PONTIFÍCIA UNIVERSIDADE CATÓLICA
DO RIO DE JANEIRO

Raiane Oliveira Salles

UNDERSTANDING RECURSION AND LOOKING FOR SELF-EMBEDDING IN PIRAHÃ
The case of possessive constructions

Dissertação de Mestrado

Thesis presented to the Programa de Pós-Graduação em Estudos da Linguagem of the Departamento de Letras, PUC-Rio as partial fulfillment of the requeriments for the degree of Mestre em Letras/Estudos da Linguagem.

Advisor: Prof. Cilene Aparecida Nunes Rodrigues
Co-Advisior: Prof. Marcel Den Dikken

Rio de Janeiro
August 2015

PONTIFÍCIA UNIVERSIDADE CATÓLICA
DO RIO DE JANEIRO

Raiane Oliveira Salles

UNDERSTANDING RECURSION AND LOOKING FOR SELF-EMBEDDING IN PIRAHÃ
The case of possessive constructions

Thesis presented to the Programa de Pós-Graduação em Estudos da Linguagem of the Departamento de Letras do Centro de Teologia e Ciências Humanas da PUC-Rio, as partial fulfillment of the requeriments for the degree of Mestre.

Profa. Cilene Aparecida Nunes Rodrigues
Advisor
Departamento de Letras – PUC-Rio

Prof. Marcel Den Dikken
Co-Advisor
The City University of New York, EUA

Prof. Marcus Antonio Rezende Maia
UFRJ

Profa. Kristine Sue Stenzel
UFRJ

Profa. Denise Berruezo Portinari
Coordinator of the Setorial do Centro de Teologia
e Ciências Humanas – PUC-Rio

Rio de Janeiro, September 28th 2015.

Raiane Oliveira Salles

Graduated with a degree in Letters Portuguese/English from Veiga de Almeida University, in 2008. Participated of various events and courses and presented works in events in the área of Language and Cognition, Generative Linguistics and Indigenous Languages. Areas of interest: Language and Cognition, Generative Linguistics, Syntax and Indigenous Languages.

Bibliographic data

Acknowledgments

First of all, I want to thank the Pirahã people for welcoming a stranger in their territory. Thank you for *kapiigakagakai* with me, giving me the privilege to understand your language. You asked me to show you pictures from my family, friends, my house and my city. Now, my family and friends want to see pictures from you. They are marveled by the way you live. I am marveled by how much better your society works when compared to mine. Thank you for sharing your laughter with me and for teaching me that life is more than what can be written in a book or shared online. You never told me so, but observing the way you value the experience of having good food, company and laughter, made me see that it is all that really matters in life.

The second time I went to the Pirahã village of Piquiá, Augusto Pirahã gave me all the support I needed in terms of logistics. He makes all the efforts he can to help his people protect their territory, traditions and language. He knows the importance of work conducted by linguists and he is happy with those who want to be involved in projects for the benefit of the Pirahã people. He also tries his best to guarantee that no right is denied to Pirahãs as an indigenous society. Thank you Augusto, for helping the Pirahã people and for trusting in my work.

Another friend, who made me feel home was Renata Leite, the teacher in the Pirahã school at Piquiá. We had a lot of fun working and living together in the Piquiá village. Thank you, Renata, for letting me stay over when I was in Humaitá and for our fun talks about life. Thank you for taking care of me when I got sick in the village. I learned a lot from you!

Many thanks to Funai for the support they offered me in Humaitá and also for having invited me to work on the project of documenting cultural and linguistic aspects of all the Pirahã villages.

Cilene Rodrigues, you inspired me to pursue dreams which were bigger than everything I could have dreamed by myself. You brought to my life the most inspiring work I have already done. The excitement with which you conduct your work has made me see how special Linguistics can be. Doing formal linguistics and fieldwork, especially in Pirahã, has been a great adventure. You have not only been a professor to me, but also a friend. I would never find the best words to thank you for all the things you taught me. You did way more than what was expected from you if we just followed protocols. Thanks, Cilene, for encouraging me when not even I myself believed I could. Thank you for caring so much about me and the work we were doing. Thank you for having seen in a simply very curious girl a potential linguist. My debt with you will certainly be remembered when I become a professor myself. I hope my future students will thank you for have been my model!

Marcel den Dikken also helped me a lot, even before my fieldwork started. His ideas presented in Glow 37, in Brussels, 2014, made me travel from Rio to São Paulo, just to talk to him over lunch about my data. He did more than that. He gave special attention to my inquiries and pointed out important issues I should look for during fieldwork. Whenever I had internet access during fieldwork, Marcel was on the other side of the world ready to talk to me and help me understand what is going on in the Pirahã language. Thank you, Marcel, for sharing with me so many good insights.

I was very lucky to have many excellent and dedicated professors, who are to be responsible for every good aspect to be found in my career. Many thanks to them all. Letícia Sicuro Corrêa, you opened my research horizons during the course *Introduction to Psycholinguistics*, which was indeed one of the courses which helped me make the transition from a curious student to a researcher. Professor Erica Rodrigues, your passionate classes on the evolution of linguistics made me fall in love again with the field, even when Cilene's hard core syntax classes were killing me softly! Cilene Rodrigues, your provocative classes made our brains burn, but we definitely learned by force that in science thoughts have to be structured. Sabrina Anacleto, you are a professor to me! Sabrina is our genius and lovely friend,

who translated for us when we could not follow Cilene's class. Thanks, Sá!

Luckily, I also had the opportunity to do some courses in other universities, which expanded my horizons and put me in contact with other areas and ways to look at language. I want to thank professor Marcus Maia, for telling me that research with indigenous languages involve much more than going to a people's land and asking them if something is or is not possible in their language. I needed to have a more complete understanding of their picture. Otherwise, I would be doing what the colonizers did and just be trying to "rob their souls" now. Professors Marcus Maia, thank you for our inspiring talks and for indicating to me the courses I should take that semester I spent at UFRJ.

Professors Andrew Nevins and Gean Damulakis, thank you for teaching us an insightful way to look at phonology. It was an honor to be able to study with you. I need to thank my colleagues in this course, especially Elis Barros, who helped me when things were blurred up in my mind!

Professors Bruna Franchetto, Mara Santos and Kristine Stenzel thank you for sharing your fieldwork experience with my colleagues and I. We not only learned how to use linguistic tools, but we also had the opportunity to build important elicitation strategies and to think beyond our areas of expertise. The semester we spent together in the National Museum was great. Again, thanks Elis Barros, and also Nathalie Vlckova, for sharing fieldwork and computer tools experiences with us!

Thank you, Professor Kristine Stenzel, for showing me languages are so beautifully different, and at the same time, not that apart from each other. I fell in love with each morpheme I had to analyze in your typology exercises. Marcia Nascimento and Ananda Machado, thank you for sharing with me lovely afternoons making cross-linguistic comparisons!

My family deserves my acknowledgment too. My father always did the best he could to offer me good education and is definitely responsible for my having achieved the position I am in now. He always made me feel strong and helped me with difficult decisions. My mother taught me to fight for what I want and to never forget about dreaming. She is always there when I need mommy back, and taught me that hoping and fighting is better than giving up. To the rest of my family and all my friends: thank you for the positive thoughts and for understanding that I cannot be with you in the family parties or go out with you as much as I wanted. I promise I will try to compensate when that big house with a pool and a barbecue grill is mine!

Brayan Seixas, thank you for being the most perfect boyfriend, fiancé, and now husband. For ten years we have been growing up together and helped each other to dream and to have the strength to carry on when the dreams finally come true (or not!). You have been amazing, extremely helpful while I was in the bubble writing this thesis. Thanks for helping me to see I could reach the things I wanted when I was about to give up with not even trying. You are my fellow for life, more than my husband, my family, my best friend. Thank you.

Abstract

Salles, Raiane Oliveira; Rodrigues, Cilene Aparecida Nunes (advisor); Dikken, Marcel den (co-advisor). **Understanding Recursion and Looking for Self-embedding in Pirahã: the case of possessive constructions.** Rio de Janeiro, 2015, 129p. Master Thesis - Departamento de Letras, Pontifícia Universidade Católica do Rio de Janeiro.

It has been claimed that Pirahã, a Brazilian native language spoken in the Amazon region, is non-recursive, disallowing syntactic self-embedding altogether (Everett, 2005). This thesis investigates this claim. First, a formal definition of recursion is necessary, and we examine how this term appeared within mathematics and logic and how it made its way to formal linguistics. Our conclusion is that within Generative Grammar, recursion is to be understood as a finite set of functions that calls for itself, taking its previous output as its input (i.e. the operation Merge). It is responsible for the combinatorial nature of Grammar, a universal cognitive capacity, which confers I-language with its core property: discrete infinity. The unavailability of self-embedding in a given language is not evidence against the universality of recursion. Self-embedding representations are one, but only one of the possible external outcomes of the recursive computations within I-language. The second goal of this thesis is to present new fieldwork data on possessive DPs in Pirahã, showing that self-embedding in this structural domain is possible and productive. Putting it all together, our conclusion is that Everett's claim about the universality of recursion is both theoretically and empirically incorrect. The contribution of this thesis to the field of formal linguistics is twofold: it offers a better understanding of computability within I-language, and a new empirical-based assessment of syntactic embedding in Pirahã.

Keywords

Resumo

Salles, Raiane Oliveira; Rodrigues, Cilene Aparecida Nunes (orientadora); Dikken, Marcel den (co-orientador). **Entendendo Recursão e Buscando por Auto-encaixamento em Pirahã: o caso das construções possessivas.** Rio de Janeiro, 2015, 129p. Dissertação de Mestrado - Departamento de Letras, Pontifícia Universidade Católica do Rio de Janeiro.

Tem-se argumentado que a língua Pirahã, da família Mura, falada na Amazônia brasileira, é não-recursiva, não apresentando nenhum tipo de auto-encaixamento sintático (Everett, 2005). O objetivo dessa dissertação é investigar essa questão. Primeiramente, uma definição formal de recursão é necessária. Apresentamos, assim, uma análise histórica do termo, considerando sua origem no campo da matemática e da lógica e sua trajetória dentro da linguística formal Chomskyniana. A conclusão a que se chega é que dentro da Gramática Gerativa, recursão deve ser entendida como um conjunto finito de funções que chamam por si mesmas, podendo tomar seu outup prévio como input (i.e. operação *Merge* do Programa Minimalista (Chomsky, 2015)). Esse recurso cognitivo universal é responsável pela natureza combinatorial da Gramática (Língua Interna), conferindo a ela a propriedade de infinitude discreta. Assim, a ausência de auto-encaixamento em uma determinada língua não é evidência contra a universalidade da recursão. Representações com auto-encaixamento são apenas um dos possíveis resultados externos do sistema combinatorial recursivo da Gramática interna. O segundo objetivo dessa dissertação é apresentar novos dados de pesquisa de campo em estruturas nominais possessivas em Pirahã, demonstrando que auto-encaixamentos nesse domínio sintático são possíveis e produtivos na língua. A conclusão geral da nossa pesquisa é que a posição de Everett contra a universalidade da recursão está tanto teoricamente, quanto empiricamente incorreta. Duas são as principais contribuições dessa dissertação para a teoria linguística formal: oferecer um melhor entendimento de computabilidade interna à Gramática, e apresentar novos dados empíricos sobre a sintaxe de encaixamentos em Pirahã.

Palavras-chave

Gramática; infinitude discreta; recursão; auto-encaixamento; Pirahã; expressões nominais possessivas.

Table of Contents

List of Glosses

1 – 1st person

2 – 2nd person

3 – 3rd person

ATEL – atelic

COMPLETE CERT – complete certainty

CONT – continuative

COP – copula

DEM – demonstrative

DUR – durative

FEM – feminine

GEN – genitive

IMPERF – imperfective

INGR – ingressive

ITER – iterative

LOC – locative case

NEG – negation

NMLZR – nominalizer

NR – nominalizer

OBL – oblique

PERF – perfective

PRES – present

PROX – proximal

PUNCT – punctiliar

RELATIVE CERT – relative certainty

REM – remote

SG – singular

WH – question word

1
Introduction

Philosophers and scientists have been marveled by our species' unique capacity to acquire a set of finite symbols and creatively produce sentences with them. Galileo Galilei (1632), for example, observed that the ability to communicate our thoughts through written language, producing different arrangements out of a finite set of alphabetic letters is an outstanding human capacity. Arnauld & Lancelot (1660) (the grammarians of Port Royal) also noticed this combinatorial power of language on our ability to produce infinite words using finite sounds. Von Humboldt's (1836) renowned sentence generalizes this creative aspect of language: it can "make infinite use of finite means"[1]. This has been called the discrete infinity property of language.

Empirical observation also shows that, although languages may differ, discrete infinity is a universal human capacity. All societies have their linguistic code, which can be naturally acquired by children. Other animals, on the other hand, may develop their communicative systems, but those are radically different from language. As a naturalist, Darwin (1882) observed a great variety of species, noticing that many animals are able to express emotions through sounds, but a human peculiarity is "the habitual use of articulate language" (Darwin, 1882: 85). According to him, there is a larger human capacity of combining diverse sounds to ideas if compared to other species'. Discrete infinity is, thus, what makes us so unique with respect to communication.

Hockett (1960) also conducted empirical observation of the characteristics shared among human languages and their peculiarities before other animal communication systems. He reached to thirteen design features of human language, namely: the vocal-auditory channel, broadcast transmission and directional reception, rapid fading, interchangeability, total feedback, specialization, semanticity, arbitrariness, discreteness, displacement, productivity, traditional transmission and duality of patterning. It has been long disputed whether all these features are human specific and universal (see Lobeck & Denham, 2012). Nevertheless, as Hockett observed, it is only in humans that one can find all the

---
[1]Chomsky (1965: 8)

thirteen features. It is not hard to link features such as productivity and discreteness to the universal property of discrete infinity.

Hauser, Chomsky & Fitch (2002) propose a different way to see the design of the faculty of language. The authors divide it into a broad part – including capacities which are shared with other animals or other cognitive systems – and a narrow part – including only what is unique to language. Following cross-linguistic analysis and comparative studies with other species, the authors conclude that <u>recursion</u> is the human capacity conferring our linguistic system with one of its core properties: discrete infinity.

This claim has been disputed by linguists, psychologists and biologists, who either find recursion not to be unique to language (Pinker & Jackendoff, 2005), or universal (Everett, 2005), or not even specific to humans (Gentner et al, 2006). The problem is that the concept of recursion is too broad, allowing for different interpretations for Hauser, Chomsky & Fitch's (2002) proposal. The authors themselves did not provide a fine definition of what they meant by recursion.

This brings us to one of the goals of this thesis, which is to reconstruct the way made from the mathematical concept of recursion to its understanding within linguistics. From this historical viewpoint, we reach Chomsky's theory and its developments from the 1950's to 2000's, concluding then that the recursive mechanism of the faculty of language is the mechanism behind the combinatorial power of Grammar, which, within the minimalist program, is called <u>Merge</u>, the operation of the linguistic computational system that applies over its own outputs. As Chomsky (2007) observes, Merge is <u>the</u> property which yields a discrete infinity of structured expressions in human language.

However, this capacity for general recursion also allows for representations involving recursive Merge of different tokens of the same type, that is, Merge of syntactic objects with the same label. That is to say that self-embedding within one category is possible. This subtype of recursion, also called specific recursion (van der Hulst, 2010), has been taken by many authors to instantiate what Hauser, Chomsky & Fitch (2002) meant by recursion. One of these authors is Everett (2005) who reports a Brazilian Amazonian language, Pirahã, which supposedly does not have self-embedding at all. This is a misinterpretation of the term as we will see in this thesis (see chapter 2). As for lack of self-embedding in Pirahã, we will argue that Everett's analysis does not contemplate all the grammatical data found in the

language (see chapter 3).

Everett attributes the absence of syntactic embedding in Pirahã to a cultural constraint, claiming that culture is able to interfere with grammar. He concludes that Pirahã is a major counter evidence for recursion as a universal human capacity. Nevins, Pesetsky & Rodrigues (2009a&b) dispute Everett's cultural explanation for Pirahã lack of self-embedding, showing that languages from societies as different as Germany from Pirahã present similar restrictions in their syntax, such as German ban on prenominal recursive possessors within possessive phrases.

More recent fieldwork have gathered data pointing to self-embedding in several phrasal domains in Pirahã, such as Postpositional Phrases (Amaral et al, forthcoming) and Verbal Phrases (Rodrigues et al, forthcoming). In order to contribute to this literature, one of the goals of this thesis is to present data from my recent fieldworks in Pirahã[2], suggesting the availability of self-embedding within the nominal domain.

In the next sections, I will present a general introduction to the Pirahã society (section 1.1) and some aspects of their grammar (section 1.2). Then I briefly summarize our discussion on the concept of recursion, and the importance of a formal definition of this term to the understanding of Pirahã as a human language (section 1.3). Finally, I state the contributions of this thesis and present its organization (section 1.4).

1.1 General aspects of the Pirahã society

Pirahã is the name commonly used to refer both to an Amazonian indigenous people and to their language. Using their own language terms, they are *hiatihi* "Pirahã people" and they speak *apaitiso* "Pirahã language", literally "that which comes from the head" (cf. Gonçalves, 2001).

The Hiatihi were first mentioned by Ferreira Pena (1853), who called them "Pirianaus". Later, James Orton (1870) referred to them as "Piarrhaus" (Nimuendaju, 1982). In 1920, Nimuendaju found a group of Hiatihi and called them Pirahãs, recognizing them as descendants from the Mura people[3]. According to Gonçalves

---

[2]Fieldwork September, 2013 and October-December, 2014.
[3]According to Nimuendaju (1948), the Mura started their migration from Peru to Brazil in the seventeenth century, and successive attacks promoted by the Portuguese in the colonial period

(2001), the most plausible hypothesis in trying to trace back the Pirahã separation from Mura is that, while a Mura group left their territory towards the Madeira river, another group of Mura settled down in one of their transitory but original territory, today known as the Pirahã territory. This group came to be known as Pirahã.

Pirahãs live in different villages along the Maici and Marmelos rivers. There is also a little group living by the Ipixuna river, inside Parintintin (another indigenous society) territory. All these areas are located in the municipality of Humaitá, in the state of Amazonas, Brazil, and were already demarcated by Funai as Pirahã territory. According to the last census registered by IBGE[4] (2010), there were 420 Pirahãs in their whole territory. The map below, provided by the State University of Amazonas - UEA (2013), shows us the main Pirahã villages[5].

