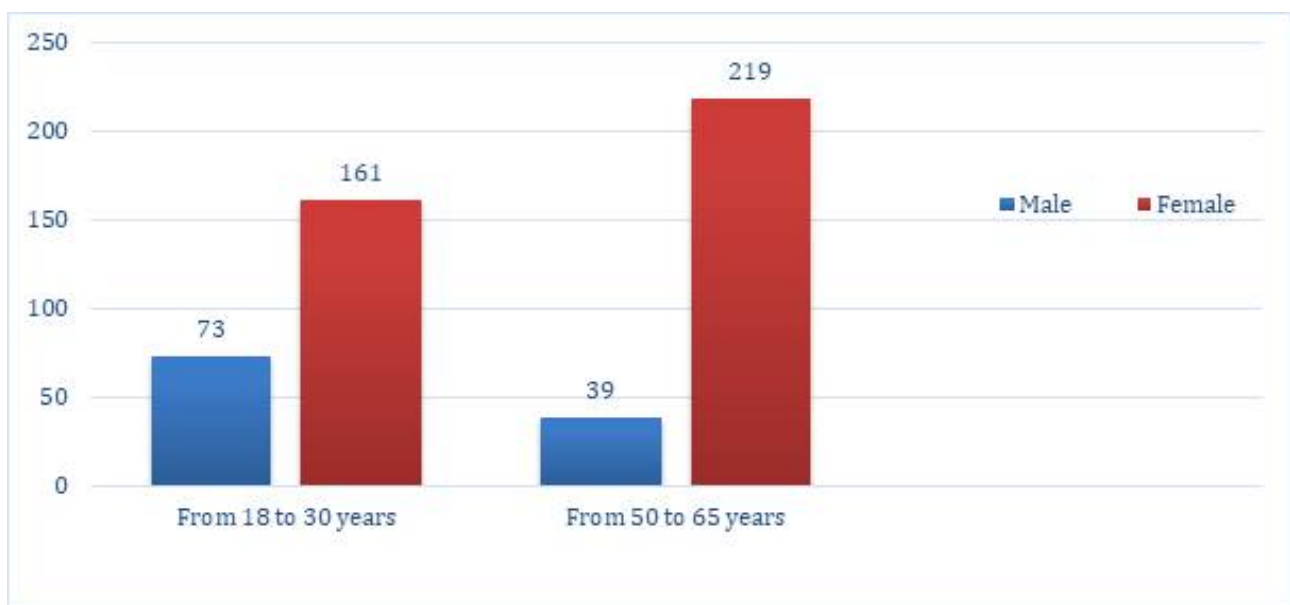

Source: ALiB data.

Graph 1 – Number of occurrences of the term *daí* according to the diasexual variable

In Graph 1, we present the data according to the variable gender. We noted that women performed 380 occurrences, while men performed 112. This context makes it possible to state that the use of *daí* is much more recurrent in the speech of female informants and can be considered a marker of the female speech of Curitiba informants. For that matter, we emphasize that the analysis of the variation between sex and linguistic variation reflects a relevant scenario due to the social organization of the context in which the speech is inserted. When considering the statement by Paiva (2015) that the speech of women is directed to the status attributed to the form, we can consider a certain prestige to the term *daí* in the spoken expressions, in addition to enabling an affirmation that the expression is gradually gaining space in a process of linguistic change, that is, of its inclusion. This socio-cultural construction is advocated by constructions in which women are expected to have a more appropriate and standards-oriented language.

We point out that the item occurred in the responses to the most controlled Questionnaires, the Phonetic-Phonological (QFF) and the Semantic-Lexical (QSL), but it was much more productive in the Themes for semi-directed speeches, that is, in a more informal speech modality, and, therefore, less monitored, being used in the functions of the sequencing microdomain.

Then, we cross the data considering the age range and the sex of the informants, as shown in Graph 2.

Source: ALiB data.

Graph 2 – Number of occurrences of the term *daí* according to the diasexual and diagerational variables

Computing the absolute numbers, we have 73 achievements of *daí* in the speech of young men and 161 in the speech of young women, which gives a significant difference between both groups, because, of the 234 records, men performed only 31.2% of *daí*, while young women did it in 68.8% of cases. If, on the contrary, we compute the data of the second age range, for men, we will have 39 achievements, and for women, 219, demonstrating, once again, the predominance of the item *daí* in the speech of women, since in the age range II we found 84.9%, in contrast to the insignificant presence in the speech of men of the same age group, with 15.1%.

