The spread of cultural vocabulary in Rondônia

A study of borrowability in the semantic fields of religion and agriculture

Sandra Cronhamn
ABSTRACT

This thesis is an investigation of the borrowability of cultural vocabulary among languages of different genetic origin in the Brazilian state of Rondônia, one of the most linguistically diverse regions in the world. Since cultural vocabulary is more likely to be borrowed compared to other lexical domains, if there has been language contact in this area, this is expected to be reflected in these terms. Two semantic fields of study were chosen: religion and agriculture.

The study concerns 23 languages of various genetic origin: 11 Tupían, 6 Chapacuran (dialects), 2 Nambiquaran, 1 Macro-Gean, 1 Panoan and 2 isolates. The two semantic fields turned out to show very different results. The religion terms showed almost no borrowability at all, which contradicted previous findings strongly. For the agriculture terms, the results were highly individual and cannot easily be generalized, but were given a more extensive description in the Results and discussion chapter. The maize terms have been attributed an extensive section since they showed the most complex results and were connectable to earlier research.

KEY WORDS

Amazonia, Rondônia, borrowability, loanwords, areal diffusion, language contact, cultural vocabulary

ACKNOWLEDGEMENTS

I would like to express my thankfulness to Love Eriksen for providing a non-linguistic perspective and for making all the lovely maps; to Vera da Silva Sinha and Wany Sampaio for giving me access to their material; and to Vera and Chris Sinha for taking part in my progression and for valuable comments. Last but not least, I want to thank my supervisor Gerd Carling for introducing me to the subject, guiding me through the jungle of cognate identification and motivating me with her enthusiasm for my research.
# Table of Contents

1 INTRODUCTION .............................................................................................................. 5  
2 BACKGROUND ............................................................................................................... 7   
  2.1 The Amazonian Landscape ......................................................................................... 7   
  2.2 Rondônia .................................................................................................................. 10   
    2.2.1 Rondônia – a linguistic area? .............................................................................. 11   
  2.3 Presentation and Classification of Languages .......................................................... 12   
    2.3.1 Tupian languages ............................................................................................... 13   
    2.3.2 Panoan languages ............................................................................................... 14   
    2.3.3 Nambiquaran languages .................................................................................... 15   
    2.3.4 Chapacuran languages ...................................................................................... 15   
    2.3.5 Macro-Gean languages ...................................................................................... 15   
    2.3.6 Linguistic isolates: Aikanã and Kanoé ................................................................. 16   
    2.3.7 Summary ............................................................................................................. 16   
3 THEORETICAL BACKGROUND .................................................................................... 18   
  3.1 Genetic relationship: The Comparative Method ......................................................... 18   
  3.2 Areal Linguistics and Linguistic Areas .................................................................... 19   
  3.3 Borrowing ................................................................................................................. 20   
4 METHOD ....................................................................................................................... 23   
  4.1 Material ..................................................................................................................... 23   
  4.2 Method ...................................................................................................................... 24   
  4.3 Problems .................................................................................................................. 25   
5 RESULTS AND DISCUSSION ...................................................................................... 27   
  5.1 Religion Terms ......................................................................................................... 27   
  5.2 Agriculture Terms ................................................................................................... 28   
    5.2.1 To cultivate; to sow; seed; harvest, crop ............................................................ 29   
    5.2.2 Tobacco; to smoke; smoking pipe ..................................................................... 29   
    5.2.3 Sweet potato ..................................................................................................... 29   
    5.2.4 Manioc, cassava ............................................................................................... 30   
    5.2.5 Calabash ........................................................................................................... 30   
    5.2.6 Pumpkin ............................................................................................................. 30   
    5.2.7 Roots ................................................................................................................ 30   
    5.2.8 Maize, corn ....................................................................................................... 31   
    5.2.9 Presentation of the maize roots ......................................................................... 35   
6 CONCLUSION .............................................................................................................. 43   
7 REFERENCES ............................................................................................................... 46   
8 ATTACHMENT ............................................................................................................. 51
ABBREVIATIONS

To facilitate the understanding of the genetic origin of the languages, the language names are followed by letters indicating the language family (and branch if there is any), wherever this is not made clear from the immediately surrounding context. The key to the abbreviations is given below.

C    Chapacura
I    Isolate
M    Macro-Ge
N    Nambiquara
P    Pano
TA   Tupí-Arikém
TG   Tupí-Guaraní
TM   Tupí-Mondé
TP   Tupí-Puruborá
TR   Tupí-Ramaráma
TT   Tupí-Tuparí
1 INTRODUCTION

The Amazonian region is one of the least explored areas in the world when it comes to linguistics, and due to the endangerment of the majority of these languages, the task of investigating them is urgent.

Studying languages is not just interesting from a purely linguistic point of view; linguistic evidence also constitutes an important complement to findings from other disciplines such as history, archaeology, anthropology and human ecology. Combining these findings makes up a more complex and more realistic interpretation of the object of research.

The choice of cultural vocabulary as the research object is based on the hypothesis that this part of the lexicon is more likely to be borrowed than e.g. the basic vocabulary of Swadesh lists. The Swadesh lists are constructed to consist of concepts existing in more or less every culture, thus being less subject to borrowing. Spreading of cultural phenomena, however, usually adds something to a culture that did not exist before and thus did not have a name, which is why the term from the donor language of the cultural phenomenon is more likely to be borrowed along with the new concept.

Presupposing that the higher borrowability of culture terms could help to identify language contact, the aim of this study is to make a contribution to the mapping of earlier connections between different ethnolinguistic groups in Rondônia by examining a domain of the vocabulary which – if any – is likely to show traces of the contact situation in its loanwords.

The subdomains of agriculture and religion were chosen because they have an important place in the culture of both diachronic and synchronic Amazonia. As will be seen in the analysis, the two domains showed a very unequal amount of borrowed items. The results of the religion terms strongly contradicted previous research. The agriculture terms showed a complex pattern on which it is hard to make any generalizations. This will be further described in the chapters 5 and 6.

This thesis does not have a single, explicit research question; rather, it is an explorative study built on the hypothesis that cultural vocabulary, along with newly introduced cultural concepts, is more likely to be borrowed than basic vocabulary. If we can trace
loanwords, they can tell us not only about linguistic exchange, but thereby also about the history of humankind.

The disposition of the thesis is as follows. Chapter 2 introduces the background of the subject. First, the Amazonian region as a whole is presented, before the focus is narrowed down to Rondônia. Then, the languages investigated are presented, grouped by their genetic origin.

Chapter 3 provides the necessary theoretical background, starting with genetic relationship and the comparative method. After that, areal linguistics is gone though, followed by a section about borrowing. In chapter 4 the material and method for this thesis is described. Some problems which have occurred during the analysis are also brought up.

In chapter 5 the results of the research are presented and discussed. The chapter is divided into two basic sections, one for each semantic field investigated, and the section for the agriculture terms is further divided into a section each for every term (or group of terms). Since I regarded the maize terms the most fruit-bearing, these were assigned an extensive section in the end of the chapter. Finally, in chapter 6 the results of the thesis are summarized.

The attachment comprises two language lists: one with the languages in the material, and one with all the other languages mentioned in the thesis. The alternative names of the languages and their Ethnologue codes are given in the lists.

All the maps in the thesis have been constructed by Love Eriksen, PhD in Human Ecology at Lund University, and are based, in addition to the material of this thesis, on the material in his dissertation (Eriksen 2011: material described in 12-15).
2 BACKGROUND

This section provides an overview of various aspects of the subject of the thesis. It begins with describing Amazonia as a region, both linguistically and non-linguistically. Thereafter, the focus is then narrowed down to Rondônia. Finally, the languages in the material are presented family by family.

2.1 THE AMAZONIAN LANDSCAPE

The Amazonian linguistic landscape could be described as a mosaic, with large, more or less discontinuous language families stretching over vast areas and small families and linguistic isolates scattered in between. It is evident that several factors are important contributors to this complex distribution pattern. Figure 1 shows the distribution of the largest Amazonian language families.