---

led to their expansion to other territories, such as the region of the Madeira river.

[4]Instituto Brasileiro de Geografia e Estatística (Brazilian Institute for Geography and Statistics).

[5]The Pirahã family living within Parintintin territory is not shown in this map.
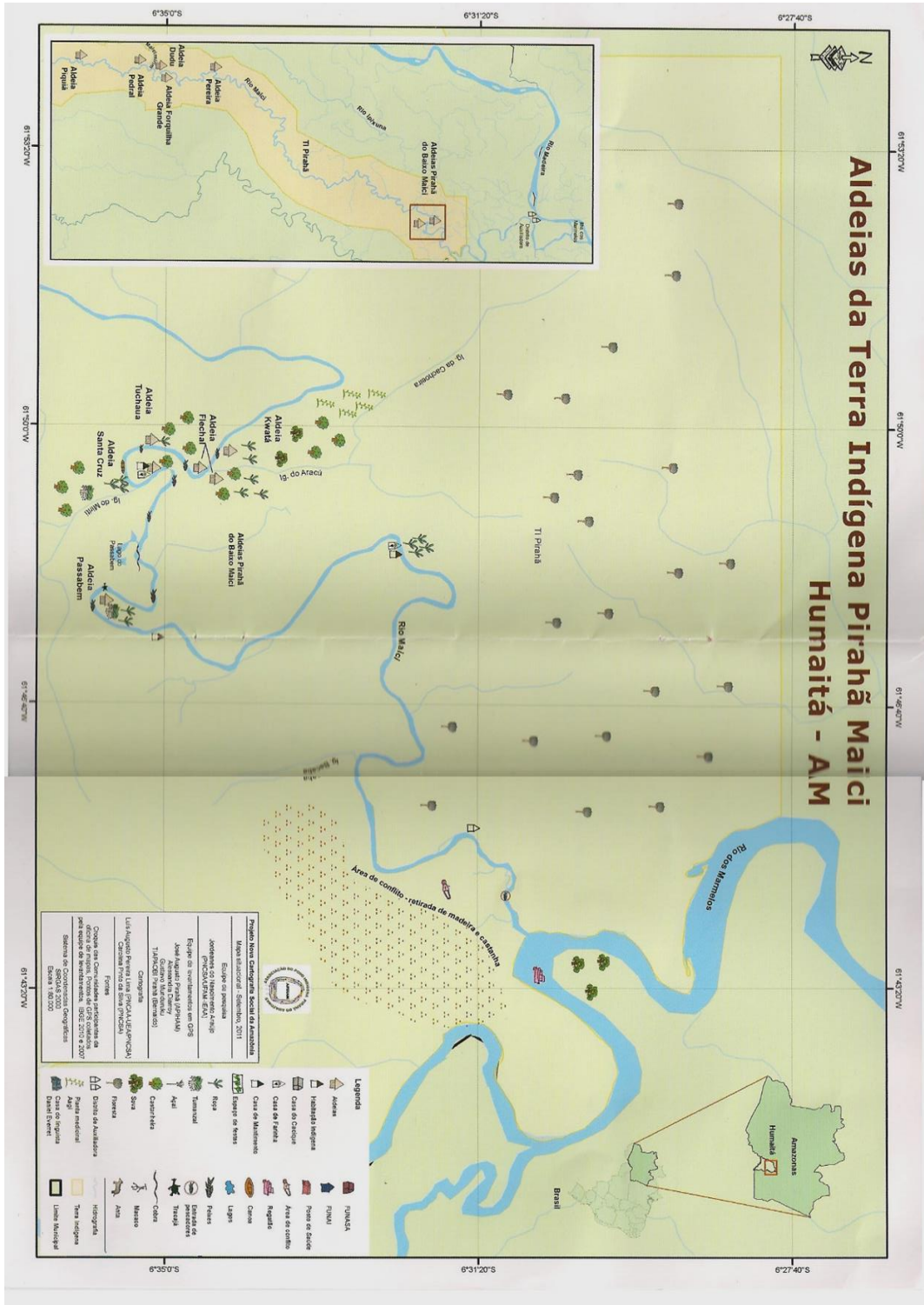
Figure 1 - Villages of the Indigenous Land of Pirahã, Maici, Humaitá- AM[6].

---

[6]The final elaboration of this map was done based on drawings made by the Pirahãs themselves, under the coordination of a team of cartographers from UEA. See the appendix photos 1, 2, 3 and 4. Photo 1 is the original map drawn in Piquiá. Photo 2 shows a Pirahã man drawing the map. Photo 3 shows Pirahãs from Piquiá exhibiting their map, which was the basis for the high Maici part of the map in Figure 1. Photo 4 shows Pirahãs from the lower Maici exhibiting their map, the basis for the lower Maici part of the map in Figure 1.

Sixty Pirahãs live in the main village, called Piquiá, where my fieldwork was mainly conducted. This is one of five villages[7] located in the *high Maici*, the part of the river closer to the south limit of the Pirahã territory (the bridge over Maici, on the *Transamazônica* road). Piquiá received the first Pirahã school, which is listed as one of the indigenous municipal schools from the Municipal Secretary of Education of Humaitá. One teacher has been working there since June, 2014, trying to alphabetize Pirahãs in Portuguese. However, neither Pirahãs speak Portuguese fluently, nor does the teacher speak Pirahã. She has been making some progress, though, since some of them can already write down their names and even recognize quantities in the Portuguese system from 1 to 10. The students are mostly men and children, although some young married women, with their husbands, may attend the classes sometimes. This state of affairs points towards the necessity of linguistic intervention in the local educational system, to ensure alphabetization in their native language.

Most of the Pirahã men can communicate with Brazilians using a pidgin composed by Portuguese words and some words from Nheengatu, an Amazonian lingua franca used in the 19th century (Navarro, 2011). As an example, consider (1), where the word *kunhã* "wife", from Nheengatu, is used together with Portuguese words but the grammatical structure of Pirahã is somehow preserved. [Source: Salles fieldwork, 2014].

(1)   kunhã aqui disse   aqui eu quer guaraná tamém
        wife   here say.3.PAST here 1   want guaraná   too
        'My wife here said: I want guaraná[8] too'

Women are less communicative with foreigners, but I witnessed them pronouncing Portuguese words too. Among themselves, children seem to speak only Pirahã, and, as far as I can tell, they are uncommunicative with foreigners[9].

Regarding their food habits, Pirahãs mostly eat fish, although they sometimes

---

[7]The other villages are Pedral, Forquilha Grande, Dudu and Pereira. There is also a small village near Piquiá which is called Pagão, but the two families who live there go almost every day to Piquiá to take the children to school. The adult men also study.

[8]A Brazilian soft drink.

[9]Although, in one situation, when I gave a toy to one of the children, he spoke to me in Pirahã: *maxa* 'beautiful'.

hunt small mammals such as agoutis and pacas. They also grow manioc and some fruit (such as mango, guava and banana), collect bee honey, copaiba honey and Brazilian nuts. Coffee, sugar, salt, rice, oil and other industrialized products have been introduced to them through the contact with Brazilians. The most common practice is to exchange their products for the things they want. They are also interested in hooks, fishing line, flashlights, batteries, clothes, perfume and beads which adorn their necklaces and bracelets, hanged in a fishing line or in a thread[10].

They face two seasons during the year: a dry one (from May to October) and a rainy one (from November to April). During the dry season, the Maici river is low and sand tracks appear, so they build their houses and rest on the beaches. In the rainy season the river flows, and they build their houses on the high banks.

Piquiá and Forquilha Grande are the only two villages with fixed health centers built by Funasa[11]. Nurses work there in the control of diseases such as malaria and tuberculosis, also leading vaccine campaigns and giving first aid when needed. Whenever an injury or disease needs specialized treatment, they take the patient to public hospitals in Humaitá or Porto Velho - RO.

The contact with Brazilians in the high Maici villages is mainly with the professionals of the school and the health centers, although the Pirahãs also travel to the bridge over Transamazônica, where there is a little Brazilian restaurant for travelers. They go to this place to collect mangoes, beg for soda and cookies in the restaurant and remain there for a while in transitory houses. The Pirahãs from the lower Maici (north limit of the territory, where the Maici flows into the Marmelos river), have more frequent contact with Brazilians because they exchange their products (e.g. copaiba and Brazilian nuts) for non-native products (such as coffee, sugar and tabaco). These products are brought by little merchants from nearby villages and districts, such as Auxiliadora. The communication in those situations is established with the pidgin mentioned before.

Given the little amount of time we spent in the village so far, we do not have yet a complete picture of their culture. However, we witnessed several cultural manifestations such as traditional chanting and dancing, and the elaboration of necklaces and other adornments[12]. For a detailed ethnology of this society, we

---

[10]See the appendix, photo 5.
[11]Fundação Nacional de Saúde (National Health Foundation).
[12]See the appendix, photos 6 and 7.

recommend the work done by the Brazilian anthropologist Marco Antonio Gonçalves (1993, 2000, 2001, 2005)[13], who defends that Pirahã is not a cultural exception within the Amazonian context.

In the next section we give a general overview of *Apaitiso*, the Pirahã language.

1.2 General aspects of the Pirahã grammar

While the majority of the indigenous languages in Brazil descends from the the stocks Tupi and Macro-jê, Pirahã was classified as a language from the Mura family by Nimuendaju (1982). Mura is an isolated family consisting of only the Mura and Pirahã languages. The Mura language is now extinct. Nowadays the Mura people are monolinguals, speaking only Brazilian Portuguese (Amoroso, 2009). As for Pirahã, Brazilian Portuguese has not yet been inserted in their society and their monolingualism makes the alphabetization in Portuguese inefficient. Observing the preservation of Pirahã is highly important, since it is the only surviving language in the Mura family and it is spoken by a few number of individuals, who do not use a written system. Thus, the lack of alphabetization in Pirahã may represent a future threat into this language, as children are now being inserted in a new oral and written system. An educational intervention enforcing alphabetization in Pirahã is one of our goals in the next years. A proper documentation of their language and a good understanding of their grammar are *sine qua non* conditions for elaboration of didactic materials to be used in the local school.

Although many studies have been conducted on Pirahã, such as Heinrichs (1964), Sheldon (1974, 1988), Sandalo (1989), Keren Everett (1998), Topintzi (2004), Gordon (2005) – most of these on Pirahã phonology – the most complete

---

[13]Personally, I could briefly access the Pirahã cosmology in my last fieldwork. One day, I inquired Kobio Pirahã, an elder, about the identity of the entity *xigagai* (see Gonçalves, 2001). He explained to me that it was similar to what Brazilians call *papai do céu* "the heaven daddy" (God). Then, he told me Xigagai takes care of the children only. According to him, another entity is *kaoaibogi* (also cited in Gonçalves, 2001) which comes to play with them during the festivals. To call for this entity, they stump their feet on the ground calling its name. When using the pidgin, they referred to Kaoaibogi as *Orupari*, or Jurupari, cited in the literature as a God worshiped by Amazonian indigenous people. He was later associated with the christian devil by missionaries (cf. Cascudo, 2002). When I was stomach sick in Piquiá, Kobio also advised me to wear perfume in order to send away the bad spirits that were causing my illness. This suggests that Pirahã people have their own cosmology which is partially shared with other Amazonian people as argued by Gonçalves (2001).

work on Pirahã grammar was conducted by Daniel Everett, with the most notable publications being Everett (1983, 1986 & 1992). He started his contact with the Pirahãs as a missionary of the *Summer Institute of Linguistics - SIL*, and during the 1980's he did his PhD at Unicamp, in the state of São Paulo, working on aspects of the Pirahã grammar. The features of the Pirahã grammar presented in this section are both based on Everett's data and the data I collected during my fieldwork in Piquiá, from October to December, 2014.

Pirahã has a small segment inventory, with eight consonants: /p/, /b/, /t/, /k/, /g/, /m/, /n/, /ʔ/, /s/, /h/, and three vowels: /i/, /a/, and /o/[14]. Some facts are worth noting. First, /b/ and /m/ are used interchangeably by speakers (2), as well as /g/ and /n/ (3). Second, women do not use /s/, using rather /h/, while men used both /s/ and /h/, also interchangeably (4). [Source: Salles fieldwork, 2014].

(2)   a.   maosai/baosai
clothes
'Clothes'

    b.   moogi/boogi
Moogi
'Moogi'

(3)   a.   gaihi/naihi
DEM
'That'

    b.   gahiao/nahiao
plane
'Plane'

---

[14]Everett states that all the vowels and some of the consonants of Pirahã can be nasal phonetically, but there are no intrinsically nasal vowels or consonants in this language (Everett, 1983:208). Since I am not discussing the phonetic supra segmental level in Pirahã, I am not representing nasality in the orthography I chose.

(4) a. maohai (women/men)/maosai (men)

clothes

'Clothes'

b. ioihoi (women/men)/ioisoi (men)

Ioisoi

'Ioisoi'

According to Keren Everett (1998), Pirahã is tonal, with high and low tone, and stress assignment on syllables independent from tone. Since we did not conduct any analysis of Pirahã tones, we will not represent them graphically in the data collected by me[15]. In the next section, we will establish a typological description of the language.

1.2.1 Typological profile of Pirahã

The verbal morphology in Pirahã is predominantly agglutinative. According to Everett (1986), there are nineteen suffixes, which morphologically manifest aspectual, temporal and adverbial notions. The arguments of the verb, on the other hand, are not marked in verbs. There is no tense marking in Pirahã verbs, being the temporal notion expressed by the combination of aspect, context and adverbs. Aspect is marked through the verbal suffixes *-b* perfective (5), *-p* imperfective (6), *-áo* telic (5), *-ái* atelic (6), *-iig* continuative (6), *-ta* iterative (7) and *-hoag*/*-hói* ingressive (8). [Source: Everett, 1986: 290, 292: (333), (343), (345), (348)].

(5) ti xis ab **-áo -b** -í -haí kaahaixá

*1 ANIMAL catch**-TELIC-PERF**-PROX-RELATIVE CERT macaw*

'I will have caught a macaw'

---

[15] Also important to observe is that we do not represent a complete morphological segmentation of the words in my data, given the shallow understanding of Pirahã morphology. We also maintain Everett's *x* as the orthographic representation for the glottal plosive.

(6)   hi xopáoho **-ái**   **-p**   **-iig** -á

    3   *work -**ATELIC -IMPERF -CONT** -REMOTE*

    'He was/will be working'

(7)   hi kohoi **-tá** -há

    3   *eat **-ITER** -COMPLETE CERT*

    'He is eating again'

(8)   ti soxóá xait -á **-hói**

    1 already sleep-(?)**-INGR**

    'I already am going to sleep'

Everett also describes referential aspects, which are: *-i* proximate (9), *-a* remote (10), *-ab* durative (10) and *-áp* punctiliar (11). [Source: Everett, 1986: 293-294: (351), (356), (358)].

(9)   hi   gáí-sai   xaoói ti kap**-í**   baaí

    *3 say-NOMLZR foreigner 1 shoot**-PROX** wild pig*

    'The foreigner said, "I am going to shoot a wild pig"'

(10)   taoá   oho   **-ab**   **-a**   -áti

    *Taoá   search for **-DUR** **-REMOTE** -UNCERT*

    'Perhaps Taoá will continue searching'

(11)   boitó soxóá   xab -óp **-áp** -á

    *boat already turn-go**-PUNCT**-REMOTE*

    'The boat already arrived'

The author lists other eight different verbal suffixes, among which are: *-sog* desiderative[16], *-áti* uncertainty, *-haí* relative certainty, *-há* complete certainty, *-sai* nominalizer.

The canonical constituent order in Pirahã is Subject-Object-Verb (SOV), as

---

[16]But see Rodrigues et al (forthcoming) for evidence against analyzing *-sog* as a verbal suffix.

in (12)-(13). [Source: Salles fieldwork, 2014].

(12)   ti pioahai ogai
         1 guaraná want
         'I want guaraná'

(13)   ti pioahai ogi-hiaba
         1 guaraná want-NEG
         'I don't want guaraná'

Everett (1986) also claims that the unmarked (basic) word order of a Pirahã clause is SOV. The author considers this to be the basic order, since putting those words in the OSV order without a pause between O and S, the meaning would be 'milk drinks me', while a pause leads to a topic interpretation: 'milk, I drink a lot'. Dryer (2007), while making comments on the ways to classify a language's basic order, describes frequency as one of the techniques commonly employed. Everett seems to have used this principle, since he states that 90% of Pirahã sentences in his transcribed material are SOV.

The fact that Pirahã is an SOV language is revealing in other aspects. Since Greenberg's universals (1966), the typology of word order has been studied as a predictor of word order parameters shared by languages with the same pattern between subject, object and verb (cf. Comrie, 1989; Dryer, 2007).

Take the word order between subject, verb and object within a clause. Although the six logically possible combinations are SOV, SVO, VSO, VOS, OVS and OSV, Comrie (1989) points out that SOV and SVO orders are the most commonly found, followed by VSO. There are solid examples of VOS and OVS languages too, while OSV order is only found in preliminary examples, which Dryer (2007) claims to be little convincing. Dryer classify languages in four different types: verb final, verb initial, SVO languages and the less common OVS and OSV, object initial.

Verb final languages are those whose verbs follow the subject and the object, more specifically, those which are SOV. As mentioned before, it is interesting to investigate the word order patterns shared by SOV languages. Dryer observes that verb final languages tend to present the following word orders: Manner Adverb >

Verb; Postpositions; Genitive[17] > Noun; within comparatives, Comparative pattern > Comparative marker > Adjective; and adverbial Subordinators following the Subordinate clause. Let us review each case in Pirahã, according to Everett's description of the language.

1.2.1.1 Manner adverbs

According to Everett (1986), adverbial notions are normally expressed by verbal affixes in Pirahã. The author points out that the same words used to modify names, also modify verbs, that is, there would be a class of general modifiers instead of adjectives and adverbs separately (notice the usage of *báíhiigí* as an adverb in (14) and an adjective in (16)). Even so, he dedicates one section of his work to the structure of adverbial phrase, where we find the examples below [Source: Everett, 1986: 302-303: (392)-(393), (396a)].