In this sense, we can see that the realization of *daí* is much more dependent on the gender variable than on the age range variable. However, we note that the difference between the achievement of men and women in the age range I, with 88 achievements, is smaller than in comparison to the age range II, with 180 achievements. This scenario makes it possible to infer that the expression *daí* has gradually been gaining space in the speech of the informants.

The next step concerns the analysis of the influence of the variables sex and school level on the use of *daí*, as shown in Graph 3.
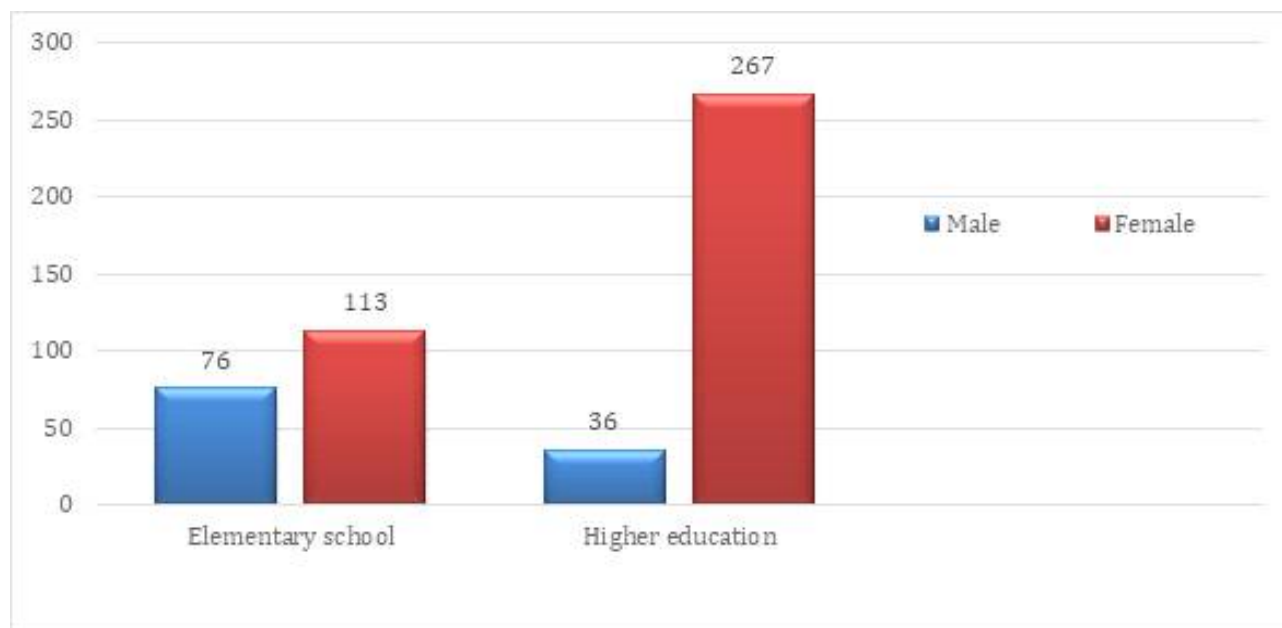
Source: ALiB data.

Graph 3 – Number of occurrences of the term *daí* according to the diasexual and diastratic variables

Analyzing the data in Graph 3, we check that women with higher education have a high frequency of use of *daí*: 267 achievements, while men with the same level of

education register only 36 occurrences. When the gender and the school level variables are crossed, elementary school women also lead the use of *daí*, with 113 records, whereas in the speech of men of the same level we obtained 76 records, results that are closer to each other than among the informants of higher education.

Thus, we highlight that the realization of *daí* presents an interesting scenario when observing the crossing of the school level variable and the gender variable, since the difference in the speech of the informants of elementary level presents only 37 occurrences, while the difference between informants of the highest level is 231 achievements. This background allows us to infer that women with higher education have the greatest indications of the phenomenon's realization, bearing the mark of achievement endowed with linguistic prestige.

CONCLUSIONS

The study demonstrated that the expression *daí* can be used with several meanings, depending on the context in which it is inscribed. Limiting it as just an adverb of time, as normative grammars do, would not be the appropriate or the only function of it in the speech, that is, we understand that the expression assumes a polysemic characteristic, depending on the situation in which it is being used, thus we were able to associate the meanings of *daí* based on the notion of textual sequencers (TAVARES, 2006), as explained in the analysis.

