Abstract:
This paper points to the theoretical and methodological issues involved in the subproject on Portuguese Language Atlas in Indigenous Areas (ALiPAI) part of the project GeoLinTerm coordinated by Abdelhak Razky (UFPA/UnB). The first results of ALiPAI are part of the doctoral thesis of Guedes (2017), which mapped the geosociolinguistic profile of Portuguese in contact with Tupí-Guaraní languages in indigenous areas of Pará and Maranhão. The first experiences in the geolinguistic field research in Brazilian indigenous areas provided a confluence of research methodologies, especially Geosociolinguistics (RAZKY, 1998), Pluridimensional and Relational Dialectology (THUN, 1998) and Anthropological Linguistics (RODRIGUES, CABRAL, 2012).
Ten informants were selected from each of the five network of points investigated (Suruí Aikewára, Asuriní do Tocantins, Tembé, Guajajára and Guaraní Mbyá). The questionnaires applied in these areas were developed by the National committee of the Linguistic Atlas of Brazil - ALiB: Phonetic-Phonological questionnaire (QFF) and Semantic Lexical Questionnaire (QSL). These questionnaires were adapted to include the correspondence request in the indigenous language for each one of the answers obtained in Portuguese. In addition, a complementary QFF and a Sociolinguistic Questionnaire were used. The paper also reflects on the necessary adaptations made in the methodology of geolinguistic research, to account for the geosociolinguistic characteristics of the ALiPAI target communities.

Keywords:
Geolinguistic Research in Brazilian Indigenous Areas: challenges and strategies

Abdelhak Razky; Regis José da Cunha Guedes; Eliane Oliveira da Costa

INTRODUCTION

Great advances have been achieved both theoretically and methodologically in Brazilian geolinguistic studies. Several papers described this reality by distinguishing the phases of the dialectal studies based on the research carried out in the form of linguistic atlases and monographs.

However, these studies about Brazilian Portuguese (PB) didn’t take into consideration some sociolinguistic spaces which would motivate the expansion of the geolinguistic practice to get a real view of PB, a language spoken in a country open to new technologies and characterized by linguistic diversity.

In this context, the group of researchers from the Geosociolinguistics and Socioterminology project (GeoLinTerm) sought to expand its research perspectives through the project of the Linguistic Atlas of Portuguese spoken in Indigenous Areas (ALiPAI). The project tends to broaden the horizons of geolinguistic research by focusing on indigenous Brazilian areas starting by the states of Pará, Maranhão, Amapá and Amazonas.

Regarding the relationship between Portuguese and indigenous languages, it is important to point out that “the study of the possible phonological and grammatical influences of indigenous languages in Brazilian Portuguese should be done not only on the basis of a good knowledge of the language or indigenous languages involved, but also taking into account the particular dialects of Brazilian Portuguese “and that until now we have no studies combining knowledge of both the indigenous languages and the dialects of Brazilian Portuguese” (RODRIGUES, 2014, p. 11).

The ALiPAI project intends to fill this gap in the context of geolinguistic studies and, like any pioneer project, has presented a series of theoretical and methodological questions, which are the object of this paper. The text is structured in four sessions: (1) typologies involved in the fieldwork of traditional dialectology and geosociolinguistics (2) the concepts of isogloss and grouping, and their implications for fieldwork (3) fieldwork in monolingual and bilingual / plurilingual areas and (4) perspectives for fieldwork in indigenous areas, followed by the final considerations.

TYPOLOGIES OF FIELDWORK IN TRADITIONAL DIALECTOLOGY AND GEOSOCIOLINGUISTICS

In traditional dialectology, the methodology of fieldwork was already based on the tripod: network of geographical points, informants and questionnaires. In the
geosociolinguistic dimension, when dealing with the network of geographical points and the use of questionnaires, one observes a continuity among these instruments. However, some methodological adaptations have been made, especially in the diversification and application of the questionnaires. On the other hand, with regard to the profile of the informant, the geosociolinguistic approach predicts an improvement of field observation for this profile.