Figure 1. The distribution of the largest language families in Amazonia. (Map compiled by Love Eriksen.)
Two of the largest and geographically most widespread language families of Amazonia are those of Arawak\(^1\) and Tupí, the expansions of which have been historically important for the entire Amazonian region. Hornborg and Hill (2011) describe earlier views on the matter as regarding this the result of demic migrations, where ethnolinguistic groups expanded into new locations along with their language and culture, pushing other groups aside. Their own hypothesis serves a more complex explanation of the situation. Recognizing that ethnicity and language are separate (although closely linked together), the expansion of a language does not necessarily imply the migration of the specific individuals speaking that language. They were “not expanding into empty space” (Hornborg & Hill 2011:6), nor is any displacement of other peoples necessarily implied; it is important to realize the part played by the surrounding peoples and to take into consideration sociocultural factors such as language shift, trade-induced language contact, multilingualism, intermarriage, political hierarchies, military conquest, captive taking, ethnogenesis etc. Thus, there are several ways for a language to expand beyond the limits of the ethnolinguistic group, and the linguistic diversity of Amazonia is a complicated matter which requires a nuanced explanation. Hierarchical, asymmetric relations appear to have been common, and changing over time.

By the end of the first millennium AD, the Arawaks dominated large areas of Amazonia with their trade networks (Hornborg and Hill 2011:6). A possible proto-home for the Arawakan family has been suggested somewhere in Western Amazonia (Walker & Ribeiro 2011). According to the article by Danielsen, Dunn and Muysken (2011:186) in which a comparative analysis of structural features in Arawakan languages has been made, “there is evidence for multiple southerly migrations, in that the southern groups do not cluster together as a whole”. This notion will be brought up again in chapter 6.

The Arawakan expansion did not only involve spread of the Arawakan languages, but of a cultural complex referred to as the Arawak matrix by Eriksen (2011). He lists nine features characteristic of Arawakan groups (Eriksen 2011:8-9; based on Schmidt 1917; Izikowitz 1935; Santos-Granero 1998 and 2002; Heckenberger 2002; Hornborg 2005; Hill 2011), for instance suppression of endo-warfare, descent as the main basis of social

\(^{1}\) Although there are no Arawakan languages in the data for this essay, they deserve a brief description due to their huge, long-lasting impact on the whole Amazonian region.
and political life, an elaborate set of ritual ceremonies including sacred instruments, high-intensity landscape management, and settlements, trade and transportation along rivers.

The great Tupian expansion occurred later, after AD 1000 (Hornborg and Hill 2011:3), and was still in full action at the time of the European arrival. They seem to have employed the pre-existing Arawakan routes. The Tupian family is believed to have its origin in Rondônia where five of its ten branches (the whole Western subgroup), are spoken exclusively, while the Eastern subgroup began spreading eastwards through the southern parts of Amazonia. The Tupí-Guarani branch of the Eastern subgroup is by far the most extended one, spread out over vast, discontinuous geographic areas (some Tupí-Guarani languages are also spoken in Rondônia). The Tupían languages will be further described in section 2.3.1.

Contrary to the Arawaks, whose primary spreading mechanisms were “trade, kinship, prestige and ceremonialism”, Tupíans employed more violent methods in their conquests such as warfare and even cannibalism, although “the attraction of Tupían identity, language, cosmology, and material culture was still an important factor behind the dispersal of the languages” (Eriksen 2011:245).

As pointed out by Hornborg and Hill (2011:5), a common misconception of the Amazonian tribes is that they are small, simple, isolated and limited by their natural surroundings, in contrast to e.g. the “sophisticated” Andean cultures. But as a matter of fact, it has more recently been suggested that Amazonia is the origin of many of the goods traditionally associated with e.g. the Incas. For at least 2000 years before the European arrival, western Amazonians inhabited “sedentary, densely populated settlements that were economically and culturally connected to societies of the Andean slopes and highlands” (Hornborg & Hill, 2011:14). As for the isolation, it is to a large extent a result of the European invasion. Not only was violence evoked in some situations of contact, but the Europeans carried bacteria to which the native Americans had no resistance, leading to a drastic decrease of the population. This in turn made the survivors flee into the forest out of fear for their lives, leaving their settlements, in some cases for a more primitive existence (Hornborg & Hill 2011:5).
2.2 RONDÔNIA

The languages investigated in this thesis have the common denominator of being spoken in the state of Rondônia in Brazil. Rondônia is situated in southern Amazonia, bordering to the other Brazilian states of Mato Grosso and Amazonas in the east and north respectively, and to Bolivia in the south-west. The Bolivian border is constituted by the Guaporé river. The Madeira river flows through the north-western parts of Rondônia, and its tributaries drain large parts of the region.

Rondônia is adjacent to, and partly overlapping with, the Brazilian highlands in the south-east, accounting for some considerable alternation in height in comparison to the more inner parts of Amazonia. Tropical rainforest used to cover the majority of the region, but it “is now deforested for about 50%” (Crevels and van der Voort 2008:155). The native population has suffered hard from this, and today most of them live in indigenous reserves which, although being under external political control, allow them to maintain their lifestyle.

Rondônia is an area of considerable linguistic diversity, with many small language families and several linguistic isolates. A part of these languages are poorly documented and the ways and directions of language contact are yet to be established. The languages in the material of this analysis2 and their present-day locations are presented in Figure 2. Many of the groups were not contacted until fairly recently, and there might still be uncontacted groups in Rondônia. “The right side of the Guaporé, i.e. the Brazilian side, was confronted with Western culture really only in the beginning of the twentieth century, with the rubber boom” (Crevels and van der Voort 2008:155). Although the population has decreased, the Rondônian linguistic diversity is relatively intact and few languages are completely extinct. But as many of the languages are severely endangered, documenting them before they pass into oblivion is still a very urgent task.

2 Four of the languages on the map have not been analysed: Wayoró, Oro At, Cojubim and Sabanês. These languages contained no data for the given fields of analysis for this thesis. They were sorted out as the research object was established, but by then the map had already been made.

3 It is, of course, not only the current locations of the languages that matter when it comes to investigating the contact between them; it is clear that these have varied in the past. A much more unclear matter is how and when this has happened.
2.2.1 RONDONIA – A LINGUISTIC AREA?

The Guaporé-Mamoré valley has fairly recently been suggested as a potential linguistic area, comprising Rondonia as one of its subareas (Crevels and van der Voort 2008). Amongst the putative areal traits we find both grammatical, phonological and lexical features. Much more research is however needed before any definite establishments can be made. The concept of linguistic areas will be reintroduced and further described in section 3.2.

It is common for the groups in a linguistic area to share a number of cultural traits too. Crevels and van der Voort (2008:152) mention the Marico cultural complex⁴ as defined by Maldi (1991). This cultural complex comprises speakers of Tupían, Jabutían and

---

⁴ There are other examples of cultural complexes in Rondonia, including Tupí, Chapacura and Guaporé (Crevels and van der Voort 2008:152-155; based on Lévi-Strauss 1948; Galvão 1960).
isolated languages in Rondônia. The area is characterized by a combination of the following traits:

- seminomadic swidden agriculture combined with hunting and gathering
- relatively small egalitarian societies
- territorial subgroups that often bear animal names
- territorial subgroups could form alliances with others across linguistic borders
- religion involves shamanism and hallucinogenic substances
- material culture is characterized among others by the marico, a crochet carrying net made of the fibres of specific palmtree leaves
- the local fermented alcoholic brew called chicha, which is a drink mainly based on maize, yam, manioc or fruits such as banana, is mashed, fermented and sifted in a specific way

2.3 Presentation and classification of languages

The material of this study consists of 23 languages altogether; 11 Tupían, 6 Chapacuran (dialects), 2 Nambiquaran, 1 Panoan, 1 Macro-Gean and 2 isolates. Although more indigenous languages are spoken in Rondônia, this troop is fairly representative for the linguistic diversity in the region.

When it comes to classification of the Amazonian languages, there is a great amount of disagreement between scholars. Many languages are poorly investigated and therefore there is much left to do in this field. Here I will follow the classifications that Campbell (2012:59-166) presents. On the other hand, the naming of Ethnologue will be used to the highest possible extent, since the code system will allow for easier identification of the languages. For the Oro dialects (Chapacuran), which are not mentioned either in Campbell or Ethnologue.com, the naming of the data collectors Wany Sampaio and Vera da Silva Sinha will be followed. If it its not clear from the context, the languages will be marked with letters indicating their genetic origin (the key to the abbreviations can be found on p. 4). A list of all the languages in the material, their alternative names and spellings and their Ethnologue codes can be found in the attachment. A similar list is also provided for the languages which are mentioned in the thesis but are not part of the material.
2.3.1  **Tupían languages**

The Tupían language family is one of the largest in South America. It consists of ten branches comprising nearly 70 languages. The Tupí-Guaraní branch has spread not only over large areas of the Amazon basin, but also along the Paraguay river and the Atlantic coast. When the Europeans first arrived at the Atlantic coast of South America in the early 16th century, the first tribes they met were coastal Tupían peoples (Rodrigues & Cabral 2012:495). These languages have been known to Westerners for a long time and are among the most well-documented of the South American language families.