(14)   kaioá hi   báíhiigí   xis
        Kaioá 3SG   slow   animal
        ibóít-ai-p-á-há kabattií
        cut-ATELIC-IMPERF-REMOTE-COMPL.CERT
        'Kaioá was cutting the tapir slowly'

(15)   hi   xaibogi   xaháp-i   hoasaisi
        3sg   fast   go-PROX   nambu (bird species)
        'The *nambu* left rapidly'

(16)   boitóhoi báíhiigí
        boat slow
        '(a) slow boat'

We observe that these modifiers, which work as manner adverbs, precede the verb, following the order Adv>V, typical pattern in SOV languages.

---

[17]Genitive refers to possessors in possessive noun phrases and N to the possessum.

1.2.1.2 Postpositions

As expected for an SOV language, Pirahã makes use of postpositions (17)-(18). [Source: Salles fieldwork, 2014].

(17)   kapigaitoi tabo **apo**
      pencil   table on
    'The pencil is on the table'

(18)   kapigaitoi ti **igihio**
      pencil   1 near
    'The pencil is near me'

1.2.1.3 Genitives (or Possessives)

Another SOV word order pattern observed in Pirahã is the order GN (or possessor>possessum) within noun phrases. This order is shown in (19) and discussed in detail in chapter 3. [Source: Salles fieldwork, 2014].

(19)   ti ibaisi
    1 spouse
    'My wife'

1.2.1.4 Comparative constructions

Under Everett's analysis, comparison is expressed through parataxis in Pirahã. Actually, the author explains that the tests he tried to use to elicit comparatives were frustrated, resulting in the paratactic structures in example (20) [Source: Everett, 1986: 221: (87)].

(20)   kapíígaxiítoii xogií   gáihi   kapíígaxiítoii koíhi   gáihi
    pencil   *big*   DEM *pencil*   *small* DEM
    'That pencil is big; that pencil is small'

Even in situations in which Pirahãs borrow the Portuguese word *mais* 'more', the structure, according to Everett, is juxtaposition of sentences, without any comparative marker. See (21) [Source: Everett, 1986: 223: (96)].

(21)   batío   pága   póoko xoogiái hi mais paga   bíi
              Martinho pay.PRS little Xoogiái 3 more pay.PRS well
              'Martinho pays little; Xoogiái he pays better'

Example (22) [Source: Everett, 1986: 221: (85)] uses the comparative marker *xigiábií*, analyzed as a verbal complement which plays the same comparative role as 'as' or 'similar/like'.

(22)   giopaí gáihi   kapióxio xigiábií
              dog   DEM   other   like
              'That dog is like the other'

If we want to account for the word order pattern within comparative structures in Pirahã, (22) does not help us because, even though *xigiábií* may be considered a comparative marker, no adjectives are involved in the construction. However, if we look at (21) and consider the comparative pattern *xoogiái* 'Xoogiái' and the expression *mais paga bíi* 'pays better' as a quality of 'Xoogiái' compared to 'Martinho', the pattern, then, would be comparative pattern > adjective. Important to notice is that, according to Dryer's (2007) typology, head initial languages follow the pattern adjective>comparative marker>comparative pattern, while head final languages follow the opposite order. Thus, the position of the comparative marker is irrelevant to differentiate those two groups of languages. Therefore, the order presented in Pirahã comparative structures is the expected word order for an SOV language.

1.2.1.5 Adverbial subordinators

Now, let us see how adverbial subordinators behave in Pirahã subordinate clauses. Everett (1986) describes the suffix *-saí* as a conditional suffix, analogous to 'if', in English. See (23) [Source: Everett, 1986: 264: (239)].

(23)   *pii   boi-hiab-i-saí   ti   ahá-p-i-í*

      water   come-NEG-EP-COND   1sg   go-IMPER-PROX-COMPL.CERT

      'If it doesn't rain, I'll go'

If we are looking for the subordinators' position, we can say it is final in relation to the subordinate clause, since it is suffixed to the verb, which follows the subject. Therefore, the order is Clause > Subordinator, exactly as expected for an SOV language.

In this section, we gave a general account of word order typology in Pirahã. In chapter 3, we will present some syntactic aspects of nominal expressions in Pirahã, as well as discuss copular verbs, and constructions involving existentials and topicalization in Pirahã. A minimum understanding of these three types of structure is important for us in order to understand recursion within the nominal domain in this language.

As a final observation, let me just remark that this thesis does not give a thorough analysis of Pirahã grammar, rather focusing on recursion and possessive constructions.

1.3 Recursion and the Pirahã debate

In order to grasp recursion as understood by Hauser, Chomsky & Fitch (2002), it is necessary to revisit this notion within mathematics first. This is done in our chapter 2. Here, we present a summary of this chapter.

Recursive functions are used in number theory (arithmetic) to study the infinity of the set of natural numbers. Grassmann (1861) noticed an inductive property of natural numbers, that is, from zero, all the other numbers may originate through the repetition of the operation *add* one.

This observation led Peano (1889) to postulate a set of axioms to define the infinite set of natural numbers. The fifth axiom, usually called the *Axiom of Induction*, captures Grassmann's inductive property, since it predicts that the successor of a number is the number itself plus one ($x' = x + 1$). The infinity of the set of natural numbers is guaranteed, then, by the recursive operation *add one*, which generates each and every successor of a number in this set.

The path of recursion from mathematics to the analysis of linguistic structure was logical, since infinity was already attested in human language. In 1953, Bar-Hillel encourages the axiomatization method to analyze language and a few years later, Chomsky (1956) claims that a linguistic theory is adequate if and only if it provides a finite set of formal steps (axioms) to generate all and only those sentences of a given language. With that in mind, he proposes the 1950's transformational grammar model, where the phrase structure rewriting rules provide a set of finite axioms, which are able to produce an infinite number of kernel sentences. These sentences may, then, be modified by transformational rules (e.g. the derivation of passives from the transformation of their active counterpart).

Forty years later, Chomsky (1995) proposes a program which investigates the mechanism behind recursion in human language. The Minimalist Program understands language as composed by a dynamic combinatorial system, according to which lexical items are combined through the concatenating operation Merge. Since Merge may apply again over its own outputs, it is a recursive combinatorial mechanism.

Using van der Hulst's (2010) terminology, one can define Merge as general recursion. However, since Merge does not consider the labels of the objects it is concatenating, a specific kind of recursion, self-embedding, may result from concatenating two tokens of the same syntactic type. Self-embedding, thus, is a possible product of the combinatorial system of language. Having said so, we can define self-embedding as possible structural representations. This is what linguists and non-linguists have frequently misunderstood, taking a specific kind of recursion, self-embedding, for general recursion, Merge.

One example of this misinterpretation of recursion is Everett (2005), who wrongly claims that lack of self-embedding in Pirahã is a massive counter evidence to the universality of recursion in human language.

Everett purportedly presents several examples showing absence of self-embedding, in nominal phrases, clauses, as well as the absence of any kind of subordinating functional elements, such as complementizers. His data is given in (24)-(25). [Source: Everett, 2005: 624,628: (24), (35)].

(24) ti   gái-sai   kó'oí   hi   kaháp-ií
      1SG speak-NR   Kó'oí   3SG leave-INTENTION
      'I said (that) Kó'oí intends to leave'
      (lit. 'My saying Kó'oí leave-intend')

(25) *kó'oí   hoagí   kai   gáihií   'íga
       Kó'oí   son   daughter   DEM   true
      'That is Kó'oí's son's daughter'

In (24), the sentence following *ti gáisai* "my saying" is analyzed by Everett as a juxtaposed sentence which, although interpreted as the complement of *gáisai*, does not involve explicit embedding, being only a paratactic conjunction. The unavailability of data like (25), is, according to Everett, evidence for absence of self-embedding within nominal expressions.

Everett associates this lack of self-embedding to Pirahã cultural facts. For example, since all Pirahãs know each other, it is argued that it is unnecessary to use more than one level of embedding to identify someone in this language. This, according to Everett's line of reasoning, would explain the unavailability of data like (25). Even though in some contexts, such as talking about foreign families, unknown to Pirahãs, extra information may be needed to identify someone, there is no extra level of embedding, but rather the juxtaposition of one more clause (26) [Source: Everett, 2005: 628: (37)].

(26) 'ísaabi   kai   gáihií   'íga   kó'oí   hoagi   'aisigí-aí
      Ísaabi   daughter   DEM   true   Kó'oí   son   the same-be
      'That is Ísaabi's daughter. Kó'oí's son being the same'
      (That is Ísaabi's daughter and Ísaabi is Kó'oí's son)[18].

The author under consideration does not explain to us how in (26) the proper noun Ísaabi is recovered in the interpretation of the second sentence. He just says it is. This lack of detailed analysis suggests that a more fine-grained study of self-embedding within nominal phrases (as well as within other domains) is necessary

---

[18]See Everett (2005: 628).

in Pirahã. This is what we do in chapter 3, showing with possessive constructions that self-embedding is not only possible but also productive in Pirahã.

The main unwarranted conclusion of Everett's, is that culture constrains the Pirahã grammar through the *Principle of Immediacy of Information Encoding*. Information beyond that which can be obtained through immediate experience is not encoded within the same sentence. This is taken to be an argument against the assumption that recursion is universal and also against the notion of Universal Grammar, one of the car chiefs of the Principle & Parameters theory (Chomsky 1981; 1986). This clearly impacts modern studies of human language, and Everett's paper caused reactions not only in the academy, but also inflamed the mainstream media[19].

This repercussion led Nevins, Pesetsky and Rodrigues (2009a) to reassess Everett (2005). They clarify that recursion is defined by the operation Merge and point to the weakness of Everett's cultural principle to deal with his own Pirahã data. Moreover, the authors find many examples of self-embedding in previous work done by Everett (e.g. Everett, 1986).

Although the lack of self-embedding in a language such as Pirahã or any other is not evidence against Merge as the recursive combinatorial operation behind human language, the investigation of what could constrain this operation as to disallow self-embedding representations is still meaningful. First, however, it is necessary to sift the Pirahã data, verifying if self-embedding is indeed impossible in this language. This is clearly not the case, as shown in chapter 3.

1.4 Contribution and organization of this thesis

Given everything we said here, the contribution of this thesis is twofold. First, it gives us a better understanding of recursion, allowing for the correct comprehension of the design of the Faculty of Language proposed by Hauser, Chomsky & Fitch (2002). Although controversial with respect to considerations about the evolution of language (cf. Pinker & Jackendoff, 2005a&b; Fitch, Hauser

---

[19]Examples are articles published in *The New York Times* and *The Guardian*. Links are, respectively: http://www.nytimes.com/2012/03/22/books/a-new-book-and-film-about-rare-amazonian-language.html?pagewanted=all&_r=0 and http://www.theguardian.com/technology/2012/mar/25/daniel-everett-human-language-piraha.

& Chomsky, 2005), the hypothesis that recursion is the human specific capacity distinguishing us from other animals has not been denied by comparative studies (see 2.3.1). Moreover, by clarifying the misunderstanding involving general and specific recursion within linguistics, we are left with no reasons to believe Pirahã is a counter evidence to the universal status of recursion in language.

Second, by observing evidence of specific recursion (or self-embedding) within Pirahã possessives, we assure that no ban on Merge is at work in this language, at least in the nominal domain. The data resulting from our fieldwork provides the literature on recursive representations with more possibilities for cross-linguistic comparative analysis. Furthermore, the issues we discuss in chapter 3 contribute to a wider understanding of the structure of determiner phrases in Pirahã and open new research demands for future works.

Besides the theoretical contributions of this thesis, it has also a positive impact to Pirahã society. The more linguistic knowledge of their language is produced, the easier it will be in the near future to contribute with projects such as dictionaries and pedagogical grammars. Since they are being alphabetized in Portuguese, the construction of a written system for Pirahã together with the speakers is highly significant as a language preservation strategy. The elaboration of pedagogical material will certainly rely on academic research conducted in this language.

The thesis is organized as follows. In chapter 2, we present the debate around the capacity of recursion and its status within the study of language and its design. First, we introduce language as understood by Hauser, Chomsky & Fitch (2002) in the proposal that the unique component of language is defined in terms of recursion. Second, we offer an overview of the mathematical definition of recursion and how this concept made its way to formal linguistics. Settling down this issue, we reassess Hauser, Chomsky & Fitch, discussing recursion as being (non)-specificity to humans, and (non)-universal.

In chapter 3, we concentrate on Pirahã possessive noun phrases, beginning with a brief introduction to nominal expressions in the language. The new data is presented involving from one to three levels of embedding. After that, we discuss possible alternative non self-embedding analysis for the data involving covert possessive verbs such as "have" or topicalization within the nominal domain.

The conclusion (chapter 4) summarizes the discussion presented in this work

and points towards future issues related to self-embedding in Pirahã. This involves constructions with the bridge verb *gai-sai* "to say". As we will discuss, in this type of construction, self-embedding seems to be impossible in elders grammar, possibly related to interface constraints involving evidentiality. In youngers' grammar, however, evidence for self-embedding in these constructions were spotted, suggesting grammatical changes in Pirahã, which seems to be leveling off their grammar, removing any interface ban on self-embedding.

2
What is Recursion?

This chapter explores the debate around the property of recursion and its status within the study of the nature of human language (I-language). It is organized in the following way: section 2.1 presents I-language as understood by Hauser, Chomsky & Fitch (2002 - HCF), who proposed that the unique component of language is to be exhaustively defined in terms of recursion. However, as we shall see, the authors did not define recursion properly, starting a debate full of misunderstandings among linguists, biologists and psychologists devoted to the study of language and its nature. Section 2.2 gives us an overview of the mathematical definition of recursion and how this concept has been understood within linguistics. In section 2.3, we reassess HCF and discuss recursion in two different spheres: (1 - section 2.3.1) recursion as (non)-specific to humans; (2 - section 2.3.2) recursion as (non)-universal: the Pirahã debate.

2.1 Discrete infinity and the Faculty of Human Language

HCF (2002) proposed an interdisciplinary investigation on the nature and the architecture of I-language [1] trying to bring together biology, anthropology, psychology and neuroscience. Clearly, with respect to language there is a lot of variance among human beings, what makes it hard to construct a biolinguistic project. However, as HCF's allegory of a Martian naturalist visiting our planet suggests, there are certain aspects of internal language that are undeniably universal. One of these universals is the existence of a combinatorial system, which is able to create an infinite number of complex linguistic objects out of a finite number of lexical items. Given the combinatorial power of human language, the Martian may conclude that human language is significantly different from other coexistent communicative systems, being perhaps "organized like the genetic code - hierarchical, generative, recursive, and virtually limitless with respect to its scope

---

[1] Based on Chomsky's previous work, HCF define I-language as a component of the mind/brain responsible for grammatical computations and its interfaces with other related cognitive modules. They assume that I-language is the primary object of interest for the study of language evolution and function of the language faculty.

of expression." (HCF: 1569).

Before HCF, other humans have already noticed the capacity of our language to give infinite expressive power to a finite set of sources. Galilei's book on the dialogue about the two chief world systems (Copernican vs. Ptolemaic) (1632) is the first reference on this property of language. Through the figure of Sagredo, one of the interlocutors of his dialogue, Galilei reflects upon the inventions the human mind is capable of and observes that one stands among all the others, which is the ability to produce different arrangements out of a finite set of alphabetic letters in order to communicate our thoughts to other people with written language. The grammarians of Port Royal, Arnauld & Lancelot (1660), were also marveled by the combinatorial power of language and went further in their analysis. On our ability to produce speech, they observed that there might be a "spiritual element of speech", which enables us to use a finite variety of sounds to produce an infinite variety of words that communicate our inner thoughts. This creative aspect of language was much later rescued by von Humboldt (1836): a language can "make infinite use of finite means"[2].

After observing the different communicative systems used by different species on Earth, HCF's Martian could actually ask the following question: is there any universal property on language that would make it radically different from other communicative systems? Trying to compare the mental faculties of humans and other animal species, Darwin (1882) approached this issue noticing that many animals are able to express emotions through sounds, but that "the habitual use of articulate language is, however, peculiar to men" (Darwin, 1882: 85). As noticed by Darwin, we can use inarticulate sounds to express emotions just as other animal species do (e.g. baby cries). Some animals also present a fairly good capacity to understand and articulate human sounds (e.g. Rico, the dog (Kaminski et al, 2004), and Alex, the parrot (Pepperberg, 2008)). However, these events may not be connected to a higher intelligence. Connected to this intelligence, is probably our "almost infinitely larger power of associating together the most diversified sounds and ideas" (Darwin, 1882: 86). Therefore, being HCF's Martian a good observer, he would fatally conclude that discrete infinity is the universal property that makes us so unique with respect to language.

---

[2]Chomsky (1965: 8)

Whatever happened in our evolutionary lineage to differ our system from that of birds, other mammals or other primates, it has to do with the computations underlying our cognitive system. Clearly, our ability to make infinite use of finite means poses a logic problem: how can a finite input be mapped into an infinite output? To solve this, we need to posit a combinatorial system as the main mechanism behind language. That is why HCF claim for a division of questions involving language into those concerned with the communicative issues and those concerned with the computational capacities of the system. They make it clear, though, that they are not claiming language not to have communicative purposes, but that there is a possibility that the evolution of these computational capacities has happened for reasons other than communication (such as navigation or numbers), being only afterward proved useful for communication.

In order to separate communicative aspects from computational aspects in the investigation of the evolution of language, HCF propose two senses of FL, "one broader and more inclusive, the other more restricted and narrow." (HCF: 1570). The broader portion is the Faculty of Language - broad sense (FLB), which is the combination of a computational system with a sensory-motor system (regarding the articulation of language) and a conceptual-intentional system (regarding the meanings conveyed by language). They explain that FLB includes the biological capacities peculiar to humans' ability to master languages, but excludes other capacities such as memory and respiration, which are necessary for language, but not enough in themselves. The computational system they refer to is what they call Faculty of Language - narrow sense (FLN), which is "the abstract linguistic computational system alone, independent of other systems with which it interacts and interfaces." (HCF: 1571).

That is to say that FLN contains narrow syntax, a key component which generates linguistic representations that are mapped to interface systems which interact with the sensory-motor and the conceptual-intentional systems. Being the combinatorial system behind grammar, FLN is thus responsible for discrete infinity.