According to Zagari (1998, p. 36), in traditional dialectology the normally required informant was called MAFIS (“man, adult, farm worker, illiterate and sedentary”). This profile was chosen in order to register the most “preserved” variants of the language in a given space. However, in view of the rapid changes observed in society, especially since the twentieth century due to an accelerated growth of urban areas (rural exodus) motivated by the search for better conditions of employment, a different perspective is required. In this context, the first studies in the field of urban dialectology and sociolinguistics, which aimed to investigate the new linguistic configurations of variation in this new social context, the MAFIS informant profile became obsolete.

Another difference is related to the number of informants per each geographical point to be investigated. In traditional dialectological studies, one or two informants per point were usually considered. This has historically been a reason for criticizing geolinguistic methodology that has focused for a long time on the study of diatopic variation, thus favoring the expansion of the number of geographical points of inquiry to the detriment of the number of informants per point. The emergence of modern dialectological studies in the second half of the twentieth century brought about an increase in the number of informants from two to at least four per locality. This change was motivated by the interest of dialectologists in mapping social variables: diageneric, di generational, etc. However, this number is still considered low when compared to the sample of informants taken into account in many of the sociolinguistic studies. In order to minimize this aspect, some studies in modern dialectology increased the number of informants as in the case of ALiB (CARDOSO et al., 2014) by including the diastratic variable (schooling) in Brazilian capitals.

In the scope of the geosociolinguistic studies, like the one realized in the doctoral thesis of Guedes (2017), the number of informants was increased. This study tried to map ten respondents in each geographical point of inquiry by including two children to the sample, a choice that was motivated by the nature of the research, which dealt with the vitality of the indigenous languages and the variation of the Portuguese spoken in indigenous Brazilian areas.

THE CONCEPTS OF ISOGLOSS, GROUPING AND THEIR IMPLICATIONS FOR FIELDWORK

Since the mid-nineteenth century, traditional dialectologists were concerned by describing the diatopic variation of languages, especially delimiting the geographical boundaries between languages, dialects and/or speech in a given region. Hence came the
conception of isogloss as an imaginary line drawn on a map in order to establish the geographical boundaries of linguistic phenomena.

Isoglosses can be classified according to the nature of the linguistic phenomenon as isolexical, when delimiting the lexical variation in a certain region; or isophonic, when they refer to the phonetic variation, etc.

This technique of representation was sufficient for the configuration of the geographical distribution of linguistic phenomena in a territory. It was a moment in the history of mankind in which distances and geographical features (seas, rivers, mountains, etc.) and climate (frost, desert heat, etc.) actually isolated communities, or at least hindered frequency of communication between people, which, as is well known, is a factor that directly influences diatopic linguistic variation.

However, the design of isogloss lines has become out-of-date for the representation of the reality of linguistic variation, since the homogeneity that these imaginary lines proposed to represent has increasingly become a historical fact. This is the outcome of the great changes undergone by a society which is increasingly globalized, especially with the development of communication and transportation since the twentieth century.

The reduction of distances between populations is a characteristic of modern society. It’s result of the development of more efficient means of transportation such as cars, trains and airplanes; democratization of access to media such as radio, television; and more recently, the internet and smartphone, which in turn have led to the emergence of social networks that combine several of these technologies.

In this context where communication between people is intense, there is a dilution of the geolinguistic borders, once represented by the isoglosses, especially in the more urbanized regions and with greater access to these means of communication and transportation.