Following the classification by Rodrigues and Cabral (2012:496-499), the Tupían family initially divides in two: an eastern and a western subgroup. The subgroups contain five branches each, and the Tupí-Guaraní branch is divided into eight sub-branches. A simplified Tupían family tree is provided below, with the languages in the material given in italics to the right of their respective branches.

**Western subgroup**

- Puruborá
- Ramarâma
- Mondé
- Tuparí
- Arikém

**Eastern subgroup**

- Jurúna
- Mundurukú
- Mawé
- Aweití
- Tupí-Guaraní

**Figure 3** shows the distribution of the Tupían branches in Amazonia. The red line indicates the border of Rondônia. As can be seen, all the five western branches are concentrated to Rondônia, as well as some Tupí-Guaraní languages, which has lead to a theory about Rondônia being the original homeland of Proto-Tupí. Archaeological evidence such as pottery characteristic of Tupiâns has also been found in the area, dating Proto-Tupí 5000 years back in time (Rodrigues & Cabral 2012:499-500). The Tupían language family thus plays an important part in Rondônia, and has been doing so for a long time.
2.3.2 PANOAN LANGUAGES

According to Loos (1999:227), the close similarities and relative mutual intelligibility of the Panoan languages indicate a relatively recent split and expansion. Most languages in this family are spoken in Peru, with an additional few in adjacent regions of Bolivia and Brazil. Kaxarari is the only Panoan language in the material of this study. It is located in the westernmost corner of Rondônia, quite far from the other languages in the material and separated by the Madeira river.

Eriksen (2011:36) describes a fundamental socio-cultural divide between two Panoan groups: one group of large-scale societies situated along the lower part of the Madre de Dios River practising riverside agriculture, and one group of small inland societies of hunter-gatherers, inhabiting areas east of the Ucayali River. The latter group is poorly known and some societies are perhaps still uncontacted, whereas societies the former
group have coexisted in mission settlements during the colonization. According to its location, Kaxararí should belong to the river-based group.

Although Loos (1999:227) wrote that more work needs to be done in this field, a genetic unity comprising both the Panoan and the Tacanan family (spoken between the Madre de Dios and Beni rivers in Bolivia and Peru) has been proposed by several scholars and appears to be generally accepted today (e.g. Campbell 2012:100-102).

2.3.3 NAMBIQUARAN LANGUAGES
The closely related Nambiquaran languages, many of which are mutually intelligible, constitute a small family located primarily in the Brazilian state of Mato Grosso but also in adjacent regions in Eastern Rondônia. Lowe (1999:270-271) writes that “until about 20 years ago, the Nambiquara lived a traditional hunting and gathering life, supplemented by some slash-and-burn agriculture.” They had little contact with the Western culture and the knowledge of Portuguese was very limited. Nowadays, however, most of the young speak good Portuguese, but the vernacular is still used by all Nambiquara and all children learn it.

2.3.4 CHAPACURAN LANGUAGES
This small family has its centre in the Guaporé valley and along the tributaries of the Madeira river, from the Brazilian state of Amazonas through western Rondônia down to adjacent regions in Bolivia. Few extensive descriptions are available of these languages (an exception being a grammar of Wari’ by Everett and Kern from 1997), but the genetic relationship of the family has been known since d’Orbigny (1839), as put forward by Aikhenvald and Dixon (1999:359). Judging from the striking similarities which become obvious when comparing the terms, the Chapacuran “languages” in this material should more adequately be referred to as dialects.

2.3.5 MACRO-GEAN LANGUAGES
The Macro-Gean language family is a group of branches whereof many different constellations (and names) have been proposed since at least the mid 1920’s. The history of this family is wrapped in mystery since many of the languages, and in some cases whole branches, are already extinct and very poorly documented. The genetic unity of these languages is not uncontroversial; Campbell (2012:93) calls the Macro-Gean family a “hypothesis”.

15
As can be seen in figure 1, the geographical spread of this language family is vast. Its primary area of distribution is in the eastern parts of Brazil, stretching all over the country in latitude. A few languages are also spoken in Bolivia.

The Jabutian languages (Jabuti, Arikapú: both spoken in Rondônia) have been argued to belong to this family by some scholars (Nimuendajú 2000 [1935]; Ribeiro and Van der Voort 2010). The classification is justified by Ethnologue (online source 1), but not by WALS (online source 2). In this thesis, Jabutí will be considered a Macro-Gean branch.

2.3.6 Linguistic isolates: Aikanä and Kanoé

There are a high number of unclassified languages in Rondônia, many of which are poorly documented. Two isolates are represented in the material for this thesis: Aikanä and Kanoé. As indicated by figure 2, they are spoken very near each other. The Aikanä are reported to live together with speakers of another linguistic isolate called Kwaza (Aikhenvald & Dixon 1999:363) and in the same indigenous reserve as LatundêN (online source 3). Kanoé is spoken today only by four or five elderly people whose first language is Portuguese, and by three monolingual speakers first contacted in 1995 (Bacelar, 2005).

2.3.7 Summary

To conclude, we expect Tupían languages to have been spoken in Rondônia for at least 5000 years. For the other, less researched (and in some cases quite recently contacted) languages, the history and earlier locations is harder to establish. Crevels and van der Voort (2008:172) write that the diversity of languages in Rondônia and the whole Guaporé-Mamoré region must have gathered up and increased during several centuries, without mentioning any particular linguistic groups taking place in it. The region is described as a residual zone by Crevels and van der Voort (2008), using the terminology and definition of Nichols (1992). Residual zones are “regions in which linguistic diversity has been accumulating as languages were pushed aside by those of politically and economically powerful populations in adjacent linguistically homogenous ‘spread zones’” (Crevels & van der Voort 2008:167) Some characteristics of a residual zone are for instance high genetic density, high structural diversity and local multilingualism rather than an areal lingua franca (Nichols 1992:21). Rondônia seems to have functioned as an area of refuge from the seventeenth century and onwards: “Populations from the lower Madeira region through time migrated south and drove the existing
populations even further south, and formed part of a push-chain mechanism that may have been created ultimately by the encroachment of western civilization from the northeast, channelled though the Amazon river and its tributaries” (Crevels and van der Voort 2008:172; based on Leonel 1995). Furthermore, refuge is not the only reason for the accumulation of this large number of ethnolinguistic groups – the Guaporé-Mamoré region is likely to have attracted populations because of its fertile lands and fishing grounds.

It is important to realize that even languages which are not currently spoken in Rondônia may have had a finger in the pie. If the Arawaks migrated southwards, as suggested by Danielsen et al. (2011), it is possible that they were at some stage spoken in the region. This theory will be revisited in chapter 6.
3 THEORETICAL BACKGROUND

3.1 GENETIC RELATIONSHIP: THE COMPARATIVE METHOD

When we talk about languages being *genetically related* to each other, we mean that they – analogously to biological genetic relationship – descend from a common ancestor, a *proto-language*. This view on language change likens language development to a family-tree, where the proto-language generates dialects which in time become distinct *daughter languages* (and *sister languages* to each other) (Campbell 2004:122-123).

The Comparative Method seeks to establish genetic relationship by reconstructing the proto-language based on comparison of as many as possible of the descendants. “The work of reconstruction usually begins with phonology, with an attempt to reconstruct the sound system; this leads in turn to reconstruction of the vocabulary and grammar of the proto-language. [...] If the reconstruction is successful, it shows that the assumption that the languages are related is warranted” (Campbell 2004:123).

This method is the most important tool for proving genetic relationship. By providing valid reconstruction and sound correspondences, one can eliminate other possible reasons for similarities between languages: “accident (chance), borrowing, onomatopoeia, sound symbolism, nursery forms, and universals and typologically commonplace traits” (Campbell 2008:165). A combination of considerations is important to pay attention to: both finding *cognates*, that is, similar and related words in sister languages, and establishing sound correspondences. As Campbell (2008:193-194) points out, applying this method properly is not a simple step-by-step procedure but rather about working in parallel tracks:

> It is not as easy as first observing some similar words, calling them etymologies, classifying the languages, and then belatedly seeking sound correspondences and attempting reconstruction. Rather, given the interrelatedness, both recognizing possible cognates and working out sound correspondences are necessary to confirm each other. To show words compared in different languages are true “etymologies” (cognates), it is necessary to show support of regular sound correspondences, since otherwise it is not possible to know whether the similarities exhibited might not be due to some other factor than inheritance. Similarly, true sound correspondences can be established only by comparing cognate forms. Since these depend on one another, it becomes necessary to work back and forth, checking correspondences within probable cognates and refining cognates against their fit with sound correspondences until in the end both are understood and mutually support one another, and exclude material that does not fit the system. This is not circular, but rather requires both to be worked out before final judgments are warranted.
Campbell (2008:195-196) further states the importance of keeping the semantic content compared as alike as possible. If one allows comparison of non-synonyms, the possibility of the similarity being due to chance grows. Another way of reducing the risk of chance is to compare as long segments as possible: “Monosyllabic CV or VC (or V) forms may be true cognates, but they are so short that their similarity to forms in other languages could also easily arise due to chance” (Campbell 2008:200).