We might at this point ask what the underlying mechanisms of FLN that yield discrete infinity are. HCF do not define these mechanisms, but they say how they operate. For them, in order to achieve discrete infinity, FLN includes at least the capacity of recursion. Nevertheless, what is recursion? HCF do not define recursion neither. However, given the research program established by generative

grammarians since the fifties, it is arguably the case that the FLN mechanism responsible for recursion is combinatorial in nature, which within the Minimalist Program (Chomsky, 1995) is defined as the operation Merge, which takes two syntactic objects and combine them to form a new syntactic unit. This combinatorial operation is arguably triggered by lexical properties (features) of the lexical items being combined (see Chomsky, 1995, chapter 4). In order to understand the link between recursion and Merge, in what follows we will recover the history of recursion from its origin in mathematics and logic to linguistics.

2.2 MERGE: Understanding recursion as a mathematical procedure

In this section, we are giving a historical overview on how the usage of recursive functions came to have a central role to mathematical theories and how this concept gave the basis to the development of generative grammar theory. We are going from Grassmann to Peano's axioms, from Gödel to Church, Turing and Post, until we get to Chomsky's (1955, 1957) usage of recursively enumerable sets to analyze syntactic phenomena of human languages. After that, we follow some important developments within Generative Grammar until we reach the Minimalist Program and the concept of Merge as the recursive operation of the computational system of human language. In the end, we will review two theories, which discuss the ways in which specific recursive representations can be constrained in language.

In 1861, Hermann Grassmann, a German mathematician and linguist, made an important contribution to number theory[3]. Observing the set of natural numbers, he noticed that zero is the base number from which all the other numbers originate, through the repetition of the operation *add one*. This is a key property of natural numbers, called *inductive property*. Natural numbers can be defined by a function *f (n)* in which *f (0)* is the base and *f (n+1)* (a successor function) is given in terms of *f (n)* already defined. For example, take the sequence in (1), a subset of the natural numbers from 0 to 4. We can define the base 0 as in (2) and the numbers from 1 to 4 as in (3), (4), (5) and (6). Notice that the subsequent number is defined by the operation *add 1* applied over the previous defined number in the sequence.

---

[3] Or Arithmetic.

(1)   {0, 1, 2, 3, 4}

(2) $f(0) = 0$

(3) $f(1) = f(0 + 1) = 1$

(4) $f(2) = f(1 + 1) = 2$

(5) $f(3) = f(2 + 1) = 3$

(6) $f(4) = f(3 + 1) = 4$

As geometry, arithmetic started to be exposed to the axiomatic method and treated as a deductive discipline, i.e. the axioms should be accepted without proof constituting the foundations of the number system, from which theorems are obtained using the principles of logic (Nagel & Newman, 1957). After Grassmann's observations, other mathematicians worked in the axiomatization[4] of arithmetic, such as Peirce (1881), Dedekind (1888) and Peano (1889). Although the essential ideas of Peano's axioms for natural numbers are in Dedekind, the originality of his work is undeniable (see Kennedy, 1973). Peano's nine axioms were basically postulates which together defined the infinite set of natural numbers. The five first axioms are the most important for us here. They are organized from (a) to (e) in (7)[5].

(7)   (a) 1^6 is a natural number;

       (b)   For each *x* there exists exactly one natural number, called the successor of *x*, which can be denoted by *x'*;

       (c)   1 is not the successor of any natural number;

       (d)   For any given number, there is no number whose successor is exactly that given number;

       (e)   If a set contains 1, *x*, and the successors *x'*, then the set contains all the natural numbers.

The fifth axiom in (7 (e)) is usually called the *Axiom of Induction*. The axioms

---

[4]The definition of mathematical systems by a set of propositions assumed to be true so the consequences that follow from them can be studied in theorems. For more, see Axiomatic Method in the *Encyclopedia of Mathematics* URL: http://www.encyclopediaofmath.org/index.php?title=Axiomatic_method&oldid=17770

[5]Using as a reference Landau (1966), since Peano's original presentation of the axioms only uses mathematical notations. To explain every mathematical symbol used would be a task which extrapolates the purposes of this thesis.

[6]In the original formulation, Peano (1889) used *1*, but more recent presentations use *0* as the first natural number (e.g Kleene (1952)).

have to define the set of all natural numbers, but the set of natural numbers is infinite. Then, we need a property that is able to account for this infinite set; this is the inductive property above. The fifth axiom implies that numbers have the inductive property, since $x' = x + 1$. That is, every defined number and its successor will be part of the infinite set of natural numbers[7]. We can infer then that the inductive property of numbers is responsible for infinity in the numerical domain. Take (8) as an illustration of how this property works.

(8)   (a) $x' = x + 1$

      (b) $x' + 1 = (x + 1) + 1 = x + 2 = x''$

      (c) $x'' + 1 = (x + 2) + 1 = x + 3 = x'''$

      (d)   ...

Step (d) in (8) indicates that the property can apply indefinitely. As we can observe, x''' in (8 (c)) is the product obtained by the application of the inductive property over x'', which is, on its turn, the product obtained by the application of this same property over x' in (8 (b)). The property, then, is applying *recursively*, since it uses as inputs its own outputs. Thus, we can say that a finite number of axioms and the recursive application of the inductive property of numbers express the infinity of the set of natural numbers.

Now, let us turn to some developments that happened regarding axiomatic arithmetic that influenced the thinkers who provided the basis for the generative linguistics endeavor. Let us start within *metamathematics*. Any theory can be proved by a *metatheory*, a theory that investigates the properties of another one (see Kleene, 1952). Peano's axioms constitute a number theory, which can be analyzed by theorems that must hold true for all the objects of that system. One might ask how to prove a theorem if we cannot access all the numbers in the infinite set of natural numbers. Considering Peano's number theory, this should be easily answered by the fact that it works with *primitive recursion*, a procedure which uses the value of a previous argument to define the value of its successor (Odifreddi, Piergiorgio & Cooper, 2012)[8]. If a number is defined by the same function that

---

[7]For more, see Kleene (1952).

[8]For more, see Odifreddi, 1989. See also Odifreddi, Piergiorgio and Cooper, S. Barry, "Recursive Functions", *The Stanford Encyclopedia of Philosophy* (Fall 2012 Edition), Edward N. Zalta

defined its predecessor, then a theorem which holds for a number and its successor should be considered proved, "since any case can be derived from the initial instance by repeated applications of the inductive step" (Tomalin, 2006: 63).

Gödel (1931), Turing (1936) and Post (1943; 1947) presented important developments of the notion of mathematical induction. Gödel, made explicit the properties of any recursive function: it had to be finite in nature, and had to be defined in terms of either preceding functions or the successor function. Turing demonstrated how these functions could be implemented in terms of computation (Turing machine), a generative procedure. Post, differently from Turing, saw recursive functions as rewrite production systems, a set of logic independent conditions, analogous to proof conditions. The notion of recursive inductive systems as defined by these authors is the foundation of the Chomskyan notion of Grammar.

In 1953, the linguist and logician Yeoshua Bar-Hillel published a paper in which he defended that recursive definitions should be used within empirical sciences, including linguistics. His claim is for a formal analysis of linguistic phenomena. Using the primitive kind of recursion encountered in Peano (1889) and English as a metalanguage, Bar-Hillel presents a formal analysis of French sentences, suggesting that a sentence can be split into smaller units until the basic constituents are encountered:

(9)   x will be called a sentence (in French) if (and only if) x is a sequence of a nominal and a (intransitive) verbal, or a sequence of a nominal, a (transitive) verbal, and a sentence, or ......, or a sequence of a sentence, the word "et", and a sentence, or ......
(Bar-Hillel, 1953: 163)

This definition can be converted in the following set of rewriting rules:

(10)   S → N V
        S → N VN
        S → S et S

(ed.), URL = <http://plato.stanford.edu/archives/fall2012/entries/recursive-functions/>.

Thus, within a historical perspective we can place Bar-Hillel's work in the following way:

Bar-Hillel's use of recursive definitions to analyze the structure of sentences in natural language can be viewed as one manifestation of this pervasive desire for the mathematisation of syntactic analysis, which became such a characteristic feature of certain kinds of linguistic research in the mid-twentieth century. Significantly, Bar-Hillel's ideas intrigued Chomsky in the early 1950s, ... (Tomalin, 2006: 67)

Chomsky (1955) claims for a general linguistic theory, which in 1956 is defined as a metatheory that is concerned with the problem of how to choose a theory of the structure of a given language based on a finite corpus of sentences. According to the author, a theory (or a grammar of a language) is adequate if it provides a finite set of formal steps to generate all and only those sentences of a given language.

Chomsky (1956) analyzes three possible models of grammar under his metatheory criterion. The models are: (a) finite-state grammar (Markoff processes); (b) phrase-structure grammar; and (c) transformational grammar. These are considered under the three following questions:

(11)   Are there interesting languages [RS - real languages] that are simply outside the range of description of the proposed type [model]?

(12)   Can we construct reasonably simple grammars for all interesting languages?

(13)   Are such grammars "revealing" in the sense that the syntactic structure that they exhibit can support semantic analysis, can provide insight to the use and understanding of language, etc.? (Chomsky, 1956: 114)

If any of the theoretical models above fails to answer (11) negatively, (12) and (13) do not even have to be asked. The model is inadequate.

The first model analyzed by Chomsky is the finite-state grammar, also called

Markoff Process. This process can be defined as in (14).

(14)   a system with a finite number of states $S_0,..., S_q$, a set A = {a_{ijk} | 0 $\leq$ i, j $\leq$ q; 1 $\leq$ k $\leq$ N_{ij} for each i,j} of transition symbols, and a set C = {(S_i, S_j)} of certain pairs of states of G that are said to be <u>connected</u>. As the system moves from state S_i to S_j, it produces a symbol $a_{ijk} \epsilon A$. (Chomsky, 1956: 114).

In other words, a Markoff process produces a set of symbols as the states run from S_0 to S_n. But the transition symbols are produced as the connected pairs (S_i, S_j) of states run. That is to say that a finite-state grammar can only produce new symbols applying a new step over the symbol produced by the last step applied. Taking this process, elementary as it is, as a grammar of a language L, it follows that we will have a finite amount of apparatus giving us an infinite number of sentences. To see this, consider Chomsky's (1957) example:

(15)   The old man came.

If a looping is added in the transition from the second to the third state, we will have (16).

(16)   The old old man came.

Since there is no upper limit on the number of loopings the mechanism can do, then it follows that the sentence can go on and on. This gives us iteration. Hence, a grammar can be infinite in length. In addition, since there is no restriction on the number of states the process can run, it follows that a large amount of symbols can be inserted by the process. Therefore the number of symbols in (16) can be increased:

(17)   The old old man came, the young woman sang, the child cried, the adults applauded...

Real languages, however, including English, have some universal properties that a Markoff process fails to capture. These properties are constituency and dependencies. To understand the issue involving constituency, consider (15) again. Given that the Markoff process gives us a linear sequence of symbols, it will allow us to make any grouping using the symbols in (15). That is, it will give us all the constituents in (15) (e.g. *the old man*, *came, the old man came*), but it will also give us all the non-constituents (e.g. *the old*, *man came*). This follows from the fact that in a linear sequence the condition for grouping is *being linearly adjacent*. To avoid this problem, we need a mechanism that produces hierarchies of symbols instead of mere sequences. That is why Chomsky abandons a Markoff process in favor of a phrase-structure grammar.

As already said, another property of human language is that it involves dependencies among the constituents. To see this consider another example of Chomsky's (1957):

(18)   If S1, then S2.

         (cf. If it rains, then I'll bring an umbrella)

In this example there is a dependence between the symbols *if* and *then*. Notice that there is no limit on the amount of symbols that can be placed in between *if* and *then*.

(19)   If it rains and the sun does not come out or the clouds are too
         dense... then I'll bring an umbrella.

The power of a Markoff process as a grammar is also limited because it cannot accommodate dependencies. Another example of dependence involving relative clauses is given in (20), where a complex sentence is inserted between the subject and the predicate.

(20)   The man, who said that if it rains he will bring an umbrella, is
         arriving today.

The second model analyzed by Chomsky is the Phrase Structure Grammar. A

sentence is not just a linear grouping of words from left to right. It is divided into smaller constituents, such as noun phrases (NP), verb phrases (VP), prepositional phrases (PP), etc., which are hierarchically organized as in (21).

(21)   [[the man]$_{NP}$ [[took]$_{Verb}$ [the book]$_{NP}$]$_{VP}$]$_{Sentence}$

A phrase-structure grammar captures this notion of constituency. It is defined in (22).

(22)   A phrase-structure grammar is defined by a finite vocabulary (alphabet) VP, a finite set Σ of initial strings in VP, and a finite set F of rules of the form: X → Y, where X and Y are strings in VP. Each such rule is interpreted as the instruction: rewrite X as Y.
(Chomsky, 1956: 117)

In the terms of Peano's arithmetic, Σ is the initial symbol S, an abstraction which would correspond to zero (or one). The rewriting rules F are the finite set of recursive axioms that will map an S into the set of sentences of the language L. Starting with the initial symbol S, the rewriting rules allow subsequent insertions of phrases in the derivation until the terminal strings (words from the language's VP) are inserted. A phrase-structure grammar is context-free when the application of a rewriting rule does not take into consideration the structural environment in which it takes place[9].

Although the phrase-structure grammar captures constituency, it fails to account for dependency. To generate sentences with auxiliaries, for instance, this grammar would have to include too complex rules. Consider (23), for instance.

(23)   The man has taken the book.

Under this model, the derivation of (23) would have to include rewriting rules

---

[9]Context-sensitive phrase-structure rules are phonological rules of the type proposed by Chomsky and Halle (1968). In Portuguese for example, we have the following rewriting rule: /s/ → /z/ / v___v (v = vowels). As the rule makes it explicit, the rewriting process occurs only in a structural environment of /s/ being surrounded by vowels.

of the types in (24).

(24)   a.   Verb → Auxiliary V

          b.   V → take

          c.   Auxiliary → Present (have) take + en

As Chomsky observes, a rule such as (24c) requires knowledge of the history of the derivation (have, for instance, will have to be rewritten as has because the subject is third person singular). Considering passives (25), he observes that a phrase-structure grammar would not be able to deal with the rearrangement of the constituents of the sentence, under the assumption that passives are structurally related to its active counterpart, especially because the subject of the passive is the underlying object of its active counterpart.

(25)   a.   The girl read the book.

          b.   The book was read by the girl.

It seems that, although more productive than a finite-state grammar because it captures the notion of constituency in sentences, a phrase-structure grammar still does not work because it is unable to capture the fact that a given position in a structure (subject position in (25b)) may be in a structural dependency with another position (the object position in (25b)).

To solve this problem, Chomsky proposes the supplementation of the phrase-structure system with *transformational rules*. He defines grammatical transformations as in (26).

(26)   Each grammatical transformation T will essentially be a rule that converts every sentence with a given constituent structure into a new sentence with a derived constituent structure. (Chomsky, 1956: 121).

In this so-called transformational grammar, the rewriting rules of the phrase-structure component generate kernel sentences, which serve as the input for the transformational rules. The transformational rules are able to look at the derivation

and rearrange the strings of a sentence. These are rules that apply over the product of other rules.

In sum, in Chomsky's 1950's model, recursion is given by the phrase-structure rewriting rules. These rules provide a set of finite axioms, which are able to produce an infinite number of kernel sentences. Importantly for our discussion, in Chomsky (1955), complex sentences such as (27a), involving self-embedding, were not generated by the rewriting rules. The rewriting rules would generate the kernel sentences in (27b,c) which would then be put together by a transformational rule inserting (27c) as the complement of (27b).

(27)   a.   John knew that Mary understood the theory.
       b.   John knew it.
       c.   Mary understood the theory.

The idea that a grammar is a finite procedure that is able to generate the sentences of a language L and only those sentences brings us back to the recursive function of *the successor of* in mathematics. The rewriting rules in the 1950's model are a set of axioms that maps an S into the set of sentences of L. Hence, the set of sentences of L is an enumerable set generated by those axioms. In order to know if a new S is a member of the set of sentences of L, all you have to do is to ask yourself if that S is derived through that given set of rules. In other words, as well discussed in Watumull et al. (2014), in order to know the extension of a given recursive set of rules, all you need to do is to look at its intension.

Chomsky seeks for explanatory power in the models of grammar he analyzes and his conclusion is that, as a general theory of language needs to provide grammars that will account for the infinite number of sentences in a natural language, these grammars need recursive steps to produce infinity without being extremely complex. We can conclude then that *recursion is the means by which the grammar of a natural language reaches discrete infinity*. The only model for description of language that Chomsky found explanatory adequate was that of transformational grammars.

The 1950's model failed, however, because it was not transparent with respect to the universals observed in language. The rewriting rules as well as the transformational rules were language specific. In addition, generalizations with

respect to the application of certain transformational rules (e.g. passive and raising) were not captured by this model, given that each of these constructions were generated by an independent transformational rule.

By 1965, the model proposed by Chomsky was basically organized as follows:

(28)   base → DS → Transformations → SS

The base is composed by the categorical component and the lexicon, which does the same as the set of rewriting rules in the 1950's model, giving us kernel sentences. Those kernel sentences are the DS, the Deep Structure of the derivation of a sentence. DS is the input for the transformational rules of that grammar, giving us the Superficial Structure, the SS in (28). DS was taken to be representations directly feeding the semantic component where meaning was composed. SS, on the other hand, delivered representation for the phonological component, which would be responsible for the vocalization of the utterances.

The problem with this model lies in the direct link between DS and the semantic component. This link is short in its empirical coverage. To see this consider passive structures. Remember how the derivation of passives like (25) was treated in the 1950's, assuming that they were structurally (and semantically) related to their active counterpart, being (25b) the passive counterpart of (25a)? This did not change in 1965. However, this semantic relationship does not hold entirely true for cases like (29) in which the argument of the verb are quantified expressions.

(29)   a.   Everybody loves someone.

         b.   Someone is loved by everybody.

(29b) is not semantically equivalent to (29a). When we say (29a), we mean that there is a set A of lovers and a set B of loved ones. For each lover in A there is a loved one in B. However, for (29b) the interpretation is that there is a unitary set A with a sole person and a plural set B of people who all love that one member of A. Thus, if we take (29a) as the DS or the kernel sentence of (29b) and if we assume that DS feeds the semantic component directly, then the difference in meaning between (29a) and (29b) is unexpected.