In addition to the concepts of Tavares (2006), there are other theories that explain the various meanings present in the Lexicon, one of them is the concept of context applied to meaning, exposed by Stephen Ullman, in which he presents Malinowski's ideas as a basis:

> The concept of context [...] must burst the bonds of mere linguistics and be carried over into the analysis of the general conditions under which a language is spoken... The study of any language, spoken by a people who live under conditions different from our own and possess a different cultur, must be carried out in conjunction with the study of their culture and of their environment (MALINOWSKI apud ULLMAN, 1964, p. 50).

Ullman also rests on Wittgenstein: "he does not merely say that we can establish the meaning of a word by obsreving its use; he boldly asserts that the meaning of a word *is* its use" (MALINOWSKI apud ULLMAN, 1964, p. 64).

The meaning of the words also depends on the influences suffered according to the context in which it is used, that is, in the author's conception, the words extract from the context certain determinants that will only make sense if they are in specific situations of use. Ullman exemplifies the fact by stating that "Even proper names, the most concrete of all words, have a variety of aspects only one of which will be relevant

to a particular situation (MALINOWSKI apud ULLMAN, 1964, p. 52). Furthermore, the ideas of Wittgenstein and Malinowski, endorsed by Ullmann, present the relationship between the meaning of the word and its dependence on usage.

Thus, we interpret that *daí*, besides being characterized in the microdomain of sequencing, employing several functions within the speech of Curitiba, which represents the use and its various possibilities of meaning, can serve as a basis for other studies with the objective of identifying marks in the speech of men and women.

From the analysis according to social variables, we observed that women performed 380 occurrences of *daí*, while men only 112, a scenario that allows the statement that the use of *daí* is much more recurrent in the speech of female informants, which can be considered a marker of female speech in Curitiba.

When considering the crossing of the diasexual and diastratic variables, we found, in the age range I, 73 achievements of *daí* in the speech of young men and 161 in the speech of young women, computing a significant difference. Similarly, in the age range II, for men, we have 39 achievements, and for women, 219, demonstrating, once again, the predominance of the item *daí* in the speech of women. Still, we verified that women with higher education have a high frequency of use of *daí*, with 267 achievements, while men with the same level of education performed only 36 occurrences.

Thus, we might think *daí* as a discursive mark in the speech of women in the age range II, with a higher level of education, which may indicate that the change has not been consolidated, but the frequency of its usage remains at the level of the diageneric and diastratic variations.

Considering the speech styles, we may say that the occurrences of *daí* in the corpus were registered both in the spontaneous style (themes for semi-directed speeches) and in the more controlled style (of questions and answers), with predominance in the spontaneous style. This scenario is due, in particular, to the less rigid form of discourse.

References

ALKMIN, T. M. Sociolingüística. *In*: MUSSALIM, F.; BENTES, A. C. *Introdução à lingüística:* domínios e fronteiras. 5. ed. São Paulo: Cortez, 2005. p. 21-47.

BRIGHT, W. As dimensões da sociolingüística. *In*: FONSECA, M. S. V.; NEVES, M. F. (Org.). *Sociolingüística.* Rio de Janeiro: Eldorado, 1974. p. 17-24.

CAMACHO, R. G. Sociolingüística parte II. *In*: MUSSALIM, F.; BENTES, A. C. *Introdução à lingüística:* domínios e fronteiras. 5. ed. São Paulo: Cortez, 2005. p. 49-75.

CARDOSO, S. A. M. O Projeto ALiB e sua trajetória. *In:* MOTA, J. A.; CARDOSO, S. A. M. (Org.). *Documentos 2:* Projeto Atlas Lingüístico do Brasil. Salvador: Quarteto, 2006. p. 27-34.

CARDOSO, S. et al. *Atlas Linguístico do Brasil*. Introdução. Londrina: EDUEL, 2014a. v. 1.

CARDOSO, S. et al. *Atlas linguístico do Brasil*. Cartas linguísticas 1. Londrina EDUEL, 2014b. v. 2.

CASTILHO, A. A gramaticalização. *Estudos Lingüísticos e Literários,* Salvador, n. 19, p. 25-63, mar. 1997a.