3.2 Areal linguistics and linguistic areas

In order to be able to identify genetic relationship and to understand the full history of languages, it is important to consider areal traits and separate them from inherited features. Campbell (2004:330-331) gives the following definition of a linguistic area, or a Sprachbund: “The central feature of a linguistic area is the existence of structural similarities shared among languages of a geographical area (where usually some of the languages are genetically unrelated or at least are not all close relatives). It is assumed that the reason why the languages of the area share these traits is because at least some of them are borrowed”. Linguistic areas are usually associated with multilingualism.

Furthermore, Campbell (2004:338-339) lists three considerations and criteria for the determination of a linguistic area: the number of shared traits; the clustering of them in some specific way; and the weight of different areal traits. The weight of an areal trait refers to the fact that different features are differently well suited to count as evidence when establishing a linguistic area. The clustering of the traits concerns the fact that each feature typically has its individual spread, but that a core exists in every linguistic area where a number of traits are diffused. As for the number of traits, Campbell says it best himself: “the more, the merrier” (Campbell 2004:339).

A reliable way of determining whether or not a feature is a trait belonging to a linguistic area is to look at related languages spoken elsewhere. If these do not share the feature, the likeliness for it being an areal trait is higher. If we apply this to the Guaporé-Mamoré region, Aikhenvald (2012:85) points out that “The question of whether the Guaporé-Mamoré basin has ever been an established linguistic area is complicated by the fact that many groups – such as Nambiquara, numerous subgroups of Tupí, and Tacana – are not spoken anywhere outside the area”.

19
3.3 Borrowing

Linguistic borrowing is carried out by the following formula: a recipient language borrows a loanword created from a source word in a donor language (Campbell 2004:62). Borrowed words are often remodelled to fit the phonological & morphological structure of the recipient language. Importantly, non-lexical features can be borrowed too, such as e.g. phonological and typological traits. However, these are not relevant for the present thesis and will not be further described.

The amount of borrowed words varies greatly from language to language, but the phenomenon of borrowing seems to be universal. A legitimate question is: Why do languages borrow altogether? After all, any language contains the necessary material to creatively produce new words. Campbell (2004:64-65) divides the reasons for linguistic borrowing into three: 1) need (often along with new concepts), 2) prestige (foreign term highly esteemed), and 3) negative evaluation (foreign term derogatory). The two reasons mentioned first are the most important.

Haspelmath (2009:47) mentions widespread bi- or multilingualism as a possible (and probable) reason: “When many people know a concept by a certain word but not by another word, even if the better-known word belongs to another language, it becomes more efficient to use the better-known word”. He does however emphasise the distinction between loanwords and code-switching, a not so different phenomenon also requiring bi- or multilingualism, stating that loanwords indicate language change whereas code-switching represents a kind of speech behaviour, often highly individual. For something to count as a loanword, much further integration is needed.

Although it is hard to make any absolute predictions as to what can be borrowed or not, there are clear differences between different parts of the vocabulary. Greenberg (1957:39) claimed that “fundamental” (i.e. basic) vocabulary is less likely to be borrowed than cultural vocabulary. He asserted that “fundamental vocabulary is proof against mass borrowing”, although he admitted that any lexical item might be borrowed. The famous Swadesh list, named after its originator the American linguist Morris Swadesh, consists of basic vocabulary and is often used as strong evidence for genetic relationship. It is constructed to comprise terms which are basic to all human societies and are the least likely to be borrowed. Tadmor (2009:65) writes that “measuring the
borrowability of lexical meanings is not entirely straightforward”, and criticizes the Swadesh list for being based on intuition rather than empirical evidence (although he admits it is the intuition of a brilliant scholar). He later concludes that the application of borrowability “becomes much more meaningful when used in conjunction with other variables, such as universality, stability, and simplicity” (Tadmor 2009:74). Haspelmath and Tadmor (2009)’s Loanword project results in an alternative version of the Swadesh list called the Leipzig-Jakarta list.

Normally, for a loanword to be recognized, we need to establish both a source word and a donor language. To determine the direction of borrowing for a loanword and identify the donor language and the recipient language, Campbell (2004:69-74) mentions some clues to take into consideration:

1. **Phonological clues.** A loanword can contain sounds which do not normally exist in the native vocabulary, in which case it is easy to determine the direction. In case the phonological history of the language is known, previously undergone changes might also help to identify loanwords. These clues constitute the strongest evidence.

2. **Morphological complexity.** If a form is morphologically complex in one language, but monomorphemic in the other, the former is likely to be the donor language. For example, articles are sometimes incorporated into loanwords, losing their definite or indefinite connotation when merging together with the main word in the new language.

3. **Clues from cognates.** If cognates can be found among the relatives of one language but not the other, the language in whose family the word is represented in more than one language is the likely donor language.

4. **Geographical and ecological clues.** In case of e.g. cultural concepts or local plants and animals, the geographical regions and cultural traits associated with the languages may reveal a loanword’s origin.

However, in some cases all of this criteria is still not enough to determine the direction of borrowing. Words appearing in a number of different families in a certain area are often called *Wanderwörter* or *areal roots*. In other cases the specific donor language can be hard to isolate, in which case a donor family with similar cognates might be proposed (Haspelmath 2009:45).

When identifying cognates with the purpose of tracing genetic relationship, it is important to take into consideration the risk of *false cognates*. As Haspelmath (2009:36) writes, “Especially in less well-researched languages and language families, and at older stages of history, it is often unclear whether a word is a loanword or a native word that is cognate with its putative source.” For this thesis, however, the aim is not to establish
genetic relationship but to investigate possible areal diffusion, so “false cognates” are more or less what is being looked for.

Languages that on a superficial level seem similar in their grammar, phonology and lexicon might turn out to be genetically unrelated at a closer look, the similarities being due to borrowing. Establishing what is inherited and what is borrowed can be an enormously complicated task, especially when written sources - which for instance researchers of Indo-European languages have been spoiled with - are absent and data is sparse, as is the case with the Amazonian languages. Who knows what Amazonian languages have been in contact before the European arrival?
4 Method

4.1 Material

The data provided is in the form of word lists of about 400 words in 23 different languages. The languages are genetically divided into four language families and three isolates. For the analysis, the data has been narrowed down to two domains of culture terms: agriculture and religion. These have been chosen because of the fact that both agriculture and religion are traditionally important markers of ethnicity and identity in Amazonia. Each semantic field comprises 13 terms.

The terms that the informants were asked to translate were given in Portuguese, since the data was collected by Brazilian scholars and the informants were bilingual in (at least) Portuguese. Since translation usually entails a certain degree of semantic change, the gloss is presented in Portuguese here too (with an additional English translation). All terms of the two semantic domains that have been analysed are presented below.

<table>
<thead>
<tr>
<th>Religion terms</th>
<th>Agriculture terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>religião ‘religion’</td>
<td>cultivar ‘to cultivate’</td>
</tr>
<tr>
<td>deus ‘god’</td>
<td>semear ‘to sow’</td>
</tr>
<tr>
<td>templo, igreja ‘temple, church’</td>
<td>semente ‘seed’</td>
</tr>
<tr>
<td>sacrificio, ofrenda ‘sacrifice, offering’</td>
<td>milho ‘harvest, crop’</td>
</tr>
<tr>
<td>orar, rezar ‘to pray’</td>
<td>tabaco ‘tobacco’</td>
</tr>
<tr>
<td>sacerdote ‘priest’</td>
<td>fumar ‘to smoke’</td>
</tr>
<tr>
<td>pregar ‘to preach’</td>
<td>cachimbo ‘smoking pipe’</td>
</tr>
<tr>
<td>jejuar ‘to fast’</td>
<td>batata doce ‘sweet potato’</td>
</tr>
<tr>
<td>demônio ‘demon’</td>
<td>mandioca ‘manioc, cassava’</td>
</tr>
<tr>
<td>ídolo ‘idol’</td>
<td>cabaça ‘calabash’</td>
</tr>
<tr>
<td>magia, feiticeira ‘magic, witch’</td>
<td>abóbora ‘pumpkin’</td>
</tr>
<tr>
<td>feiticeiro, bruixo, bruxa ‘wizard, witch’</td>
<td>raízes ‘roots’</td>
</tr>
<tr>
<td>fantasma, alma penada ‘ghost,’</td>
<td></td>
</tr>
</tbody>
</table>

The collection of the data has been carried out by linguistics professor Wany Sampaio and anthropologist Vera da Silva Sinha as a part of Project Açaí, an action by the Brazilian government to promote, create and implement indigenous education in Brazil. The data was gathered in 2011 from students taking a two week course. They came from
different parts of Rondônia and were between 17 and 30 years old. The male/female distribution was equal.