To solve this problem within the theory, Chomsky (1970a) proposes a model

in which the transformations still do not alter the meaning of the sentences, but interpretive rules can operate both in DS and SS. It means that the semantic component can assess both DS and SS. Thus, the difference in meaning between (29a) and (29b) is due to the position of the quantifier after the transformation. That is, at SS. Hence, SS was taken to contribute to the composition of the meaning. This is the Extended Standard Theory.

Another turning point for generative theory is the X-bar theory. Chomsky (1970b) (and eleven years later Stowell (1981)) proposed that all categories project in accordance with the so-called X-bar scheme.

(30) XP
     ／＼
   Spec X'
      ／＼
     X Complement
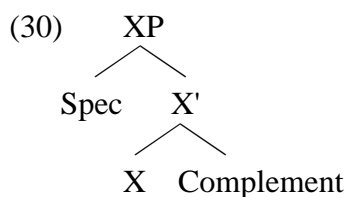
Being X any possible lexical category, the idea is that the head combines with its complement to form a unit (X-bar), which in turn combines with another category (specifier) to form the XP. The X-bar scheme was taken to be a frame available in the grammar to form DS representations. This scheme was part of the computation of a sentence in the sense that it provided a frame for any category drawn from the lexicon. Notice that it is a rigid scheme. The lexical properties of the projecting head did not define any of the parts of its maximal projection.

Given the end of the rewriting rules, in the1970's/1980's model, recursion is not captured by a set of axioms anymore, being rather captured by the way categories are combined to form a sentence. That is, the fact that projections can be combined in an iterative way gives us recursion, and consequently, infinity. Hence, X-bar theory may be less transparent with respect to the mechanism behind recursion in human language, but it still accounts for recursion, as it is constituted by a finite set of frames available for the combinatorial engine of Grammar.

Looking back at the rewriting rules, it was easy to see recursion as a derivational process that takes the output of its previous application as its input. That is, recursion could be easily defined as a function that at step 1 takes x as its argument and y as its value, then on step 2 takes y as its argument and z as its value. In the 1980's we have a combinatorial system operating without the aid of a set of

axioms. Nevertheless, this combinatorial engine is still recursive as it operates upon the result of previous applications of itself. Take the scheme in (31) as an example. *x* combines with α to form an X-bar unit. Then the X-bar unit thus formed combines with β to form XP. This can go on and on in an iterative way.

(31)

```
          XP
         /  \
        β    X'
            /  \
           x    α
```
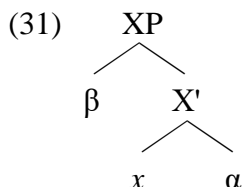
It is out of the scope of this chapter to make any more detailed comment on the Extended Standard Theory, Government and Binding Theory in the 1980's. The main big point in the 1980's was probably the introduction of Principles and Parameters, which made questions about cross-linguistic variation and about acquisition of language more interesting and easy to deal with. This is unfortunately, also out of the scope of this work.

The Minimalist Program (MP - Chomsky, 1995) understands the Grammar as composed by a dynamic combinatorial system (computational system), according to which lexical items drawn from the lexicon project based on its formal features. Therefore, there are no rewriting rules or X-bar scheme guiding the projection of lexical items within syntax. The operation responsible for combining lexical items is <u>Merge</u>. This operation consists of concatenating two independent objects and labeling the new object thus formed. Merge works in a recursive way. To see this, consider the abstract example in (32). First a Lexical Array (Chomsky, 1993) (or a Numeration (Chomsky, 1995)) is selected from the lexicon. Then, the computational system starts to work concatenating α and β. At this step of the derivation the object {α, β} is labeled K. After that, K merges with γ, forming the object labeled M. According to MP, the concatenations in (32b&c) are demanded by the formal features of the lexical items available in (32a).

(32) a. LA = {α, β, γ}

    b. {K{α, β}}

    c. {M {γ {K{α, β}}}}

Crucial to our discussion are the derivational steps in (32). The object K, the output of Merge in (32b), is used as the input for Merge in (32c). Thus, Merge is an operation that calls for itself throughout the derivation.

In this research program there are four syntactic operations: Select, Merge, Move and Agree (cf. Chomsky, 2000). Select accesses the lexicon forming at once a numeration. Then, it selects each item of the numeration placing it into the working space. Merge then starts to work as shown in (32). Agree, as described in (33), opens a syntactic dependency between γ and δ[10]. This operation takes place so that a probe δ in (33) can value some feature of a goal γ in (33).

(33)
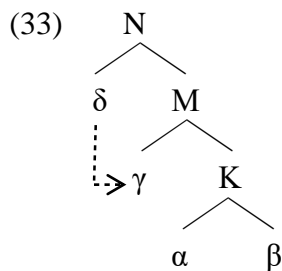
Move, as shown in (34), is a combination of copy and Merge. After Agree has taken place, γ might be copied and remerged as the specifier of δ forming the object labeled O.

(34)

---
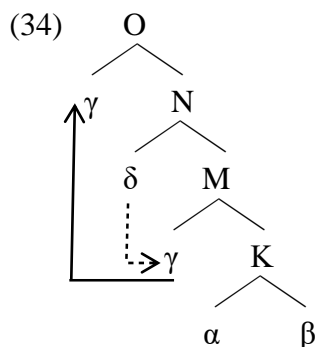
[10]I am using dotted lines to describe Agree and full lines to describe Move.

At this point of our discussion, it might be useful to make two observations: first, differently from previous theoretical models, in the MP, there is no internal level of representation to filter out derivations, and building a structure and transforming it intermingles. This gives us a dynamic system. Second, Merge does not consider the labels of the objects it is concatenating. It takes into consideration only the formal features of these objects. This second observation implies that nothing within the computational system itself prevents two tokens of the same type to be i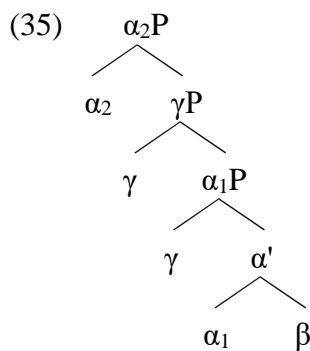mmediately or non-immediately concatenated. That amounts to say that a derivation involving self-embedding (35) is possible.

(35) $\alpha_2 P$
α_2 γP
γ $\alpha_1 P$
γ α'
α_1 β

As discussed in Van der Hulst (2010), linguists have considered two types of recursion: general and specific. By general recursion he means the process through which a sentence is built, i.e. smaller syntactic objects must be recursively combined until the sentence is formed. This is what the operation Merge does. Specific recursion, on the other hand, consists of concatenating two tokens of the same type, one within the other, as in (35). That is frequently called self-embedding. As we will see later in this chapter (section 2.3.2), linguists have often mistaken recursion by self-embedding.

Now we are reaching the main point of this thesis. Everett (2005) presents us with a language that supposedly bans self-embedding of any category. This brings us to an important observation: if syntax is autonomous and independent of other systems (such as semantics) (Chomsky, 1957), why are specific kinds of Merge being impaired? In other words, if Merge does not take into account the labels of the syntactic objects formed, how can it decide which kind of categories it can concatenate or not?

Linguists have tried to understand the syntax of self-embedding structures and why, not only in Pirahã, but also in many other languages (cf. Nevins, Pesetsky & Rodrigues, 2009a), certain kinds of specific recursion are impaired. For instance, Roeper (2011) investigates the constraints on these structures based on language acquisition experimental data involving self-embedding structures. Children need to be exposed somehow to instances of these structures so they can generalize and acquire the grammar of the language they are being exposed to.

Notice that it is not to say that Roeper consider Merge as a property that needs to be acquired, instead of being part of the human biological apparatus. He understands Merge as the "completely universal form of recursion" (Roeper, 2011: 114), but he investigates specific recursion in two different types of representations, namely: Direct Merge and Indirect Merge[11].

Direct Merge (36) would deliver a structure with a conjunctive reading (37), where the order between the repeated constituents is not fixed. [Source: Roeper, 2011: 116: (2)-(3)]. To put it in abstract terms, a structure resulting from Direct Merge would not be one as in (35), but the one in (38) below.

(36)   Direct Recursion: $X \rightarrow Y (X)$

                                      $NP \rightarrow NP ((and) NP)$

(37)   John, Bill, Fred, and Susan arrived.

(38)

$$\begin{array}{c} \alpha_2 P \\ \diagup \diagdown \\ \alpha_2 \quad \alpha_1 P \\ \diagup \diagdown \\ \alpha_1 \quad \beta \end{array}$$

Roeper claims Direct Merge and the resulting conjunctive reading to be the acquisition default. Putting it in other words, there would be no reason to believe a given language lacks Direct Merge and the resulting conjunctive readable structures. The recursive (self-embedding) reading would be a further step in the acquisition process, involving Indirect Merge.

Indirect Merge (39) [Source: Roeper, 2011: 117: (10)] does not allow for the

---

[11]Throughout the text, the author may also use the terms Direct Recursion and Indirect Recursion.

interchangeability among elements Direct Merge does. For instance, take (40a&b), which relate to totally different individuals, the different configuration of the elements causes difference in meaning.

(39)   Indirect Recursion: DP → (Determiner) NP
                                Determiner → {ARTicle POSSessive}
                                POSS → DP 's

(40)   a.   John's friend's father
        b.   John's father's friend

The structures in (39&40) are similar to that abstract structure represented in (35). There is self-embedding within the nominal domain (DPs within DPs), but another category is present in between the steps, thus this type of Merge is called Indirect. Roeper takes indirect recursion to be the locus of cross-linguistic variation. That is, while direct recursion is universally attested, specific types of indirect recursion may not be licensed within a given language.

For the study of Grammar in general it is important to know why Indirect Merge may be blocked in a given language. Assuming the recent minimalist concept of Phases (Chomsky, 2000; 2001; 2005), Roeper tries to provide an answer to this. Chomsky's recent proposal assumes that a derivation develops cyclically, as the operation *spell-out* applies more than once during the derivation, sending chunks of the structure to the interfaces. Roughly speaking, if we take (35) (repeated here as (41)), $\alpha_1 P$ would be a Phase delivered to the interfaces, then $\alpha_2 P$ would be another Phase. That amounts to say that the constraint on indirect recursive constructions such as (40) in certain languages might not be within the combinatorial system in itself, being rather on the interfaces. After phases are spelled-out to the interfaces, they are combined under the constraints imposed by the interface systems. That is, the computational system is autonomous with respect to the application of Merge, although representations can be parametrically filtered out on the interfaces[12].

---

[12]Maia et al (forthcoming), for example, argue, based on their work on prepositional phrases in Karajá, a Macro-Jê language spoken in Brazil, that structures with self-embedding might result in parsing difficulties.

(41)

$$\alpha_2 P$$

$$\alpha_2 \quad \gamma P$$

$$\gamma \quad \alpha_1 P$$

$$\gamma \quad \alpha'$$

$$\alpha_1 \quad \beta$$

Hinzen (2006, 2014), on the other hand, argues for non-autonomy of syntax. Constraints imposed on Merge would be internal to the combinatorial system, not attributed to external interface systems. Using the author's own words:

> On this model there is only one computational system and it generates all the distinctions we need for the use of language in thought: human thought, insofar as it is distinctive from non-human thought, is *inherently* linked to the use of linguistic expressions in grammatical structures. Constraints on recursion in the linguistic domain now need to follow from the workings of this one system itself: they can't be externally imposed. (Hinzen, 2014: 114).

Nevertheless, Hinzen's explanation for the constraints on Merge is not totally incompatible with Roeper's (2011) proposal. Phases are syntactic objects that correspond to propositions at LF. The concept of what constitutes a proposition can be traced back to the ancient Greek philosophers. According to Aristotle's *De Interpretatione*, only the combination (*sýntesis*) and separation (*diaíresis*) of symbols (or concepts) can establish a proposition, which can be judged for truth (when the concepts are combined) or falsity (when the concepts are separated)[13]. A proposition is, then, a relationship between symbols, forming an evaluable unit[14]. Hinzen (2014) then observes that the 'combination and separation' task (i.e. the generation of a proposition) is performed by syntax, which creates the relations

---

[13]If the concepts are combined, they correspond to an affirmative; if they are separated, they correspond to a negation.

[14]For more on the linguistic theory in Aristotle, see Neves (1981; 2002).

established in propositions. If propositions are truth-evaluable (a semantic concept), then it seems plausible to suppose that syntactic and semantic configurations are established in a sole system. This proposal is not conceptually that far from a cyclic system in which propositional syntactic objects (Phases) are constantly being delivered to semantic evaluation. Both systems are cyclic and the cycles are based on the propositionality of the chunks of symbols, the main difference being that for Roeper (following Chomsky, 2000; 2001; 2005) there are different systems involved in the derivation, while Hinzen proposes a sole system, which performs syntax and semantic evaluation.

Hinzen is interested in clausal embedding, and according to him, what limits this type of recursion/merge is the propositionality of sentences. Take, for example (42) [Source: Hinzen, 2014: 120: (14b)].

(42)   *[CP John thought [SC_2 Joan happy [SC_1 Mary sad]]

A small clause cannot be embedded within another small clause because they do not form a truth-evaluable proposition. How about self-embedding of categories which do not form propositions, say DPs or PPs? Hinzen recognizes that there is recursion where truth-evaluability does not arise, such as in example (43) [Source: Hinzen, 2014: 120: (16)].

(43)   the vase on the table in the room in the country house in France

The author explains that in these domains, recursion is working at the level of referentially evaluable objects, which work analogously to the truth-evaluability of clauses. He states that:

> Recursions of the latter sort are not self-sufficient, however, and will ultimately be encompassed by that of clauses. Clausal recursion furthermore appears to be the recursive phenomenon *par excellence*, commonly claimed to the present in all languages (which precisely makes Everett, 2005, controversial in the way it has been), while recursion in NPs appears to be subject to deeper cross-linguistic variation (Hollebrandse and Roeper, 2008). (Hinzen, 2014: 120).

In other words, recursive DPs or PPs are possible if the objects are referentially evaluable, but they only make sense, i.e. are truth-evaluable, if they

are part of a clause. Because referentially-evaluable objects are not propositions in themselves, recursive Merge in the referred domains suffers more cross-linguistic variation. What is interesting to notice is that Hinzen's model also involve cyclicity, the same concept of derivation by Phases.

> Rather than being 'free', recursion waits until a relevant form of structural completeness is reached (or sufficiently approximated) that has a substantive semantic identity: put differently, it is *cyclic*, with cycles understood as units of referentiality. (Hinzen, 2014: 120).

By recursion in the citation above, Hinzen means specific recursion (i.e. self-embedding), not Merge. Syntax forms the units of referentiality through the recursive operation Merge and these units may be embedded within each other forming specific recursive representations. Importantly, he rules out the possibility that languages have representations such as (44) (cf. Arsenijevic & Hinzen, 2012), direct embedding of X-within-X. It is implicit in Hinzen's explanation that a cyclic derivation is one which has intermediating categories between the recursion of the units being formed, that is, cyclic recursion delivers Indirect Recursion in the sense of Roeper (2011)[15]. Then, (45) is an available example in languages, because, unlike (44) it has intermediate categories between the occurrences of C.
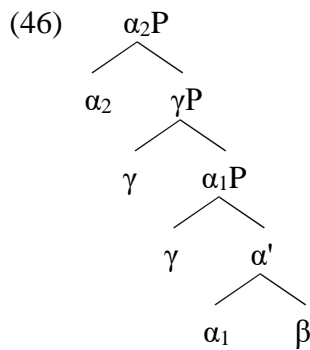
(44)   * [C[C[C]]]...]

(45)   [C-T-V [C-T-V... [C-T-V]]]

To sum up, for self-embedding to take place, a new cycle, or a new Phase, has to be started, and a category $\gamma \neq \alpha$ would intermediate the embedding of α-within-α, just as (35&41), repeated here as (46), shows.

---

[15]For a similar restriction on recursion/self-embedding, see Dékány & den Dikken (2015), who argue that self-embedding of a category is possible when the two instances of the category are separated by a phase head.

(46)

$$
\begin{array}{c}
\alpha_2 P \\
\diagup\diagdown \\
\alpha_2 \quad \gamma P \\
\diagup\diagdown \\
\gamma \quad \alpha_1 P \\
\diagup\diagdown \\
\gamma \quad \alpha' \\
\diagup\diagdown \\
\alpha_1 \quad \beta
\end{array}
$$

As a recent matter of research, the ways specific recursion can be constrained are still unclear and the analysis proposed here are just illustrative of the different approaches that can be taken into account when analyzing the specific recursion (or their absence) cross-linguistically.

At any rate, as we have seen so far, discrete infinity is the core property of human language. We saw how mathematics and logic tried to deal with infinite sets and how this concept was introduced in linguistics by Bar-Hillel and revisited by Chomsky. After understanding the developments of generative grammar and how it reached a more dynamic system with the Minimalist Program, we can safely understand recursion as the mechanism behind the combinatorial system. That is, the role of the operation Merge. Thus, Merge is what confers language discrete infinity.

Therefore, when HCF talk about recursion and its mechanisms as the content of FLN (i.e. as the human specific and language specific part of the Faculty of Language), we must understand that they are talking about an operation such as Merge, and its property of being able to apply over its own outputs to generate infinite sentences out of finite resources[16]. The specific recursion, which is self-embedding, is one possible product of such mechanism. In other words, it is natural to assume that HCF understand recursion as a language property related to Merge. Thus, let us now reconsider HCF's paper.

2.3 Revisiting HCF

We have seen in the previous section that recursive mechanisms are necessary

---

[16]This is clearly stated in Chomsky (2007) and is easily inferred from Bolhuis et al's (2014) considerations on the evolution and the nature of language.

so that a computational system is able to achieve discrete infinity. If language is specific to humans, the logical path is to see recursion as uniquely human too. HCF claim that "although many aspects of FLB are shared with other vertebrates, the core recursive aspect of FLN currently appears to lack any analog in animal communication and possibly other domains as well" (HCF: 1571). This is the challenge HCF give to researchers interested in the nature of FLN, that is, if recursion is found in other animal communicative systems, another hypothesis on the design of language should be construed.

When HCF delineated FLB and FLN they separated aspects of language also present in other cognitive domains to be part of FLB and the aspect which is unique to the domain of language to be part of FLB. What made them choose for recursion to be the content of this specialized subset of language, i.e., what are the evidences that recursion is present only in the domain of language?