In the mid of geosociolinguistic research, from our work in Guedes (2012) and especially Razky and Guedes (2015), which mapped the lexical variation in the Audio-Lexical Atlas of Pará (ALeSPA) data, we developed the concept of diatopic lexical groupings, since we observed that some lexical items were grouped in specific regions of the territory of Pará, in the form of macro-groupings and micro-groupings that form spatial intersections, as can be seen in figure 1.
The diatopic distribution of the lexical variables presented in the lexical map 105 (figure 1) made it possible to circumscribe diatopic lexical groupings in the geographical space of the state of Pará. These groups represent the predominance of some lexical items in certain geographic spaces in relation to others in which either there is no occurrence of these mapped lexical variants, or their frequency is very low. In figure 1, for example, the micro-grouping Mi4 presents the predominant zone of the item “tiú” in the mapped territory. However, the same item is also registered in the point of inquiry 8. This situation is similar to that of the micro-grouping Mi3, which circumscribes the area of “chameleon” variant, which is also found in point 11. It is precisely this representation of the dilution of rigid boundaries proposed by the isoglosses which makes the notion of grouping more faithful to the reality of the variation of linguistic phenomena in a geographical space.
We understand that this geographical configuration reflects the existence of historical isolexical zones in the territory of Pará. This configuration, in turn, is related to the forms of settlement in the region. However, the rigidity of the geographical limits was diluted due to several migratory flows that occurred at different times, from the first inhabitants of the territory (the indigenous people of several ethnic groups who migrated there for centuries) to the different historical moments of the occupation of the Amazon region by Europeans and by the mixture of white and indigenous people.

Describing the lexical variation of the Portuguese spoken in the region by means of the concept of diatopic lexical groupings represent a more adequate way to depict the reality of the linguistic facts studied. Since the diatopic homogeneity intended in the making of the isoglosses does not really correspond to the reality of the region, it’s possible to ask if this homogeneity was anything other than a myth.

**FIELDWORK IN MONOLINGUAL AND BILINGUAL / PLURILINGUAL AREAS**

The study carried out by Guedes (2017) on the Portuguese spoken in indigenous areas of the states of Pará and Maranhão allowed for a reflection on the multiculturality of Brazilian indigenous areas. This new approach in geolinguistic studies dealing with the diversity of languages spoken in the indigenous areas (Portuguese and indigenous languages) has allowed us to rethink the methodologies of fieldwork in the area of geolinguistics. It is an integration of different methodological traditions: the studies of Geosociolinguistics and Pluridimensional Dialectology (dealing mainly with variation in monolingual areas) and those of Anthropological Linguistics (which traditionally studied the languages spoken by traditional communities, sometimes bilingual, and even multilingual areas).

The expansion of geolinguistic studies, especially from the perspective of Pluridimensional and contactual Geolinguistics (THUN, 1998), led to the creation of linguistic atlas projects in Brazil that aim to map more than one language, such as frontiers communities in which there is contact of speakers from different languages in the Brazilian territory (GUEDES, 2017, p. 78).

Guedes (2017) resumes the assertion of Thun (2014) that in geolinguistic research in Brazil, like the methodology adopted in the fieldwork of the ALiB project, the knowledge about an indigenous language is not asked about. The data are treated as if Brazil were a monolingual country, but it is not. Around 180 indigenous languages are currently spoken within the geographical limits of Brazil (RODRIGUES, 2006), they coexist in different situations of contact. From this and other questions that emerged during the fieldwork, the methodology used in the research about indigenous areas needed some adaptations that are exposed in the following section.
The doctoral thesis defended by Guedes (2017) was a first step in the realization of the ALiPAI project. It is a pioneer project since it aimed at mapping Portuguese spoken in indigenous areas in Brazilian territory, a new subject within geolinguistics in Brazil. None of the state, regional and small domain atlases already elaborated, as well as the Linguistic Atlas of Brazil, included the mapping of Portuguese spoken in indigenous areas either as an option or because of methodological restrictions of the projects, or by the difficulties of access to these areas.

The ALiPAI project tries to fill this gap. In this first stage, the objective was to identify, analyze and map Portuguese in contact with indigenous languages of five ethnic groups in the states of Pará and Maranhão in order to build a database that provides information on the phonetic and lexical variation of Portuguese in situations of linguistic contact with indigenous languages of branches one and four of the Tupí-Guaraní Family, namely: Guaraní Mbyá, Suruí Aikewára, Asuriní do Tocantins, Tembé and Guajajára.

In these first experiences of doing fieldwork in indigenous areas, real methodological challenges were faced. Examples of this include: the need to collect data in locations that are difficult to reach (terrestrial, aquatic and aerial means of transportation have been used); the distance of indigenous lands from urban centers; the conditions of the roads, that in the “Amazonian winter” are practically impossible to face; the need for authorizations from FUNAI and/or indigenous leaders to gain access to indigenous lands. All this led us to substitute a programmed point of inquiry “Urubu-Ka’apor” by the “Anambe” community. In all the research points surveyed, it was a challenge to contact indigenous leaders in order to obtain access authorization, a factor that, in itself, constitutes a big challenge for geolinguistic studies in Brazil, which are generally carried out in communities without legal access restrictions.