The students were asked to fill out word lists with a total of 411 items, comprising e.g. Swadesh lists and cultural words. Those who spoke the same language collaborated in a classroom group discussion, and for a couple of languages whose condition is very critical, the students brought the word lists home to their villages and asked the elders to help them. If they did not know a word they were asked to leave a gap, but if they could think of a word with approximate meaning they were asked to submit it.

Therefore, whenever a gap occurs in the data, it should be interpreted as either non-existent in the vocabulary, or (sadly) as no longer part of the full collective knowledge in the language - it is impossible to tell which. The number of gaps varies from language to language, making the study somewhat uneven.

Since I have signed a confidentiality contract in order to have access to the material, only fractions of it will be shown, such as parts that are necessary to prove the points of the results.

4.2 Method

This study has been carried out in a comparative manner with the aim to conduct a sounding to find out what areas of the vocabulary could be further investigated, and to look for tendencies rather than to make definite conclusions.

In an initial step, the word lists have been transferred to Excel files with the languages on the x-axis (with a column each) and the terms on the y-axis (with a line each).

<table>
<thead>
<tr>
<th>English</th>
<th>Amundava</th>
<th>Kaxararí</th>
</tr>
</thead>
<tbody>
<tr>
<td>cultivate</td>
<td>materia awetym</td>
<td>banahi</td>
</tr>
<tr>
<td>to sow</td>
<td>amuyin ayWa</td>
<td>pyhahi</td>
</tr>
<tr>
<td>seed</td>
<td>ayña</td>
<td>banatu</td>
</tr>
<tr>
<td>harvest, crop</td>
<td>akambig</td>
<td>pyryhi</td>
</tr>
<tr>
<td>maize, corn</td>
<td>awatía</td>
<td>shyki</td>
</tr>
</tbody>
</table>

Figure 4. Example of initial compilation in Excel.
Thereafter, every term (i.e. every line on the y-axis) was analysed individually, with a methodology reminiscent of the Comparative Method described in section 3.1. - similar forms were being looked for, but since the aim was not to establish genetic relationship (as is usually the aim of this method) but to detect possible loanwords, no sound correspondences were accounted for and no reconstructions were made. The number of lines was extended and all supposed “cognates” were placed on the same line. When a new cognate occurred, it was moved down a line. The lines were named alphabetically.

<table>
<thead>
<tr>
<th>English</th>
<th>Amundava</th>
<th>Aikanā</th>
</tr>
</thead>
<tbody>
<tr>
<td>maize, corn</td>
<td>awatīa</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>haki</td>
</tr>
</tbody>
</table>

Figure 5. Example of cognate analysis in Excel.

The positive side of this method was that the results were quickly exposed and neatly presented. The negative side was that the file became very extended horizontally and thus hard to survey, and the limitation of only being able to find cognates with the exact same translation, i.e. on the same horizontal line. When correspondences started to appear across word borders, the words were written down on small pieces of paper and made into a jigsaw puzzle. This method made it possible to recognize similar terms with non-identical yet reminiscent semantics, so that e.g. two similar terms in two unrelated languages with the respective meanings “tobacco” and “to smoke” could be coupled.

According to Campbell, the method of analysing non-synonyms as if they were cognates is problematic, as described in section 3.1. However, as I wrote in the beginning of this chapter, the aim was not to draw any definite conclusion but rather to “see how the land lies” and scratch the surface in order to detect possible fields for future research.

4.3 Problems

When conducting research in the field of borrowed cultural vocabulary, some of the complications are obvious. For instance, there would be little need for such research if the circumstances of the cultural relations were already known, so lack of earlier investigation and sources is such an obvious difficulty. It is possible that I have made
erroneous conclusions because of the lack of literature. In the cases where I have expressed guesses or speculations, however, I have stated this very clearly.

Some of the problems I have encountered that have to do with Amazonia as a whole are the lack of written historical sources, the lack of information, description and data for many of the languages and the disagreement between scholars of language classification and naming. The same ethnic group can go by several names, making it frustratingly easy to miss important sources. Moreover, the same name may be used for different ethnic groups, making it frustratingly hard to get hold of the right information. The problem of finding sources accounts for the high number of second hand references.

Another complicating fact is the (still to a large extent) uncharted dispersion of ethnolinguistic groups, which has been going on for thousands of years. Where do they come from originally? Where have they been on their way to their present location? Who has they been in contact with whom, when, and how?

Early linguistic fieldwork in Amazonia was predominantly carried out by missionaries, who often did not have proper education. Additionally, Dixon and Aikhenvald (1999:3) write that “the standard of scholarship in South American linguistics is not high”, particularly addressing data from before ca. 1950. In other words, there is not only a lack of sources, but especially of reliable sources. A large number of the available sources are also in Portuguese, a language I do not understand. It is of course possible to look up certain words or sentences, but hardly to translate whole articles or books.

Another possible impairment of the analysis is my own modest background in comparative linguistics. This thesis is my first actual work on the subject. Therefore, the possibility that my untrained eye might have missed some important information is not insignificant.
5 RESULTS AND DISCUSSION

As I did not pose any clearly defined research question(s), it has been unclear all the way what kind of results to expect and how to present them. In this chapter, I will basically go through everything that I found that seemed interesting and meaningful for one reason or another. I realize that it sounds fuzzy, so to bring some order into the chapter, I will begin by distinguishing the two semantic fields and presenting them separately; first religion, then agriculture. The section of religion only includes a few comments. In the section of agriculture, the terms are presented one by one (or in clusters). This agriculture section is considerably more extensive and developed since these terms showed more interesting results, and, moreover, were linkable to previous research – especially the maize terms which deserved an extensive report.

As this has been an explorative study where it has been unclear what results to expect, the terms that did not show any interpretable results were sorted out and were not presented, to make room for the terms I was able to say something about. Note, however, that non-interpretable results do not necessarily mean domains which showed low borrowing, but rather results that could not be interpreted in relation to the background literature available.

5.1 RELIGION TERMS

Among the religion terms, there seems to be very little cross-genetic borrowing. This “zero result” does however suggest something significant: that the religion, or perhaps mythology, in these cultures is something that originates in the proto-cultures and has not been subject to cultural spread. It shall be pointed out that this is nothing but a guess, and that no sources have been found concerning the religions that these specific peoples profess.

It shall be noted that the number of gaps in the field of religion (67%) was twice as high as in that of agriculture (33%). For some languages no terms were submitted at all, for some there were only one or two, and for some almost full lists were provided. No Portuguese loanwords could be identified, but if there would be any, it is possible that the informants (who were all bilingual in Portuguese) left these out because they thought they were unnecessary, but this is only a speculation.
A few examples of possible borrowing can be mentioned:

1. The Kaxararí\textsuperscript{P} term tsura ‘god’ is possibly, but not doubtlessly, related to the Mamaindé\textsuperscript{N} term sūnā ‘god’. These languages are unrelated and spoken over 760 km apart, in different extremities of the state, with a great amount of languages spoken in between which do not share the term. The similarity might be due to sheer coincidence, but as will be described later, Kaxararí shares more terms with other languages in south eastern Rondônia, close to where Mamaindé is spoken.

2. The Chapacuran dialects have a common word for ‘god’ with a slight variation: \textit{iri yam’, iri jan, hriyan, iri yan}. The Aikanã equivalent bears some resemblance to these: \textit{iwerūa}. Aikanã is a linguistic isolate spoken more than 400 km away from the closest Chapacuran dialect. For this term too, there are many languages spoken in between which do not share the term.

3. A cluster of terms with similar forms and connotations (although not identical translations) is presented in the list below. All of the languages are Tupían except Kaxararí, which is Panoan and spoken quite far from the others.