What we actually see is that HCF point to the recursive aspect of FLN as *possibly* language domain-specific. When they do so, they are claiming for research endeavor to investigate if recursion is present in domains other than communication, such as number, social relationships, or navigation. They claim that it is possible that recursion as a modular system evolved in animals from a domain-specific system (e.g. specific to navigation) to a domain-general one, penetrable to other domains of cognition, which made it possible for humans to apply recursion to solve other problems, such as language. It is not clear, however, whether HCF are claiming for recursion to have made its way back to the condition of a domain-specific system, this time specific to language. The paper concentrates in showing comparative studies with other species, providing evidence to the hypothesis that recursion is a human specific system, but little is said about studies of recursion in other cognitive domains. For a more detailed discussion on domain specificity and issues involving the evolution of language, refer to the debate between Pinker & Jackendoff (2005a&b) and Fitch, Hauser & Chomsky (2005).

In this thesis we will be focused on whether recursion is specific and universal to humans, so we chose to discuss comparative studies with other species and the case of Pirahã. Thus, in the next two sections we discuss the plausibility of the hypothesis that recursion is specific to our species (2.3.1) and whether it should be considered universal to human language (2.3.2).

2.3.1 Starlings and cotton-top tamarins: is recursion not specific to human beings?

HCF's hypothesis is that FLB has mechanisms that are shared with other animal species. FLN, on the other hand, is the recently evolved computational mechanism of recursion, i.e. not inherited or adapted from an old common ancestral's characteristic, and, therefore, unique to our species. They claim that FLB has a long adaptive history before the advent of the capacity for language as we know it today, but they leave as an open question whether the recursive mechanisms of FLN and the mapping systems to the interfaces are the result of adaptation. They do state that comparative studies with different species' communication systems suggest that language has depended on uniquely human capacities, which diverged from our common ancestral with a chimpanzee six million years ago. As noted by the authors:

> ..., although bees dance, birds sing, and chimpanzees grunt, these systems of communication differ qualitatively from human language. In particular, animal communication systems lack the rich expressive and open-ended power of human language (based on humans' capacity for recursion). The evolutionary puzzle, therefore, lies in working out how we got from there to here, given this apparent discontinuity. (HCF: 1570).

Thus, the core characteristic of language is its capacity for discrete infinity, and comparative research between human and non-human animal communicative systems has shown that this ability is unique to our species. The crucial research question is what mechanism is specific to our species so that we could have developed this capacity. HCF believe this is recursion, based on the fact that no other animal species present such a mechanism.

FLB is formed of mechanisms that are shared both with other domains and with other species. For instance, categorical perception has been believed to be a uniquely human capacity (Liberman et al, 1957; 1967). However, it has been proved to be a vertebrate characteristic, present in chinchillas, macaques and birds (Kuhl & Miller, 1975; Kuhl & Padden, 1982; Kluender et al, 1987). Thus, categorical perception is not uniquely human; therefore, even being essential to human language, it is not unique to the species, so it must not be placed in FLN.

In order to claim that recursion is human specific it is necessary to attest that

this mechanism is not present in other species. As we will see in what follows, birds turned out to be an important case study. As pointed by HCF, birds have a critical period to learn birdsongs. Thus, if they are isolated from their conspecifics during that period they cannot acquire the songs. This is very similar to human language acquisition, but are all properties shared? What do humans and birds have in common regarding their communicative system and what is specifically human?

HCF claim that speech perception and production is one common characteristic between humans and birds, as the latter are also able to recognize and produce formants in their vocalizations. As birds manipulate their vocal tract to produce formants in their chant and recognize the chant of other conspecifics, this is not a human particularity.

Vocal imitation is another characteristic which HCF claim to be shared with birds, although they are able to imitate only in the modality of song, while humans can imitate in multiple modalities. According to the authors, birds also have rich conceptual representations, although their vocalizations are not completely able to transmit these concepts. Apparently, the computational system, which links both meaning and sound in human language, is missing in birds. This computational system defines FLN - recursion and the mappings to the interfaces.

These considerations lead us to Gentner et al (2006), who claim for recursion in European starlings, a species of birds, which sing long iterative songs. In order to understand their experiment, we should first recall the concepts of finite-state grammars and phrase-structure grammars, which we discussed in section 2.2. A finite-state grammar does not account for language because it does not capture the universal properties of constituency and dependency. A phrase-structure grammar, on the other hand, captures constituency, but does not account for dependency. Both were discouraged by Chomsky (1956, 1957) as models for language. After many developments, Chomsky (1995) reached a more explanatory adequate theory to account for language, with Merge as the recursive operation of FLN.

Gentner et al (2006) tested starlings on whether they are able to learn only finite-state grammars or if they can also develop for context-free grammars (i.e phrase-structure grammar, which does not take into account the structural environment in which a rewriting rule is applied).

In their experiment, they used AB sequences that could be generated by a finite-state grammar (47) and by a context-free grammar (48). The second, third

and fourth sequences of (48) entail recursive center-embedding, of the type found in English sentences like (49). In (49), a whole sentence is inserted between the subject and its predicate, in (48), a sequence AB is broken and a new sequence AB is inserted in between A and B. On the other hand, the finite-state grammar in (47) can give us tail recursion, in which a sequence AB is inserted at the end of another sequence AB, similarly to the English sentence in (50), but not center-embedding recursion.

(47)   ab

        abab

        ababab

        abababab

(48)   ab

        aabb

        aaabbb

        aaaabbbb

(49)   [The rat [the cat bit] died]

(50)   [The man said [that the rat died]]

Their results show that the starlings were able to learn both finite-state grammar and context-free grammar sequences, and the authors conclude from this that the birds have a capacity for recursion.

However, Corballis (2007) point out the difficulty even humans have to parse A^nB^n structures where $n > 3$, such as (51)[17].

(51)   [The rat [the cat [the dog liked] bit] died]

He claims that a possible strategy used by these starlings was to count the numbers of As and Bs and then match them. Corballis defend that in order to

---

[17]For a similar point, see Maia et al (forthcoming), fn. 12.

demonstrate true recursion, it would be necessary to demonstrate the parsing of the structure was done from outwards inwards, preserving constituency. Thus, he believes Gentner et al (2006) have provided no evidence that a nonhuman animal has the capacity for recursion.

Moreover, we already saw that a context-free grammar by itself is not able to account for the dependencies created in structures like (49) and (51), which corroborates Corballis argument that the birds are probably using another kind of combinatorial strategy, so the $A^2 B^2$ context-free grammar the animals acquired is not necessarily recursive.

Van Heijningen et al (2009) also dispute Gentner et al's results. They study another birdsong species, namely the zebra finch. They show that the birds can also discriminate a finite-state grammar from a context-free grammar, but in a test to see whether they could do the same when elements from unfamiliar categories were involved, most of them failed, suggesting that zebra finches were not able to generalize from familiar to novel stimuli. This study suggests that the zebra finches were able to distinguish a finite-state grammar from a context-free grammar based on phonetic generalization. They do not cognize syntactic patterns. The same might be true for starlings. Therefore, Gentner et al's data are inconclusive on whether recursion is present in this species.

Fitch & Hauser (2004) conducted similar experiment with cotton-top tamarins, a New World primate species. They point out that finite-state grammars are found in both human infants and nonhuman primates, but they are not sufficiently complex as to generate all the sentences of a human language. Human languages require a more complex grammar, a phrase-structure grammar, which can embed strings within each other, generating hierarchical structures and long-distance dependencies (Fitch & Hauser, 2004).

Although they do not use the term recursion, what they are showing is that phrase structure grammars involve a recursive mechanism to generate hierarchical structures. This ability, they say, is available to all humans, but it is not yet clear whether other animal species can parse a grammar above the level of a finite-state grammar. They demonstrated that the cotton-top tamarins were able to master a finite-state grammar $(AB)^n$, being able to recognize regularities in an acoustic stream. However, the tested animals were not able to master a phrase structure

grammar A^nB^n, since they cannot recognize hierarchically organized acoustic structures. Therefore, there is no evidence of recursion in other species so far.

To sum up this section, in order to investigate whether recursion is indeed the only content of FLN, more species should be tested for this property, using and improving the experiments carried out by the authors presented here. So far, it seems that recursion and the mappings to the interfaces are indeed FLN.

A final question, crucial to the purposes of this work remains to be answered: is recursion a universal property of language? Next section will focus on the Pirahã debate and the universality of recursion on human language.

2.3.2 Pirahã: is recursion universal?

In this section we shall review the Pirahã debate and make considerations about it *vis à vis* the discussion on recursion presented so far. We start with Everett (2005) to understand his claims about Pirahã. Then, we bring the points made by Nevins, Pesetsky & Rodrigues (NPR, 2009a&b) and Everett's (2009) response. The main goal of this section is to understand whether or not Pirahã represents counter evidence to HCF's claim that recursion is the mechanism of FLN. Is Pirahã a language which does not use recursion (understood as Merge in section 2.2) in its combinatorial system?

Everett's paper *Cultural Constraints on Grammar and Cognition in Pirahã: Another Look at the Design Features of Human Language* (2005) is an attempt to defy an approach to language that deals with universal characteristics, claiming for culture as a deterministic factor in the design of human cognitive capacities. The author affirms that "Pirahã culture constrains communication to nonabstract subjects which fall within the immediate experience of the interlocutors." (Everett, 2005: 609 – The underline is mine, RS). That is, Pirahã speakers are unable to communicate about events or situations beyond their immediate experience, here and now. Everett claims that the linguistic and cognitive constraints in (52) result from the principle in (53).

(52)　　Constraints on Pirahã grammar and cognition according to Everett (2005):

　　　　a.　　the absence of numbers of any kind, or a concept of counting and of any terms for quantification;

　　　　b.　　the absence of color terms;

　　　　c.　　the absence of embedding;

　　　　d.　　the simplest pronoun inventory known;

　　　　e.　　the absence of "relative tenses";

　　　　f.　　the simplest kinship system yet documented;

　　　　g.　　the absence of creation myths and fiction

　　　　h.　　the absence of any individual or collective memory of more than two generations past;

　　　　i.　　the absence of drawing or other art;

　　　　j.　　one of the simplest material cultures documented;

　　　　k.　　monolingualism

(53)　　Pirahã communication is restricted to the immediate experience of the interlocutors.

Everett's conclusion that the grammar can be restricted by cultural constraints is not predicted by Chomsky's universal-grammar model. The author believes that the immediate experience restriction reflects in the way Pirahã speakers encode information. Thus, the claim is that grammatical properties of Pirahã grammar follow from the cultural avoidance of what is abstract (facts unrelated to present experience). In this thesis, I am focusing on the gap presented in (52c)[18]. We are going to evaluate how the supposed Pirahã ban on self-embedding could derive from (53).

First, Everett talks about clausal embedding. In English, verbs like say and think use to ask for a clausal complement, such as in (54) and (55).

---

[18]The other claims in (18) are also controversial, and the interested reader may find important information on Gonçalves (1993, 2000, 2001), Gordon (2004), Frank et al. (2008), NPR (2009a,b), and Everett (2009).

(54)   John said [his mother is sick]

(55)   John thinks [his mother is sick]

The author claims that in Pirahã, the expression of such notions does not involve the embedding of a clause inside another, as is observed in the English examples. See (56) [Source: Everett, 2005: 624: (24)].

(56)   ti gai-sai   kó'oí   hi   kaháp-ií
         1 say-NMLZR Kó'oí   3   leave-INTENTION
         'I said that Kó'oí intends to leave' (lit. 'My saying Kó'oí intend-leaves')

According to Everett's analysis, the clause containing the verb *say*, which is nominalized, is a possessive noun phrase *my saying*, and the portion *Kó'oí intends to leave* is just a comment, a juxtaposed clause "not obviously involving embedding" (Everett, 2005: 624). His conclusion that embedding would not be obviously involved follows from word order facts in the Pirahã grammar. Pirahã is a Subject-Object-Verb, thus in Everett's rationale, for a clause to be the complement of another it should precede the verb of the matrix clause, as it functions as an object. He believes that Pirahã's grammatical choice for juxtaposition is a consequence of the principle of immediacy of information encoding, since each juxtaposed clause is a close semantic unit.

Everett also analyzes conditionals. One of his examples is (57) [Source: Everett, 2005: 627: (32)].

(57)   pii   boi-sai   ti   kahapi-hiab-a
         water   vertically move-NMLZR   1   go-NEGATIVE-DECLARATIVE
         "If it rains, I will not go" (lit. "Raining I go not")

Everett recognizes the semantic relation between the clauses, but he sees no syntactic embedding involved, analyzing (57) as the juxtaposition of two clauses.

Another example of lack of embedding would be the absence of relative

clauses in the language. The example in (58) [Source: Everett, 2005: 627: (33)] is not given a translation to English by Everett. He only states that the second clause would be, by itself, translated as *What did Chico sell*, but that in this context it functions as a correlative. He glosses the two sentences in (58) with a period between both because he sees each as an independent sentence. (59) is my attempt to translate the example to English according to Everett analysis.

(58)   ti baósa-apisí 'ogabagaí. Chico hi goó bag-áoba

         *1 cloth-arm   want.   Chico 3 what sell-*COMPLETIVE

(59)   'I want T-shirt. The one Chico sold' or 'I want the T-shirt that
         Chico sold'.

The example in (60) [Source: Everett, 2005: 628: (34)] is not considered clausal embedding either. The author analyzes the desiderative suffix *-soog* on the verb as evidence of the absence of biclausality, but he does not offer a clear syntactic reason for the desiderative suffix to exclude a self-embedding construction for (60).

(60)   'ipóihií 'í gí kobai-soog-abagaí

         *woman 3 2 see-want-*FRUSTRATED INITIATION

         'The woman wants to see you'

Trying to explain that (53) is responsible for the lack of clausal self-embedding, Everett argues that the clauses are rather juxtaposed because each of them is a semantic unit, so the information is encoded in separate utterances (Everett, 2005: 625). In other words, for Everett, each proposition constitutes a closed semantic unit with immediate information encoded. Because of (53) a semantic unit cannot be syntactically embedded within another semantic unit. Thus, another proposition must be formed, yielding juxtaposed propositions, rather than self-embedded clauses.

Everett, then, turns to possessive cases in Pirahã. He shows two ungrammatical examples, (61) and (62) below [Source: Everett, 2005: 628: (35)-(36)], and then discusses how Pirahã would express the idea of multiple embedding without the resource of self-embedding, with example (63) [Source: Everett, 2005:

628: (37)].

(61) * kó'oí   hoagí   kai   gáihií   'íga
       Kó'oí   *son*   *daughter*   *that*   *true*
      'That is Kó'oí's son's daughter'

(62) * kaóoí 'igíai   hoagi   kai   gáihií   'íga
        who   *son*   *daughter*   *that*   *true*
      'Whose son's daughter is that?'

(63) 'ísaabi   kai   gáihií   'íga   kó'oí   hoagí   'aisigí-aí
      'Ísaabi daughter   that   true   Kó'oí   son   the same-be
      'That is 'Ísaabi's daughter. Kó'oí's son being the same'

With (61) and (62), the author shows that recursive possession is not possible neither in the declarative nor in the interrogative form. Following from (53), the impossibility of more than one level of embedding here must be due to the fact that each possessive noun phrase is a semantic unit with immediate information encoded. That is what Everett claims. He asserts that every Pirahã knows every other Pirahã, so we could say that one possessive noun phrase is a semantic unit in itself, what makes an extra level of embedding unnecessary.

Everett's analysis of the possessive noun phrases in (61)-(63) as semantic units is similar to Hinzen's (2014) treatment of some DPs or PPs as referential objects, analogous to propositions. However, under Hinzen's rationale, the referentiality of the possessives in (61)-(63) would rather allow for self-embedding, a contrary result to the expected by Everett's cultural constraint (53).

Example (63) is grammatical and only occurs because the family they are referring to is foreign, but no self-embedding would be involved in the structure according to Everett's analysis, being rather used the resource of juxtaposition.

Finally, Everett talks about embedding in modification. Although he shows one example involving multiple modification (64) [Source: Everett, 2005: 629: (38)], he states that such a structure is rare and chooses to analyze it simply as juxtaposition of adjectives.

(64)   gahióo 'ogií biísai hoí-hio 'ao-'aagá
airplane big   red   two-there possess-be
'There are two big red airplanes'

How would (64) be banned by (53)? Everett does not explain it. Following from his analysis, we would have to find in each juxtaposed adjective a semantic unit, or a referential object. As we cannot evaluate adjectives in themselves as a true or false object (see this issue in section 2.2), it is not plausible to think they are disconnected to the noun they are modifying. That is, adjectives in themselves do not encode any kind of information, as long as they are modifying a noun. Thus, (64) could not be banned based on (53).

Everett tries to find in Pirahã a case of ambiguous modification, such as (65) in English.

(65)   Old men and women

The ambiguity in (65) lies in the fact that it can be analyzed as one or two DPs. The adjective *old* could then be modifying only *men* (66a) (i.e. only *men* are being referred to as old, while the group of *women* could be composed by young or young and old women). However, in (66b), both *men* and *women* are being modified by *old* (i.e. the whole group of people is referred to as old; nobody could be young).

(66)   a.   [old [men]] and [women]

         b.   [old [men and women]]

According to Everett, Pirahã would not allow for such ambiguity because the conjunction of noun phrases with modifiers is not allowed. Although he does not explain why, if we analyze a noun phrase with a modifier as a referential object, we could follow his rationale that (53) bans embedding of closed semantic units (immediate information is encoded in each unit). Thus, the Pirahã equivalent for *old men and women* would be (67) [Source: Everett, 2005: 629: (39)]. For Everett, (68) [Source: Everett, 2005: 629: (40)] would more clearly involve juxtaposition, since the modifier *old* is repeated in the construction.

(67) 'ogi-áag-aó   toío-'aagá   ígihí   'ipóihií   píaii
      big-be-thus   old-be   man   woman   also
      'Everyone (lit. "people bigness") is old. Men and women too'

(68) 'ogiáagaó   toío'aagá   'igihí   toío'aagá   'ipóihií   toío'aagá   píaii
      big   old   man   old   woman   old   also
      'Everyone (lit. "people bigness") is old. Men and women too'

After analyzing the cases above, Everett's conclusion is that there is no evidence for embedding in Pirahã's morphological structure and that this lack is due to the rising of informational flow caused by embedding, beyond the limits imposed by (53).