Another factor that emerged from the research context was the fact that the indigenous lands studied constituted bilingual or plurilingual spaces, since, in addition to the Portuguese language, the indigenous people belonging to the studied ethnic groups sometimes spoke more than one indigenous language. This led to include the mapping of the dialingual variable in the study.

We also needed a methodological adaptation to control the time used for data collection. We were instructed to be cautious when contacting the natives to select collaborators for the interviews. This demanded a greater time in relation to our previous experiences in geolinguistic fieldwork in monolingual areas, in which the investigator does not face, for example, linguistic barriers caused by the degree of bilingualism of the collaborator, or by the time necessary for the good acceptance of the researcher by the community. The challenge in this regard was to take into account these factors, to accelerate the process of the selection of collaborators and data collection in order to carry out our investigation in the other geographical points fixed in the project. The average time of each
Another methodological aspect was the replacement of the nomenclature “informant”. We chose to use the term “collaborator” to define the source of the information in the research. Our option was based on a negative semantic charge imposed on the term “informant” in the context of scientific research in indigenous lands. Due to the socio-historical context of contact and exploitation of the Brazilian indigenous peoples, collaborators felt as mere suppliers of information about their ethnic groups, cultures, the languages they speak, the environments in which they live, the nature, the climate, the geographical features, etc. In fact, the collaborator’s role goes beyond this passive posture, it’s a necessary collaboration for geosociolinguistic fieldwork for both the researcher and collaborator to work together for the construction of scientific knowledge.

The collaborators of the ALiPAI project are selected according to the following parameters: ten collaborators per geographical point of inquiry. The age factor included three age groups: 1st - from 5 to 10 years old, 2nd - from 18 to 37 years old, and 3rd - from 47 to 75 years old in order to map the representation of the youngest and the oldest. The sex variable comprised two children (a boy and a girl), two men and two women in the second age group, and two men and two women in the third age group in each locality. This is an innovation among dialectological studies in Brazil since we have not traditionally taken into account the speech of children. It was done here due to both the linguistic variation and degree of bilingualism within the indigenous communities.

As for the diastratic variable, two levels of schooling were taken into consideration. At the first level two men and two women having a level of instruction lesser or equal to the 9th year of Middle School, and two men and two women taking the 1st year of High School (including those who are taking undergraduate courses or those who are already graduate in intercultural studies for example).

With regard to the two children representing each ethnic group, the schooling factor was not taken into consideration, since all the selected children are regularly enrolled in village schools. The mapping of the speech of these children contributed to the quantification of the geosociolinguistic data regarding the phonetic variation of Portuguese, the degree of bilingualism and the linguistic behavior of the collaborators in relation to Portuguese and indigenous languages of each ethnic group.

The pluridimensional approach adopted in the ALiPAI required the elaboration of linguistic maps that registered the variables taken into account (diatopic, diageneric, diagenerative, diastratic and dialingual), as in figure 2:
In Map D01 (figure 2) the responses obtained for the lexical item “rat” were mapped in five indigenous communities (1. Tembé, 3. Asuriní, 4. Guaraní-Mbyá, 5. Suruí Aikewára and 6. Guajajára). The data in this dialling map confirm the classification of Rodrigues and Cabral (2012) of the relations of parenting of these languages within the Tupí-Guaraní family. The Asuriní languages of Tocantins (point 3) and Suruí Aikewára (point 5) are very close and share the same lexical item [anu'sa] to define “rat”. As Laraia (1967) affirms, these are two ethnic groups who, in the past, constituted the same cultural system. The same process is observed between Tembé (point 1) and Guajajára (point 6), where the lexical item [anu'za] and its phonetic variants [anu'za], [nu'za] and [anu'zah] are very similar, keeping the pronunciation of the voiced alveolar fricative [z], in contrast with the Suruí and Asuriní that present the voiceless fricative realization [s]. For Rodrigues and Cabral (2012), Tembê and Guajajára constitute two branches (the western and the eastern, respectively) of the Tenetehára language. Among the Guaraní-Mbyá, we obtained the item [ãgu'dʒa]. The data in map D01 demonstrate a trend among these indigenous languages for the variation between these sounds [s, z, dʒ], a variation that is also present in the Portuguese spoken by these communities, especially in the speech of the older collaborators, as traces of influences from the linguistic substrate of the Tupí-Guaraní matrix (GUEDES, 2017).
These methodological adaptations provided an efficient geolinguistic mapping of Portuguese in contact with these five indigenous languages. This developed model has served as a reference for the other studies that are being carried out under the ALiPAI project in other indigenous lands of the states of Amapá and Amazonas.