<table>
<thead>
<tr>
<th>Term</th>
<th>Meaning</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{paaxú}</td>
<td>‘ghost’</td>
<td>Cinta Larga</td>
</tr>
<tr>
<td>\textit{páwu}</td>
<td>‘demon’</td>
<td>Cinta Larga</td>
</tr>
<tr>
<td>\textit{paixo matesod}</td>
<td>‘demon’</td>
<td>Suruí</td>
</tr>
<tr>
<td>\textit{apaikapan}</td>
<td>‘demon’</td>
<td>Amundava</td>
</tr>
<tr>
<td>\textit{ipaji}</td>
<td>‘magic, witch’</td>
<td>Amundava</td>
</tr>
<tr>
<td>\textit{paki}</td>
<td>‘ghost’</td>
<td>Kaxararí</td>
</tr>
</tbody>
</table>

5.2 Agriculture Terms

The agriculture terms showed an internal variation which makes it very hard to make any generalizations about them. The analysis of some of the terms has led to very interesting results. The terms are to be presented individually or in groups with related semantic content. The space of every section has been adapted according to how interpretable results it has shown, accounting for the difference in size. The last and most exhaustive section will be devoted to the maize terms, since these have provided the most interesting and interpretable results, which, moreover, were connectable to earlier studies.
5.2.1 To cultivate; to sow; seed; harvest, crop

These terms did not show any clear cross-genetic connections. Since no reconstructed proto-forms have been found, it is also hard to establish possible inherited terms.

5.2.2 Tobacco; to smoke; smoking pipe

The Proto-Tupían root *pe ‘tobacco’, as reconstructed by Rodrigues and Cabral (2012:563), is found only in PuruboráTT (pete) and possibly as a borrowed item in JabutiTM (padi). The Tupí-Mondé languages employ completely unrelated terms for ‘tobacco’, perhaps borrowed from somewhere else: maxukuôn (Cinta Larga), maxo (Suruí; Gavião do Jiparaná).

The Proto-Tupí-Guaranían term for tobacco is *petým (Jensen 1999:129), but the Amundava’sTG term for both ‘tobacco’ and ‘to smoke’, aijurua tatatigawa, is possibly derived from a’yî ‘to burn’ and tata ‘fire’. Kanoél has a reminiscent term a’y ‘tobacco’, but due to its humble length, the possibility of sheer coincidence cannot be ruled out. (See section 3.1. for more information about this.) For Uru-Eu-Wau-WauTG no term is listed at all.

The terms for ‘to smoke’ are similar in the following three unrelated languages: LatundêN kuha, Kaxararlí kuhahi and Aikanãl kapuai. According to Ethnologue, Latundê and Aikanã are spoken in the same indigenous reserve (online source 3), whereas Kaxararí is spoken at the opposite end of Rondônia.

5.2.3 Sweet potato

Rodrigues and Cabral (2012:563) have reconstructed the Proto-Tupían root *wetîk, which can clearly be identified in the following Tupían languages:

<table>
<thead>
<tr>
<th>Tupí-Guaraní:</th>
<th>Amundava</th>
<th>ytyka enenaim</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Uru-Eu-Wau-Wau</td>
<td>yty’ka</td>
</tr>
<tr>
<td>Mondé:</td>
<td>Cinta Larga</td>
<td>wetingaa xiín</td>
</tr>
<tr>
<td></td>
<td>Suruí</td>
<td>watinga kakur</td>
</tr>
<tr>
<td></td>
<td>Gavião do Jiparaná</td>
<td>vitîga</td>
</tr>
<tr>
<td></td>
<td>Aruá</td>
<td>weitią’a</td>
</tr>
<tr>
<td>Ramaráma:</td>
<td>Karo</td>
<td>pé tik pewit</td>
</tr>
</tbody>
</table>

Sweet potato is described by Birket-Schmidt (1943:19) as important in the Amazon region. The stability of the term in the Tupían languages is here taken as an indicator of this.
5.2.4 **Manioc, Cassava**

The Proto-Tupían root *mani ‘manioc’* (Rodrigues & Cabral 2012:563) is traceable in the following Tupían languages:

<table>
<thead>
<tr>
<th>Language</th>
<th>Form</th>
<th>Language</th>
<th>Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tupí-Guaraní:</td>
<td>madyiwa</td>
<td>Mondé: Suruí:</td>
<td>moy</td>
</tr>
<tr>
<td>Arikém:</td>
<td>gok myn hodnom</td>
<td>Tupari: Makuráp:</td>
<td>mání</td>
</tr>
<tr>
<td>Ramaráma:</td>
<td>nanoí</td>
<td>Purubora:</td>
<td>mákya</td>
</tr>
</tbody>
</table>

In Amazonia, “the most important food plant is beyond comparison manioc” (Birket-Schmidt 1943:19). Since Rodrigues and Cabral have been able to reconstruct a form for Proto-Tupí, it is most probable that manioc was grown by them for at least 5000 years, accounting for the relative stability of the term in this family.

5.2.5 **Calabash**

Rodrigues and Cabral’s (2012:563) reconstructed Proto-Tupían root: *ʔiʔa is only clearly identifiable in one of the languages in the material, namely KaritiânàTA: ió. The proto-form, however, consisting of two vowels separated by a glottal stop, contains little "substance" and is thus hard to grasp in potential descendants; the uncertainty of too short cognates is mentioned in section 3.1.

5.2.6 **Pumpkin**

The only language in the data in which the reconstructed Proto-Tupían root *kurua* (Rodrigues & Cabral 2012:563) is definitely visible is SuruíM: youkurah.

5.2.7 **Roots**

The reconstructed Proto-Tupí-Guaraní root for ‘root’ is *apó* (Jensen 1999:129), visible in both Tupí-Guaraní languages in the material: Amundava bepya and Uru-Eu-Wau-Wau yvapoa. This can be compared to two of the Tupí-Mondé languages: Suruí ih takot and Gavião do Jiparaná ihv ta poa. These languages are all spoken in a relatively limited area, so borrowing is possible as well as the scenario of genetic inheritance (no reconstruction for roots has been found in Proto-Tupí).
5.2.8 **Maize, Corn**

The maize terms were the ones who gave the most interesting results. Since maize is a traditional trade good of Amazonia, which, even though its long presence in the area, has been introduced from outside, it was expected to constitute an example of cross-genetic borrowing. By comparison to a former attempt of reconstruction (Carling, Eriksen, Holmer & van de Weijer to appear), I was able to connect the terms to the different reconstructed roots in cooperation with my supervisor. The list below shows the maize terms for all the languages in the data\(^5\), grouped by root.

<table>
<thead>
<tr>
<th>Root</th>
<th>Term 1</th>
<th>Term 2</th>
<th>Language 1</th>
<th>Language 2</th>
<th>Language 3</th>
<th>Language 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. <em>mahikano-ti</em></td>
<td>meek</td>
<td>Cinta Larga(^\text{TM})</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>meey</td>
<td>Suruí(^\text{TM})</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>maeg</td>
<td>Gavião do Jiparaná(^\text{TM})</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>maek</td>
<td>Aruá(^\text{TM})</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. *(i)v(i)n)kie</td>
<td>xi’a</td>
<td>Puruborá(^\text{TP})</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>shyki</td>
<td>Kaxarari(^\text{P})</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>txitxi</td>
<td>Jabutí(^\text{M})</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>haki</td>
<td>Aikanã(^\text{I})</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. *keMai(ki)</td>
<td>káya’hän</td>
<td>Mamaindé(^\text{N})</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>kiekini</td>
<td>Latundê(^\text{N})</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4a. *awatíi</td>
<td>awatía</td>
<td>Amundava(^\text{TG})</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>avaxia</td>
<td>Uru-Eu-Wau-Wau(^\text{TG})</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>atiti</td>
<td>Makuráp(^\text{TT})</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>atiti</td>
<td>Kanoë(^\text{I})</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. *ngaya</td>
<td>güjo</td>
<td>Karitiána(^\text{TA})</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>nāya</td>
<td>Karo(^\text{TR})</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. *malpak</td>
<td>mapak</td>
<td>Oro Eo(^\text{C})</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>mapac</td>
<td>Oro Nao(^\text{C})</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>mapak</td>
<td>Oro Mon(^\text{C})</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>papak</td>
<td>Oro Waram(^\text{C})</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>papak</td>
<td>Oro Waram Xijem(^\text{C})</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>mapak</td>
<td>Oro Win(^\text{C})</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The spread of the different roots in relation to the language families is visible in figure 6, where the form of the icon represents the root and the colour stands for the language family.

\(^5\) The Tupári term *opap* was unclassifiable.
Figure 6. The spread of maize roots in Rondônia. (Map compiled by Love Eriksen.)