We see clearly that Everett is presenting a ban on specific recursion in Pirahã, or self-embedding, as defined in 2.2. Thus, the paper does not present any evidence against the design of language proposed by HCF. Everett does not provide any evidence that Pirahã utterances are not combinatorially combined (e.g. through an operation as Merge), nor even is he proposing a new theory to explain the mechanisms behind discrete infinity. The only point he tries to make is that culture is able to interfere with Grammars.

NPR (2009a) discuss Everett's proposal, analyzing his data under the light of his own previous works (Everett, 1983; 1986) and promoting cross-linguistic comparisons by bringing relevant examples from other languages which function similarly to Pirahã. They show that Everett's principle in (53) could not be at work in those languages as well, concluding, thus, that Everett's attempt to link culture and Grammar is not productive.

The authors were the first to observe that the concept of embedding used by Everett (2005) is actually that of self-embedding: "putting one phrase inside another of the same type or lower level, e.g., noun phrases in noun phrases, sentences in sentences, etc." (Everett, 2005: 622). NPR redefine the notion of embedding in Everett, based on the concepts of dominance and of what is to be considered a phrase. Their more precise definition of the ban presumably present in Pirahã's grammar is given here in (69) [Source: NPR, 2009a: 362: (4)].

(69)   No phrase α may dominate a multiword phrase β unless α and β belong to distinct syntactic categories.

If the ban in (69) is indeed at work in Pirahã, it is necessary to prove how it follows from (53). NPR (2009a) clarify Everett's reasoning, explaining that the immediacy of experience would require one event per utterance, then, as each utterance would encode one event only, embedding would be banned from this grammar. NPR, however, disagree with Everett, claiming that the notions of immediacy of experience and one event per utterance are independent. They give the following example:

> Imagine that X has personally witnessed Y uttering the sentence *A boat is coming*. Suppose X now reports on Y's action with an English-style embedded clause as in 7. (7) Y said [that a boat is coming].

> The boat's arrival counts as immediate experience for X since it was 'recounted as seen' by the living individual Y; and Y's speech act is immediate experience because X witnessed it personally. Example 7 does, however, violate the principle of one event per utterance, since it mentions two distinct events, Y's speech act and the boat's arrival.

> If embedded clauses like the bracketed constituent in 7 are impossible in Pirahã, as claimed in *CA* [Everett, 2005], it is at best the principle of one event per utterance that excludes it. (NPR, 2009a: 363).

In other words, immediacy of experience is not an impeding factor for (7) in English, an utterance involving two events. If Pirahã bans a structure such as (7), the constraint is not a matter of lack of immediate experience, but of the gathering of two different events in one sentence. Going further, NPR also evaluate if the number of events in an utterance is a constraint on embedding (70) [Source: NPR, 2009a: 364: (8)].

(70)   a.   The apple [that I am now looking at] is rotten.
            (clause embedded within clause)

        b.   [Mary's brother]'s canoe has a hole.
            (NP embedded within NP)

        c.   Old [men and women] arrived.
            (conjoined Ns embedded within NP)

As NPR evaluate, as neither of the utterances in (70) involve more than one event, the plurality of events could not be the reason of the ban on embedding either. Thus, Everett's principle of immediacy of information encoding involves unrelated notions and fails to explain why embedding in Pirahã would be banned.

Moreover, NPR show Pirahã is not exceptional with respect to embedding, leaving no reason to reject the existence of a universal aspect of human language. Before they consider the case of recursive possessives and the ban on self-embedding claimed by Everett (2005), they clarify how the operation Merge captures the concept of recursion treated by HCF and how it works, something we also did previously in this chapter. They add that Merge is a 'dumb' operation, because there is nothing in itself that avoids its repetition at some point.

If Merge is a 'dumb' operation, not able to discriminate what kind of elements and how many of them it can concatenate in a derivation, what factor could be constraining self-embedding in Pirahã? If syntax is autonomous, working through recursive applications of the dumb operation Merge, a ban on a specific type of embedding does not prove language is not recursive, but leave open the question on whether syntax can suffer the interference of other systems (cf. discussion on Roeper (2011) and Hinzen (2014) in section 2.2).

This interference should not be a cultural constraint such as that proposed by Everett (2005), as NPR have demonstrated the failure of that rationale. Could it be a cultural constraint at all? Could the interface systems impose their own restrictions to the combinatorial system of language?

NPR (2009a) state that Merge may be restricted, constraining the set of structures licensed by languages. As the authors show, Pirahã is not exceptional in banning self-embedding. In English, for instance, a noun cannot directly merge with another noun (*translation [poems]*). A preposition is necessary to license the NP *poems* (*translation of poems*). This and other restrictive facts are present in different languages around the globe, showing that it is not uncommon that Merge is constrained.

Let us see the case of possessive recursion. As Everett tries to show with the examples listed here in (61)-(63), more than one level of embedding of possessive NPs is not allowed in Pirahã. NPR show that German grammar behaves likewise[19].

---

[19]Dékány & den Dikken (2015) present similar data from Hungarian (i) [Source: Dékány & den Dikken, 2015: 1: (1d&e)] arguing that these are also cases of restriction of recursion/self-

See example (71) (cf. Krause, 2000a&b; Roeper & Snyder, 2005).

(71) a. Hans-ens Auto

John-GEN car

'John's car'

b. *[Hans-ens Auto]-s Motor

[*John-GEN car*]-*GEN motor*

'John's car's motor'

What is the constraining factor on multiple embedding of prenominal possessors in German? According to Krause's (2000a&b) analysis, there might be morphological interference within German's syntax, since the kind of ban seen in (71b) is also found in other genitive case environments. The mechanisms involving genitive case assignment are the ones constraining Merge in cases of multiple embedding. NPR (2009a), then, claim that the same syntactic ban at work in German prenominal possessor recursion, might also be at work in Pirahã. Since German and Pirahã seem to have the same kind of constraint at work in their grammars, NPR show that it is unlikely that German and Pirahã cultures, which are so different, have the same cultural constraint at work in their grammars.

Nevertheless, is self-embedding really banned from Pirahã? NPR analyze Pirahã data involving complement clauses published by Everett (1983, 1986). First, they present the case of nominalized verbs in Pirahã complement clauses. This strategy is very common cross-linguistically, with embedded verbs losing their valency and becoming more nominal (see Payne, 1997). I will repeat here one of Everett's examples (72) [Source: Everett, 1986: 263: (232)] using NPR's strategy of brackets to emphasize the embedded clauses.

---

embedding within the DP domain. The authors argue that these cases are out because there is no phase head between the two DPs. In a situation in which a c-commands a, a phase head must intervene between the two a if one is embedded within the other. See fn. 15.

(i) a. *ki háza?
who house.POSS
'Whose house'

b. * ki-ki háza
who-who house.POSS
'Everyone's house (distributive reading)'

(72)   hi   ob-áaxáí   [kahaí kai-**sai**]

        3   *see/know-*INTENSIVE *arrow* *make-*NMLZR

        'He really knows how to make arrows'

As noticed by NPR, Quechua (Hermon, 1985; Lefebvre & Muysken, 1987), Turkish (George & Kornfilt, 1981; Kelepir, 2001; Kornfilt, 2001), and many other languages also display such a strategy to construct embedded clauses. Everett (2005) analyzes (72) as the paratactic conjoining of 'arrow-making' and 'he sees well'. He argues for this analysis based on the word order pattern of the language. As Pirahã is an SOV language, one would expect the complement clause to precede the matrix verb of the matrix clause.

However, this inversion in the expected word order is not an idiosyncrasy of Pirahã. NPR state that this choice is very common cross-linguistically. They are referring to the typological study of Dryer (1980), who proves with cross-linguistic data that it is a tendency for OV languages to invert the order to VO when the complement of the verb is an embedded clause. But what is the reason for such an inversion in these languages? There are different explanations, such as the inability to receive syntactic case if the SOV order is maintained in German embedded clauses (Stowell, 1981). In the case of Pirahã, NPR (2009a) notice that the inversion to SVO order is a matter of *heavy NP shift*. The SVO order would be preferred when the object is too heavy/long or complex (Ross, 1967). This had already been noticed on Everett's earlier works (1983, 1986), where SVO order is attested when a heavy NP is the complement of the verb. See example (73) [Source: Everett, 1986: 226: (107)].

(73)   tiobáhai koho-ái-hiab-a tomáti gihió-kasí píaii taí píaii

        child   *eat-*ATEL-NEG-REM *tomato* *bean-name* *also* *leaf* *also*

        '(The) children do not eat tomatoes or beans or leaf'

The importance of (73) for the present discussion is that it does not involve sentence embedding, as it is a matrix clause; however, the order SVO is attested because the object is too heavy. Hence, the shift from SOV to SVO in (72) cannot be understood up front as evidence for clausal juxtaposition as argued in Everett

(2005). Thus, as noticed by NPR, it is natural to expect for a clausal complement to be postverbal, since it is a heavy object.

Therefore, Pirahã's more frequent SOV word order is not an argument against the possibility of nominalized-verb clauses to be the complement of other clauses in an SVO configuration. On the contrary, it seems to be in accordance with Pirahã's grammar, with heavy nominals postponed to the verb.

Regarding the domain of relative clauses, NPR (2009a) find surprising that Everett (2005) considers (58) above as a correlative, but denies the usage of self-embedding for such a construction. As they notice, languages such as Hindi (Bhatt, 2003) and Tibetan (Cable, 2009) also have correlative constructions, where the correlative and the phrase it semantically modifies co-occur in a sentence. Thus, there is no exceptionality involved in the way Pirahã grammar builds syntactically the notion of relativity and there is cross-linguistic evidence suggesting that correlatives involve self-embedding. Therefore, unless we want to treat Pirahã as a radically different language, we should not start a research on this language assuming that it blocks this type of recursive structure.

Everett (2009) persists on the idea that "Pirahã falsifies the single prediction made by Hauser, Chomsky, and Fitch (2002) that recursion is the essential property of human language." (Everett, 2009: 405). He believes that NPR (2009a) are the ones who misunderstood the concept of recursion and that Merge is not relevant to that discussion. The author claims that there are non-recursive approaches to syntax, such as the transformational generative grammar (Culicover & Jackendoff, 2005), linear-precedence rules with semantic linking rules (Gazdar et al., 1985) and construction-grammar (Goldberg, 1995; 2006).

Everett (2009) chooses to deny NPR's (2009a) claims about recursion as Merge and not self-embedding because he is choosing to consider other theories about languages instead of the framework upon which HCF make their claims about the design of human language. Notice however, that the frameworks cited do not deny recursion. Infinity is an empirical property of human language, and as such, it cannot be denied, independently of the theoretical framework adopted. To account for infinity one cannot assume recursion as synonymous with self-embedding. It must refer to something more general, a set of finite axioms or operations that can generate an infinite number of complex object (i.e. sentences) taking as their input a finite number of resources (words or morphemes), as discussed on section 2.2

above.

The main problem with Everett's claim about Pirahã is the non-distinction between I-language and E-language. As observed by Di Sciullo (2014) and Watumull et al. (2014), it is important to differentiate I-language from E-language in order to characterize the grammar of any given language. I-Pirahã is different from E-Pirahã. I-Pirahã is the recursive function that generates E-Pirahã and only E-Pirahã. Notice, however, that the absence of a given string in E-Pirahã does not allow us to conclude that a given structure is absent in I-Pirahã. That is, the absence of self-embedding in strings, as opposed to structures, does not mean that Pirahã grammar is not recursive[20]. The presence of self-embedding strings is a sufficient but not a necessary condition for us to assume recursion as a capacity of language.

At any rate, throughout this chapter, we have seen the observations made by scientists about language centuries ago and how they led to the development of Chomsky's theory about the innate and universal capacity for language in humans. Every scientific work is a choice among theories, and this thesis chooses the minimalist phase of the generative framework as the theory that better offers explanatory adequacy among other theories. It is not in the scope of this thesis to compare each and every other approach to phrase structure theories, such as the ones cited by Everett.

Considering Everett's claims in (2005) and (2009) are not in accordance with data presented in early work published by the same author, new fieldwork in Pirahã was necessary. In our research, we concentrate on possessive constructions, and in the next chapter, we presented the data we collected in a two-month fieldwork in the Pirahã village of Piquiá.

---

[20]See Hale (1975) for a similar claim about counting in Walbiri.

3
Recursive Possessives in Pirahã

In this chapter I report new data on recursive possessives, which I collected during fieldwork in the Piquiá village, from October to December, 2014. The informants were mostly young and old men, although some women agreed to participate in the elicitation tasks too. Children always refused to participate. A partially proficient speaker of Pirahã, Augusto Diarroi, known as Verão (Higao is the way Pirahãs pronounce it), and his mother, Ms Raimunda (called Ioisoi by Pirahãs), a Brazilian proficient speaker of Pirahã, helped me during the elicitation sessions.

The elicitations were done with the aid of drawings, pictures and scenario settings[1]. When presenting the material, first I elicited individual words related to the figures and scenarios, then I elicited the sentences. I also pronounced the sentences inverting the word order or inserting other elements so the speakers repeated them back, correcting any ungrammaticality. Unlike Everett's (2005) report, almost all Pirahã men were able to speak Portuguese, although with different proficiency levels. No one is a fluent speaker of Portuguese, although, most of times they were able to give me some translation to Portuguese of sentences they uttered to me.

The chapter is organized as follows. Section 3.1 gives a general overview of nominal expressions in Pirahã, so the reader can get familiar with the patterns involved within this syntactic domain. It also introduces possessive DPs, the topic of this thesis. Section 3.2 presents and discusses data on recursive self-embedding possessive structures. In section 3.3, we discuss alternative analysis for possessives in Pirahã, which would not involve recursion/self-embedding. Section 3.3.1 focus on discussing an alternative structure with a covert possessive verb "have", breaking the sequence noun-noun. As we will show, this analysis may run shortly in its empirical coverage, as more than two levels of embedding is possible. Then, in section 3.3.2, we discuss an analysis in which the possessor is not an argument of the possessum noun, being rather a topic within the DP. This analysis is also problematic, as it does not explain cases of more than two levels of embedding. In

---

[1]For exemplifications of drawings used during the fieldwork, see appendix, photo 8.

addition, as we will show, when topicalization happens, a pronoun reduplicates the topic. The conclusion, thus, is that the possessive constructions in Pirahã are bona fide cases of self-embedding within DPs.

3.1 Nominal Expressions in Pirahã

Nominal expressions are phrases in which a noun phrase is taken as the complement of some functional categories which may be headed by a determiner (Abney, 1987). These expressions may function as subjects, object within sentences, predicatives in copular constructions, or as the complement of adpositions. Nouns may be modified by adjectives, adpositional phrases and relative clauses. These expressions may even consist of a single noun, pronouns or proper names. In what follows, we present a brief overview of these expressions in Pirahã. Then we move to possessive constructions.

The categories that modify nouns in Pirahã are postpositional phrases (1), adjectives (2), cardinal numbers (3) demonstratives (4), possessives (5) and quantifiers (6). [Source: Salles fieldwork, 2014].

(1)   kapiigaitoi **tabo apo**

pencil   table   on

'(The) pencil (is) on (the) table'

(2)   kapiigaitoi **miisai** aagaha

pencil   red   LOCATIVE[2]

'There is (a) red pencil there'

(3)   ti hoisi **hoihio** xao aaga

1 son   two have LOCATIVE

'I have two sons'

---

[2]Everett (1983; 1986; 2005; 2009) analyzes *aagaha* as a copular "be". I disagree with this analysis and I argue for a locative reading for *aagaha*, following Freeze (1992), which I comment later on this chapter.

(4)   tiobahai **gaihi**

child   DEM

'That child'

(5)   **ti** kaai

1 house

'My house'

(6)   niai kapiigaitoi **maagiso** xao aaga

*2   pencil   a lot   have LOCATIVE*

'You have a lot of pencils'

From the examples above, we can see that modifiers in Pirahã tend to be postnominal, but in possessives, the modifier (i.e. the possessor noun) is prenominal.

According to Everett (1986), grammatical locative case and instrumental case may be adnominally fixed, see examples (7) and (8) [Source: Everett, 1986: 244: (171a); 208: (37)]. However, during fieldwork we found evidence for the locative marking -*o* only, which is given in (9) [Source: Salles fieldwork, 2014]. I tried to replicate examples involving the instrumental case in (8), but speakers never gave me the instrumental marker.

(7)   bií   xi   kaí- o xab- i-í- haí

Bií 3FEM house-LOC stay-?-PROX-RELATIVE CERT

'Bií will stay in the house.'

(8)   ti xií tó -p -á -há   taísi

1 tree fell-IMPERF-REMOTE-COMPLETE CERT axe

tagasága-xai piai xií xóihi

machete-INST also tree small

'I felled the tree with an axe and a machete. (It was) a small tree.'

(9)   ti   ibaisi kaai-o

1 spouse house-LOC

'My wife is in the house'

Pirahã has little expression of functional material, at least morphologically[3]. Neither number nor gender are marked. Grammatical relations are not morphologically expressed either (for that reason, the word order SOV is rigid, marking the thematic relation between nominal expressions and predicates). There is no morphological marker of (in)definiteness either, i.e. Pirahã is an articleless language.

The internal structure of nominal expression within Grammar, were carefully considered during the 1980's. Since Szabolcsi (1983, 1994) and Abney (1987), the NP structure has been considered analogous to the clausal domain. These works culminated in the DP (Determiner Phrase) Hypothesis, which in its modern version takes determiners to be the head of a functional category analogous to CP (Complementizer Phrase). The NP layer would contain the thematic relations inside the DP (analogous to *v*P layers inside CPs). The DP structure as proposed by Abney is given in (10).

(10)   DP
         |
         D'
        / \
       D   NP
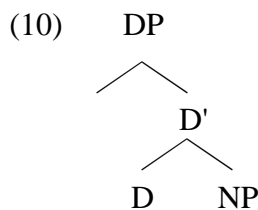
Later, other functional categories (analogous to T (Tense)), responsible for agreement in the DP domain have been suggested, such as Ritter's (1991) number functional category NumP, and Picallo's (1991) GenderP. Since functional categories such as number and gender were not observed in Pirahã, we will not discuss their syntactic nature here. The functional category D may be indirectly observed in structures with pronouns and proper names. Following Longobardi's (1994, 2005) proposal that D is the locus of referentiality, proper nouns and definite pronouns in Pirahã should occupy the D position at the surface structure. Thus, we can consider that Pirahã is a DP language[4].