ALiPAI data collection is carried out through the application of three questionnaires:

i. Sociolinguistic Questionnaire (QS) that aims at identifying the sociolinguistic situation of the community in relation to the degree of bilingualism and the linguistic behavior of the speakers in relation to the languages spoken in the community (Portuguese and indigenous languages); ii. Phonetic-Phonological Questionnaire (QFF), which is mainly oriented, but not exclusively, to identifying areas where phonic facts were already documented for Portuguese in previous research; iii. The complementary Phonetic-Phonological Questionnaire (QFFC), based on phenomena of phonetic-phonological influence, described by Silva (2010) to record the possible influences of the Tupi-Guarani origin substrate on the Portuguese spoken by the indigenous of ethnic groups in question.

The Sociolinguistic Questionnaire was based on works of other researchers in Pluridimensional and Contactual Geolinguistics such as Margotti’s thesis (2004), and questions elaborated within the GeoLinTerm project team.

The QFF and QSL questionnaires applied in the ALiPAI field research were elaborated by the team from the Linguistic Atlas of Brazil (ALiB), but they were adapted as follows: after each question the collaborator was asked “and in your language, How is this called? “, to record the knowledge about indigenous languages by the collaborator. At the end of the QSL we asked adult collaborators to tell a story about a Personal Experience, it could be the narration of a fact or a traditional story of that ethnic group, in Portuguese and later in the indigenous language of the group under investigation.

In order to investigate more extensively the lexical and phonetic variation of the Portuguese spoken by the collaborators, the “triple step” method, or “three-time research”, was adopted from Radtke and Thun (1996) questionnaires: ask, insist and suggest.

To test the questionnaires, we decided to use illustrated QFF and QSL questionnaires, especially with children and the elderly, which, because of methodological limitations, did not contemplate some items with an abstract content or that present processes: “Thank you” (QFF - 79), or “fanhoso” (QSL - 101).

The cartography is being done through Photoshop and CorelDRAW image editing software. The base map was generated from the ArcGIS georeferencing program.

Data is being collected in-locus, through interviews recorded with professional digital recorders Sony, Tascam and Zoom brands, in order to register high quality sound files to compose the ALiPAI project database.
FINAL CONSIDERATIONS

The application of the questionnaires in this first experience of field research in indigenous areas can be evaluated as satisfactory despite the physical and acoustic conditions of the contexts, the recording situations and the index of non-obtained answers for some questions, such as those of the semantic field “Urban Life” of the ALiB QSL, which seems to reflect some lexical items that do not belong to the semantic universe of the communities studied so far.

From the methodological adaptations made to the data collection in the ALiPAI, our first impressions suggest that they were very productive. Especially the inclusion of the diastratic variable (schooling) (adapted from the ALiB project), considering those with the first year of high school to higher education in order to consider this variable that is in process of expansion in indigenous communities that are having access to secondary and higher education.

On the other hand, we consider the inclusion of the two children (1st age group) as collaborators in the research to be of paramount importance. This choice made it possible to ratify other studies that have already indicated the low vitality of the indigenous languages of these ethnic groups, which are classified as “threatened of extinction”. It’s a fact that is directly related to the diffusion of the Portuguese language among the younger generations (1st and 2nd age groups) in the indigenous studied areas.

REFERÊNCIAS