The spread of terms for maize has been subject to research for a long time. Birket-Schmidt (1943) has provided a thorough investigation on which Carling et al. (to appear) have built their reconstruction and analysis. Their data, which is not presented in the article but which I have been given access to, has been combined with that of this analysis, resulting the maps which will be presented below.
Figure 7. The spread of maize roots in Amazonia. (Map compiled by Love Eriksen.)

Figure 8 is a close-up picture of Rondônia with the addition of the languages from the earlier data (represented by the icons without names) that are spoken in the same area. The following is noteworthy:

1. A Chapacuran language has been added which shares the same root as the other Chapacuran languages, and as can be seen in figure 15, this root is only found in Chapacuran languages.

2. On the Bolivian side, an isolate has been added which shares the root of Aikanã, Jabutí, Puruborá and Kaxarari.

---

6 Please note that the small families have been lumped together and share the brown coloured icon that marks the Chapacuran languages in the previous maps.
3. There are three additional Panoan languages on the map, none of which share the root of Kaxararíp (or, for that matter, of any other language in the original material).

Figure 8. The spread of maize roots in Rondônia, neighbours included (close-up of figure 7) 7. (Map compiled by Love Eriksen.)

7 Please note that the colour coding is back to normal.
5.2.9 Presentation of the maize roots


---

Figure 9. The spread of *mahikano-ti* in Amazonia. (Map compiled by Love Eriksen.)

This proto-root with its two variants is spread over a vast area and several language families in Amazonia. It corresponds to the forms of several Arawakan languages given in Birket-Schmidt (1943:20-21), and as he mentions, it is possible that this term has spread with the Arawakan expansion, since the Arawaks have long been considered the main force in the spread of maize.

Notable is that the proto-form *mahikano-ti* is found, except in the four Rondônian Tupí-languages, only in an isolate thousands of kilometres northwards. Probably, this is due to a classification error and should be regarded as a coincidence instead.
2. *(i)tii(n)k:e

Figure 10. The spread of *(i)tii(n)k:e in Amazonia. (Map compiled by Love Eriksen.)

The new data shows that the root seems to be found in languages of various genetic origin in Rondônia: PuruboráTP, JabutíM, KaxarariP and AikanãI.

This root is similar to the fourth Arawakan group of terms mentioned by Birket-Schmidt (1943:27-28), which he connects to the Panoan terms in his material (without mentioning any reconstructed root). In the analysis of Carling et al. (to appear), however, all the Panoan languages belong to a distinct root, number 14 *trukui. Since the KaxarariP term seems to be a descendant of this root *(i)tii(n)k:e, it might be possible that the two roots (2 and 14) share a common origin, especially as this has already been suggested by Birket-Schmidt – however, this is only a speculation.
3. *keMai(ki)

This root seems to correspond to the second Arawakan group of terms mentioned by Birket-Schmidt (1943:27). Interestingly, it has got a foothold in two Nambiquaran languages in Rondônia, far away from the other languages indicated on the map. A possible explanation is that this is a result of Arawakan maize trade.
4a. *awatī

Figure 12. The spread of *awatī in Amazonia. (Map compiled by Love Eriksen.)

This root is dominated by Tupí and has also been absorbed in a couple of other languages, mostly north of the Amazon River, as can be seen in figure 12. Birket-Schmidt argues that this root dominates the whole Tupían language family, but in figure 13, the map from figure 12 has been extended with information revealing which Tupían branches use this root. Interestingly, it is found only in languages from the Mawé, Awéti and Tupí-Guaraní branches⁸. These three branches have been argued to constitute a common subgroup in the Tupían language family known as the shortened form Mawetí-Guaraní (Drude & Meira 2012). Assuming that the *awatī root was borrowed into the Mawetí-Guaraní languages before they split into the three branches, this proposed

---

⁸The sole exception is Makuráp, belonging to the Tuparí branch, but the possibility of borrowing can not be ruled out in this case.
subgroup is thus supported by the results of this study (although maize terms alone, of course cannot count as evidence for genetic relationship).

Figure 13. The spread of *awati in the Tupian branches. (Map compiled by Love Eriksen.)
11. *ngaya

Figure 14. The spread of *ngaya in Amazonia. (Map compiled by Love Eriksen.)

This root is found only in Tupían languages in a very limited geographic area: KaritiânaTA, AruáTM and KaroTR-speaking groups in two different geographical locations. Cognates in other languages from the same branches, which are also spoken in the region, were found in Loukotka (1968:122). These are given below (the terms from the material of this thesis are given in brackets).

<table>
<thead>
<tr>
<th>Arikém</th>
<th>Ramaráma</th>
</tr>
</thead>
<tbody>
<tr>
<td>ngiyó (güjo)</td>
<td>nayá</td>
</tr>
<tr>
<td>Arikem</td>
<td>Itogapúc</td>
</tr>
<tr>
<td>Karitiâna</td>
<td>nanian</td>
</tr>
<tr>
<td></td>
<td>Ramaráma</td>
</tr>
<tr>
<td></td>
<td>noíábá</td>
</tr>
<tr>
<td></td>
<td>Urumí</td>
</tr>
<tr>
<td></td>
<td>náya</td>
</tr>
<tr>
<td></td>
<td>Urukú</td>
</tr>
<tr>
<td></td>
<td>(náya)</td>
</tr>
<tr>
<td></td>
<td>Karo</td>
</tr>
</tbody>
</table>
16. *malpak

Figure 15. The spread of *malpak in Amazonia. (Map compiled by Love Eriksen.)

It has been mentioned earlier that this root is only found in Chapacuran languages. When extending the area to all the available material for the whole Amazonian region, the prediction still holds. It seems as if this root has remained in the Chapacuran languages since before the split of the proto-language.

5.2.9.1 Summary: maize terms in Tupian languages

Birket-Schmidt (1943:49) argues that maize cultivation originated in Colombia around “the middle of the 2nd millennium [B.C.]” and spread to both Central and South America from there. Birket-Schmidt (1943:27) also writes that “nearly all Tupian tribes employ words of the same root in slightly different disguises. [...] Probably maize was known to the Tupians, before they spread over the Amazon area.” Since Proto-Tupí (which was spoken in Rondônia) is calculated to have started to split around 5000 years ago (Rodrigues & Cabral 2012:500), Birket-Schmidt thus indicates that the Proto-Tupians
cultivated maize before that. Assuming that Rodrigues and Cabral are right in their dating of Proto-Tupí and that Birket-Schmidt is right in his dating of the origin of maize cultivation, Birket-Schmidt must be wrong in his theory about the Tupíans cultivating maize before they split, since Proto-Tupí must have split even before the original Colombian maize cultivation had started. This is further supported by the results of the Tupían languages of this material, which have turned out to use words from four different roots. The distribution of these is shown below (the branches of the languages are indicated in brackets to the right).

1. *mahikano-ti Cinta Larga, Suruí, Gavião do Jiparaná, Aruá (Mondé)
2. *(i)vi(n)kie Purubórá (Purubora)
4a. *awatī Amundava, Uru-Eu-Wau-Wau (Tupí-Guaraní)
     Makuráp (Tuparí)
11. *ngaya Karitiâna (Arikém)
     Karo (Ramaráma)

4a. *awatī is the Tupían root that Birket-Schmidt suggests as dominating the entire Tupían family. 1 and 2 are probably loans from Arawakan languages. 11 is only found in Tupían languages. The results suggest that the Tupían languages in this study were introduced to maize from more than one source, probably at different times, supporting the “wave” theory of Danielsen et al. (2011). Root 4a, *awatī, is indeed the by far most widespread root of these, found in the three Tupían branches Tupí-Guaraní, Mawé and Awéti (MakurápTT is the sole exception and might be explained by borrowing). Rather than existing in Proto-Tupí, this root is more likely to originate in the proposed subgroup Mawetí-Guaraní which is argued to comprise Tupí-Guaraní, Mawé and Awéti (Drude & Meira 2012). Rodrigues and Cabral (2012:563) have provided Proto-Tupían reconstructions for e.g. ‘manioc’, ‘yams’, ‘sweet potato’, ‘tobacco’, ‘pumpkin’ and ‘calabash’, but not for ‘maize’ or ‘corn’. This, too, suggests that maize was not grown by Proto-Tupíans before the split.
6 CONCLUSION

This thesis has aimed to investigate the spread of culture terms among languages of different genetic origin in the Brazilian state of Rondônia. Some of the terms that were analysed showed convincing indications of language contact and borrowing, whereas other terms pointed in other directions.