CASTILHO, A. Língua falada e gramaticalização. *Filologia e Lingüística Portuguesa,* n. 1, p. 107-120, 1997b.

COMITÊ NACIONAL DO PROJETO ALiB. *Atlas Lingüístico do Brasil*: questionários 2001. Londrina: Ed. da UEL, 2001.

CUNHA, A. G. da. *Dicionário Etimológico Nova Fronteira da Língua Portuguesa*. 2. ed. Rio de Janeiro: Nova Fronteira, 1994.

FERREIRA, A. B. de H. *Novo dicionário de Língua Portuguesa*. 2. ed. Rio de Janeiro: Nova Fronteira, 1986.

GONÇALVES, S. C. L.; LIMA-HERNANDES, M. C.; CASSEB-GALVÃO, V. C. (Org.). *Introdução à gramaticalização:* princípios teóricos e aplicação. São Paulo: Parábola, 2007.

IBGE Cidades. *Curitiba*. Disponível em: https://bit.ly/2ZUeoAwCTBA. Acesso em: 12 jun. 2019.

ILARI, R.; BASSO, R. *O português da gente:* a língua que estudamos, a língua que falamos. 2. ed. São Paulo: Contexto, 2014.

KOCH, I. V. *A inter-ação pela linguagem*. 11. ed. São Paulo: Contexto, 2015.

LABOV, W. *Padrões sociolinguísticos*. Tradução de Marcos Bagno, Maria Marta Pereira Scherre e Caroline Rodrigues Cardoso. São Paulo: Parábola, 2008 [1972].

LOPES, C. *Gramaticalização:* definição, princípios e análise de casos. Niterói: EDUFF, 2008.

MOLLICA, M. C. Fundamentação teórica: conceituação e delimitação. *In*: MOLLICA, M. C.; BRAGA, M. L. (Org.). *Introdução à Sociolingüística:* o tratamento da variação. São Paulo: Contexto, 2008. p. 9-14.

MORENO-FERNÁNDEZ, F. *Princípios de sociolingüística y sociologia del lenguaje*. Barcelona: Ariel, 1998.

NARO, A. J. Modelos quantitativos e tratamento estatístico. *In*: MOLLICA, M. C.; BRAGA, M. L. (Org.). *Introdução à Sociolingüística:* o tratamento da variação. São Paulo: Contexto, 2008. p. 15-25.

PAIVA, Maria da Conceição. A variável gênero/sexo. *In:* MOLLICA, M. C.; BRAGA, M. L. (Org.). *Introdução à Sociolingüística:* o tratamento da variação. São Paulo: Contexto, 2008. p. 33-42.

SANTOS, C. D. dos. Marcadores discursivos nas entrevistas retextualizadas. *In*: STORTO, L. J.; NAKAYAMA, J. K.; BURGO, V. H. *Texto, contexto e discurso:* homenagem a Paulo de Tarso Galembeck. Curitiba: Appris, 2014. p. 187-209.

SAUSSURE, Ferdinand. *Curso de lingüística geral.* 26. ed. Tradução de Antônio Chelini, José Paulo Paes e Izidoro Blikstein. São Paulo: Cultrix: 1995.

SILVA-CORVALÁN, C. *Sociolingüística:* teoría y análisis. Madrid: Alhambra Universidad, 1989.

TAVARES, M. A. Daí (veio) o uso do daí na codificação lingüística da articulação discursiva. *Working Papers em Lingüística*, n. 6, p. 92-115, 2002.

TAVARES, M. A. Abordagem pancrônica à gramaticalização de daí como conector. *Signum:* Estudos da Linguagem, Londrina, n. 9, v. 2, p. 245-271, dez. 2006. Disponível em: https://bit.ly/2BkuqcKSigTav.

ULLMANN, S. *Semântica.* Uma introdução à ciência do significado. 5. ed. Tradução de J. A. Osório Mateus. Lisboa: Fundação Calouste Gulbenkian, 1964.

VITRAL, L. A forma CÊ e a noção de gramaticalização. *Revista Estudos da Linguagem,* Belo Horizonte, v. 1, n. 4, p. 115-124, jan./jun. 1996.

VOTRE, S. J. Relevância da variável escolaridade. *In:* MOLLICA, M. C.; BRAGA, M. L. (Org.). *Introdução à Sociolingüística:* o tratamento da variação. São Paulo: Contexto, 2008. p. 51-57.