As shown in chapter 5, the first impression of the results is the difference in the amount of borrowed items between the two semantic fields. The low amount of loanwords among the religion terms did not show any interpretable pattern. Probably mythology has been an important part of the proto-cultures of these groups and has not been subject to areal diffusion. Or else, it is possible that the word list, to a higher extent that we are actually aware of, is based on the terminology of Abrahamic religions and should be completely rethought before applied to regions such as Amazonia. It would be interesting to conduct a similar study on terms for flutes, hallucinogenic drugs or other material items which are known to be involved in religious ceremonies in this area. It is possible that these terms would show a higher incidence of borrowability as they refer to physical objects which are known to have spread, as opposed to abstract notions such as ‘god’.

Although I have no percentage numbers to present, an impressionistic comparison can be made to the findings of Haspelmath and Tadmor (2009) in their Loanword Database project. Tadmor (2009:64) lists the semantic fields by borrowability, indicating the percentage of loanwords in every field in the database. Interestingly, the list is headed by Religion and belief with 41.2% loanwords. The reason Tadmor gives for this is obvious: the spread of the large religions such as Christianity and Islam.

As is clear from the results of this thesis, this is not at all applicable to the languages of Rondônia. The religion terms showed almost no results whatsoever. Since Tadmor (2009)’s explanation for the high borrowability in this semantic field is the vast spread of the large religions, the explanation for the low borrowability among these languages could be that they have kept their original beliefs, not adapting to the Christianity of the colonizers. However, it must be strongly emphasized that this is nothing but a speculation. It would indeed be preferable to connect the results to valid, anthropological research, but not a single source has been found as to the contacts
between missionaries and Rondônian peoples. As mentioned earlier in this chapter, it is possible that the word list should be rethought to better fit the Amazonian context, which might yield different results.

The semantic field of *Agriculture and vegetation* appears as number six on the list with 30.0% loanwords. This is more in line with the results of the Rondônian terms, although very little can be said in general about these terms. They displayed a vast difference in borrowability. It was the maize terms that far and away exhibited the greatest amount of cross-genetic diffusion, which is in line with the expectations since it is known that maize was introduced through trade from outside the Amazon basin and does not belong to the original vegetation of the tropical rainforest (Birket-Schmidt 1943). This fact accounts for the higher likeliness of borrowing the term for ‘maize’.

The case study of the maize terms also suggests some language contact between Rondônian and extra-Rondônian languages. For instance, some of the maize roots found in the data were connectable to roots found in Arawakan languages. The Arawaks used to dominate the trade on the rivers between the Andes and the Amazonian basin, and maize was one of the goods they brought with them, suggesting that this is a possible interpretation. The fact that several Arawakan roots were found in non-Arawakan languages in Rondônia further supports the theory that the Arawakan languages spread in “waves” (Danielsen et al. 2011). An interesting field of future research would be to investigate other parts of the vocabulary of these non-related languages sharing maize roots to see what other goods may have played a part in these contact situations.

Root 4a. *awatǐ* showed support for the Mawetí-Guaraní hypothesis (Drude & Meira 2012), by being found only in the Mawè, Awéti and Tupí-Guaraní branches (with one exception).

On the other hand, there were Tupían terms that showed stability in several branches and similarity to the proto-forms, such as ‘manioc’ and ‘sweet potato’. First of all, the mere existence of proto-forms for these grains implies that they have most probably been cultivated by Tupians before Proto-Tupí split into its sub-branches; and second, the stability of the terms indicates that these crops were (and are) an important part of the Tupían culture, accounting for the resistance against borrowing.
Unfortunately, there is much less data available for all the other language families, so the Tupían terms are highly overrepresented in the presentation of the results. This has to do with the fact that reconstructions for the terms concerned in this thesis have only been found for one proto-language: Proto-Tupí (and to a certain extent for Proto-Tupí-Guaraní).

It should also be mentioned that Kaxararí, the only Panoan language in the data, shows similarities in several terms to various languages in the south eastern part of Rondônia: MamaindéN, LatundêN, AikanãI, PuruboráTP. Interestingly, Kaxararí has much less in common with the languages in the more central parts of Rondônia, which are located much closer. Since the Panoan languages originate in Peru and Kaxararí is one of the easternmost of these, it is highly unlikely that it has ever been spoken in south eastern Rondônia and then moved westwards. The only plausible explanation is the connection via the Guaporé River, along which both Kaxararí and the languages which it shares vocabulary with are spoken.

To conclude, the main aim of the study was to investigate the language contact in Rondônia, and the tendencies that were found were the following. The low borrowability of the religion terms indicates that the traditional religions and mythologies have been kept in these cultures, or alternatively that the list should be revised. The agriculture terms cannot as easily be generalized upon, and demand an individual presentation. Traditional Amazonian crops like sweet potato and manioc show a stability among the Tupían terms which is in line with what is known about borrowability. Maize, on the other hand, which is a trade good, shows a complex distribution pattern which was relatable to several theories. The spread indicated extensive contact with Arawakan languages. Support was also found in the spread of these terms for both the Mawéti-Guaraní hypothesis (Drude & Meira 2012) and for the “wave” theory (Danielsen et al. 2011). The spread of some terms in both semantic fields showed indications of the KaxararíP frequenting the Guaporé River.
7 REFERENCES


ONLINE SOURCES:


8 ATTACHMENT

LANGUAGES IN THE MATERIAL: SPELLINGS, CODES AND ALTERNATIVE NAMES

Aikanã [tba] Aikaná, Corumbiara, Huari, Kasupá, Kolumbiara, Masaká, Mundé, Tubarão, Uari, Wari
Amundava [adw] Amondawa, Amondáwa, Amundawa, Amundáwa
Aruá [arx] Aruaxi, Aruashí
Cinta Larga [cin] Cinta-larga
Gavião do Jiparaná [gvo] Digüt, Gavião, Gavião do Rondônia, Ikôro
Jabutí [jbt] Djeoromitxi, Dheoromitxí, Kipiú, Jabotí, Quipiú, Yabutí
Kanoé [kxo] Canoé, Canoë, Guaratégaya, Guaratêgaja, Koaratira, Guaratira,
             Amniapé, Kapixaná, Kapixana, Kapishanã
Karitiána [ktn] Caritiana, Karitiána, Karitiana
Karo [arr] Arara, Arára, Arára de Rondonia, Arára do Jiparaná, Arara-Karo,
             Itanga, Itogapuc, Itogapúk, Ntogapid, Ntogapig, Ramarama, Uruku,
             Uruukú
Kaxararí [ktx] Kasharari, Kaxariri
Latundê [ltm] Leitodu
Makurá [mpu] Macuráp, Macurate, Macurapi, Makurápi, Massaka
Mamaindé [wmd] Northern Nambiquara, Mamande, Nakarothe
Oro Win [orw] -
Puruborá [pur] Aurã, Boruborá, Burubora, Cujubi, Kuyubi, Miguelinho, Migueleno,
              Pumbora, Puroborá, Puruba
Suruí [sru] Paiter, Suruí de Rondônia, Suruí do Jiparaná, Suruí Paiter
Tupari [tpr] -
Uru-Eu-Wau-Wau [urz] Eru-Eu-Wau-Wau, Kagwahiva, Uru-Eu-Uau-Uau, Uruewawau,
                      Urueuwawáu

*Oro At Oro Eo, Oro Mon, Oro Nao, Oro Waram and Oro Waram Xijem are not found in Ethnologue and are probably considered dialects of Oro Win.*
**OTHER LANGUAGES MENTIONED: SPELLINGS, CODES AND ALTERNATIVE NAMES**

<table>
<thead>
<tr>
<th>Language</th>
<th>Alternative Names</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arikapú [ark]</td>
<td>Aricapú, Maxubí</td>
</tr>
<tr>
<td>Arikem [ait]</td>
<td>Ariken, Arikém, Ariquême</td>
</tr>
<tr>
<td>Conjubim</td>
<td>-</td>
</tr>
<tr>
<td>Kwaza [xwa]</td>
<td>Coaia, Koaiá, Koaya, Koayá, Quaiá, Arara</td>
</tr>
<tr>
<td>Ramaráma</td>
<td>-</td>
</tr>
<tr>
<td>Sabanès [sae]</td>
<td>Sabané, Sabanê, Sabanés, Sabanes, Sabones, Sowainnte</td>
</tr>
<tr>
<td>Wayoró [wyr]</td>
<td>Ajurú, Ayurú, Uaiora, Wajaru, Wayru, Wayurú</td>
</tr>
</tbody>
</table>