[3]Since Pirahã is a tonal language, careful analysis of supra segmental layers within phonology is needed to conclude that number, gender or definiteness are not expressed in PF. The expression of functional categories through supra segmental phonology is attested in other languages, e.g. Kaingang (Nascimento, 2013), in which different temporal notions are expressed through different prosodies.

[4]It is worth noting that Bošcović (2008) does not consider articleless languages as DP languages.

Pirahã also displays complex nominal expressions. In (11), for example, a relative clause modify the noun *ipoihi* "woman"[5]. Notice, that in this case, the presence of the relative clause forces a definite reading. [Source: Salles fieldwork, 2014][6].

(11)   [[ipoihi [**tabo apo abaipi**]] ti ibaisi]

        woman bench on   sit   1 wife

        'The woman sitting on (the) bench (is) my wife' [7]

Let us now consider the pronominal system used in the language given in (12) [Source: Salles fieldwork, 2014].

(12)   Pirahã pronominal system

a.	**ti**	first person singular
b.	**niai**	second person singular
c.	**hi**	third person singular (sometimes masculine)
d.	**xi**	third person singular feminine
e.	**xis**	third person singular animal

The system in (12) is impoverished morphologically speaking, as the same morphological forms are used in different structural contexts: subject, object and possessives. Examples of possessive usage of this pronominal system are given further below in this section.

The conjoined forms in (13) are used to express the notion of plurality. [Source: Salles fieldwork, 2014].

---

[5]Everett (2005) claims that Pirahã has no relative clauses, only correlatives. This is controversial. In the data he presented in 1986, he gives at least four examples of relative clauses, and during our fieldwork, we collected a data like (11), in which the constituent within brackets semantically works as a restrictive relative clause. See fn. 6.

[6]The semantic relationship between definite determiners and a relative clause was previously noticed in the literature. Kayne (1994), for instance, noticed that with proper nouns, a definite determiner is licensed only if a relative clause is present (i).

(i)   the Paris *(I knew)

[7]Iahoai Pirahã (Capixaba) saw a woman standing next to his wife, who was sitting; he then uttered this sentence to me, explaining which of the two his wife was.

(13)   a.   hiatihi xogiaagao

Pirahã   all

'Us, the Pirahãs'

b.   hi xapio

3 other

'He/They'

c.   hi xogiaagao

3   all

'They (all)'

d.   xi   xogiaagao

3FEM   all

'They (feminine) (all)'

In Everett (1983), the plural forms of pronouns may be composed by conjoining the singular forms using the conjunctive term pío 'also', as shown in (14) [Source Everett, 1983: 144: (301), translation RS].

(14)   ti gíxai pío ahá-p-i-i

1   2   also go-IMPERFECTIVE-PROXIMAL-COMPLETE CERTAINTY

'I and you are going" (= we are going)'

Another way to express plurality under Everett's analysis is through the insertion of the associative form *xigio*, as in (15) [Source Everett, 1983: 145 (304), translation RS).

(15)   ti   gíxai   xigí-o

1   2   ASSOCIATIVE-LOCATIVE(case)

xopaohoa-i-baí

work-PROXIMAL-INTENSIVE

'I work a lot with you" or "we work a lot together'

In my fieldwork I attested the usage of *piai* (= *pío*), as in (16) and *xigio*, as in (17) [Source: Salles fieldwork, 2014]. Although my examples did not involve conjunctive pronouns, I did not try to confirm (14) and (15) from Everett.

(16)   ti ogai pioahai.   kapoogo piai
       1 want guaraná. Kapoogo too
       'I want guaraná. Kapoogo, too'

(17)   tioisapao kahapi xahaigi xigio
       Tioisapao go out   sister   ASSOCIATIVE
       'Tioisapao went out with (her) sisters'

Possessive constructions in Pirahã are composed by a possessor noun/pronoun followed by a nominal possessum. Very much like in English, although without any genitive/possessive marker. Also, as said above, the pronouns are neutral with respect to its syntactic function. Thus, the canonical word order in Pirahã possessives is *possessor>possessum*. See (18)-(19) [Source: Salles fieldwork, 2014]. It is very common, though, that when the possession relation is clear, the pronoun is not used, such as in (20)-(21) [Source: Salles fieldwork, 2014].

(18)   ti apisi
       1 arm
       'My arm'

(19)   hi kaai
       3 house
       'His house'

(20)   ibaisi   xao   kah-aaga   agipai
       spouse have name-LOCATIVE Agipai
       '(My) wife's name is Agipai'

(21)   maixi magaboi ahoai
          parent   give   shorts
          '(My) father gave (me) shorts'

Before discussing the possessive data I collected, let us show you on what kind of issues we focused during elicitations. First, we needed to attest whether possessive constructions were possible in this language. Second, we looked at the word order possessor>possessum to see if it was consistent with an SOV language. In doing so, we also tested if pronouns in possessor position would yield the same word order. For instance, in Portuguese, when the possessor is a pronoun, it precedes the possessed noun. A non-pronominal possessor, on the other hand, follows the possessed noun as the complement of a preposition. See (22)-(23) from Portuguese. Therefore, we worked with both nouns and pronouns as possessors, in an attempt to capture any word order inversion or other kind of marking related to that.

(22)   minha   casa
          1.SG.POSS house
          'My house'

(23)   a   casa   de meu pai
          the house of my father
          'My father's house'

We also investigated whether (in)alienable relations were morphologically marked in Pirahã, as this is a very common possession split cross-linguistically (Nichols, 1988; Koptjevskaja-Tamm, 2001). In order to verify if (in)alienable distinctions are marked in Pirahã possessive relations, we tried to work with different relations between nouns (e.g. kinship terms, body parts, personal belongings, etc.) to see whether and how they would interfere with word order or yield any other kind of marking.

Importantly to our discussion is to notice that (in)alienable relations are

established according to cultural values semantically codified in a language[8]. A car, for example, is commonly considered alienable possession cross-linguistically, not inherent to the person who possesses it. A mother is commonly considered inherent to the possessor individual, thus, her children inalienably possess her. However, these relations are not universally codified the same way cross-linguistically, there existing no universal pattern telling us which groups of nouns are alienably or inalienably possessed (see Koptjevskaja-Tamm, 2001).

All that being said, examples (24)-(31) are the patterns of possessive constructions with one level of embedding in this language [Source: Salles fieldwork, 2014].

(24)   Pronoun + body parts:

    a.   ti apisi   b.   niai xapai

        1 arm   *2 head*

        'my arm'   'your head'

(25)   Noun + body parts:

    a.   Ioitaopai xapaitai kopaiai   b.   tiobahai xapai

        Ioitaopai hair   black   *child   head*

        'Ioitaopai's hair (is) black'   'The child's head'

(26)   Pronoun + kinship terms:

    a.   ti maixi   aaga   b.   xi   ibaisi   aaga

        1 mother LOCATIVE   *3FEM spouse LOCATIVE*

        'My mother'   'Her husband'

---

[8]This is not the same as saying that culture would be influencing syntactic mechanisms. Culture has its role on how a society organizes the world around it, establishing, in the case we are concerned here, the entities whose possessive relations are inherent or not. Thereafter, syntax will work with lexical codifications, not mattering culture anymore.

(27)   Noun + kinship terms:

    a.   Kaogiai kaai   b.   Iahoai hoisi

        Kaogiai daughter   *Iahoai son*

        'Kaogiai's daughter'   'Iahoai's son'

(28)   Pronoun + personal belongings:

    a.   ti kaáí aaga   b.   hi maosai   toio-koi

        1 house LOCATIVE   *3 clothes old-INTENSIFIER*

        'My house'   'His clothes are very old'

(29)   Noun + personal belongings:

    a.   Higao kaáí   b.   ti moitohoi

        Higao house   *1 boat*

        'Higao's house'   'My boat'

(30)   Pronoun + pets:

    a.   ti mahoisi   b.   xi   giopai

        1 peccary   *3FEM dog*

        'My peccary'   'Her dogs'

(31)   Noun + pets:

    a.   Kahaipoai giopai   b.   Higao cadero

        Kahaipoai dog   *Higao sheep*

        'Kahaipoai's dogs'   'Higao's sheep'

As we can see, no matter what set of variables is at work, the word order is always *possessor>possessum* and no morphological marking appears in the surface structure. There is, thus, no apparent evidence for an (in)alienable split in Pirahã possessives or that nominal and pronominal possessors project differently in this

language.

In the next section, I present and discuss data involving more than one level of embedding, the so-called recursive possessives.

3.2 Self-embedding in Pirahã possessives

We already saw in chapter 2 (section 2.3.2) NPR (2009a) compare Pirahã to German, showing that, regarding possessive recursion, the two grammars may behave likewise – if Everett's (2005) description of the facts is correct. The German data is repeated below (32) (cf. also Krause, 2000a&b; Roeper & Snyder, 2005).

(32) a. Hans-ens Auto
    John-GEN car
    'John's car'

   b. *[Hans-ens Auto]-s Motor
     [*John-GEN car*]-*GEN motor*
    'John's car's motor'

NPR (2009a) claim that whatever syntactic ban on recursive possessives at work in German, could also be at work in Pirahã. However, data informally elicited by Rodrigues in 2013 suggests no ban on self-embedding within possessive constructions in Pirahã. See (33) [Source: Rodrigues fieldwork, 2013].

(33) Iapohen baíxi xapaitaí kobiaí
   Iapohen mother hair white
   'Iapohen's mother hair is white'

Although here we have two levels of embedding, the attested word order maintains the order possessor>possessum discussed in the previous section.

Interestingly, languages may displace a different word order depending on what kind of possession relation is expressed (alienable vs. inalienable). That is, (in)alienable relations can cause word order variation, as observed in Tommo So, a

language from the Dogon family, spoken in the Mali region (34) [Source: Plungian, 1995].

(34)   a.   tigɛ   wo mɔ   b.   u ba
             name he GEN   *you father*
             'His name'   'Your father'

(34a) is an alienable possessive relation, displaying the order possesseum>possessor. (34b), on the other hand, is inalienable, and the word order is the opposite: possessor>possessum. Notice, however, that the alienable possession brings a genitive marking to the construction, arguably revealing the inversion syntactic process as discussed in den Dikken (2013).

Everett (1983) observes that the canonical order in possessives might be inverted for clarification purposes. See example (35) [Source: Everett, 1983: 131: (267), translation RS], where the order is inverted to possessum>possessor.

(35)   giopaí xaxái
        dog   Xaxái
        'Xaxái's dog'

In my data, (36) [Source: Salles fieldwork, 2014] corroborates Everett's point, as the inversion in the canonical order (to possessed>possessum) is accompanied by the repetition of the emphasized noun at the end of the sentence. We will not discuss this here any further, but this inversion process might be a syntactic strategy for focus.

(36)   maixi   capixaba xai xogi xai   hi xapaitai kobiai maixi
        mother Capixaba COPULA old COPULA 3   hair   white   mother
        'Capixaba's mother is old, she has white hair, mother'

Coming back now to (33), it seems to be, at least in its surface structure, a case of DP self-embedding, preserving the canonical word order predicted by typological universals. Dryer (2007) observes that languages differ primarily in the order among the main constituents of the sentence (Subject, Verb and Object).

Interestingly, he also observes that languages with the same word order within the sentence, tend to share word order patterns within other phrases.

One of the things we should ask, then, when observing word order patterns cross-linguistically is to what extent they are related to universal tendencies. For example, Dryer shows that possessives in SOV languages typically present the order possessor>posessum. Thus, Pirahã being an SOV language, as we can see in (37) [Source: Salles fieldwork, 2014], it preserves the word order possessor>possessum as shown in (24)-(31).

(37) ti pioahai ogai
   1 guaraná want
   'I want guaraná'

Another SOV language whose possessive constructions follow the order possessor>possessum, and, likewise Pirahã, presents no possessive morphological marking is Kobon (38) [Source Davies, 1981: 57].

(38) Dumnab ram
   Dumnab house
   'Dumnab's house'

Kotiria (Stenzel, 2013) is another SOV language, from the Eastern Tukano family, spoken in villages by the river Uaupés (Brazil and Colombia). As expected, the possessive constructions in Kotiria also follow the order possessor>possessum. As Pirahã and Kobon, it does not have any marking in the nouns involved in the possessive relation either. See (39) [Source: Stenzel, 2013].

(39) ka   yahiripho'na
   monkey heart
   'Monkey's heart'

Interestingly, Kotiria allows one more level of embedding in possessive constructions. See (40) and (41) [Source: Stenzel, 2013].

(40) yɨ   mahko   manuno

1SG.POSS daughter   husband

'My daughter's husband'

(41) mɨ   mahko   to   hiro

2SG.POSS   daughter   3SG.POSS   COP-SG

'Your daughter's village (lit. 'your daughter her place')'

Both (40) and (41) display the same word order in the two levels of embedding, just as (33) from Pirahã.

During our fieldwork, we considered, thus, the following three parameters (42).

(42)   a. no self-embedding (as German)

b. self-embedding, with (in)alienability distinction triggering word order inversion (as Dogon)

c. self-embedding with no word order change (as Kotiria).

With respect to these parameters, the data gathered are presented in the patterns (43)-(52) [Source: Salles fieldwork, 2014]. Some of these possessives are followed by adjectives, with no overt copula, some are isolated DPs.

(43)   Pronoun + kinship term + body part

a.   ti ibaisi xapaitai kopaiai
        1 spouse hair   black
        'My wife's hair (is) black'

b.   hi maixi xapaitai kobiai
        3 parent hair   white
        'His mother's hair (is) white'

(44)   Noun + kinship term + body part

    a.   Iahoai ibaisi xapaitai kopaiai
        Iahoai spouse hair   black
        'Iahoai's wife's hair (is) black'

    b.   Pihoio hoisi xapai koihi
        Pihoio son   head small
        'Pihoio's son's head (is) small'

(45)   Pronoun + kinship term + personal belongings

    a.   ti xahaigi kaáí naihi
        1 brother house DEM
        'That (is) my brother's house'

    b.   niai ibaisi   maosai ahoasai
        2   spouse dress   blue
        'Your wife's dress (is) blue'

(46)   Noun + kinship term + personal belongings

    a.   Iahoai ibaisi maosai kopaiai
        Iahoai spouse dress
        'Iahoai's wife's dress'

    b.   Kobio hoisi ahoai toio-koi
        Kobio son shorts old-INTENSIFIER
        'Kobio's son's shorts (are) very old'

(47)   Pronoun + kinship term + pet

        nai ti ibaisi giopai
        DEM 1 spouse dog
        'That (is) my wife's dog'

(48)   Noun + kinship term + pet

        aogi   maixi   giopai
        foreigner parent   dog
        'The foreigner's mother's dog'

(49)   Pronoun + pet + body parts

        ti mahoisi kosi
        1 peccary eyes
        'My peccary's eyes'

(50)   Noun + pet + body parts

        Poai mahoisi kosi
        Poai peccary eyes
        'Poai's peccary's eyes'

(51)   Pronoun + personal belonging + part of the object

        ti agaoa moitohoi koihi
        1 canoe motor   small
        'My canoe's motor (is) small'

(52)   Noun + personal belonging + part of the object

Kapoogo agaoa moitohoi xogi

Kapoogo canoe motor   big

'Kapoogo's canoe's motor (is) big'

In conclusion, the data above attest self-embedding in Pirahã, involving two levels of embedding, contrary to Everett (2005) claims about possessive constructions in the language. It also shows that Pirahã follows universal typological tendencies with respect to word order. It shows no (in)alienable split, thus, it is to be subscribed under the parameter (42c), being thus similar to Kotiria.

3.3 Alternative analyses

In this section, we will consider two alternative analyses for the data presented above, aiming to test the theoretical strength of the self-embedding analysis defended so far. In 3.3.1, we discuss the availability of an analysis involving a covert possessive verbal predicate, similar to "have" constructions in English, and in section 3.3.2, we consider a topicalization analysis in which the possessor is not syntactically part of the argument structure of the noun, being rather a topic placed at the edge of the DP. As we shall see, these two alternative analyses are both short in their empirical coverage, especially when three levels of embedding are considered.

3.3.1 A hidden verbal predicate analysis

One could deny a self-embedding analysis for the data presented above, suggesting rather that these constructions involve a hidden verbal predicate, similar to "have" in English, placed after the first possessor>possessum relation, as sketched in (53)[9].

---

[9]Thanks to Uli Sauerland (p.c) for having brought this alternative analysis to my attention (Abralin Congress/2014).

(53)   [[Kapoogo agaoa] [**V** [moitohoi]]]

        Kapoogo   canoe   motor

        'Kapoogo's canoe has a motor'

Assuming the analysis in (53) we would not have two levels of embedding, but rather a possessive DP with one level of embedding followed by a possessive verbal predicate. This is a feasible analysis given that verbs of this sort seem to be covert, as the example in (54) suggests [Source: Keren Everett, 1988].

(54)   oíxai xapaitai hi tihíhi

        your   hair   3 lice

        'Your hair has lice'

Regarding the copula "be" in Pirahã, it is worth noting that it is documented in Everett (1983; 1986). The author reports *xaaga* and *xiiga* as "have/be" and *xai* as "be/do". In my fieldwork, *xiiga* was not observed. The form *xai* was indeed used by speakers, as in (55) and (56) [Source: Salles fieldwork, 2014].

(55)   giopai xi   xai   taobikoi

        dog   3   COPULA lying

        'The dog is lying'

(56)   capixaba moitohoi agaoa xai

        Capixaba motor   canoe COPULA

        'Capixaba has (a) canoe (with a) motor'

As already indicated in the translations for (55)-(56), *xai* is used both as "be" (55) and "have" (56) in Pirahã, differently from what Everett described ("be/do"). According to Bach (1967), it is very common cross-linguistically for "have" and "be" predications to be expressed by the same lexical copula form. See (57) from Hindi, another SOV language (cf. Freeze, 1992: 576).

(57)   a.   kamree-mẽẽ aadmii hai

room.OBL-in   man   COP.3SG.PRES

'There is a man in the room'

b.   larkee-kee   paas   kutaa   hai

boy.OBL-GEN PROXIMITY   dog   COP.3SG.PRES

'The boy has a dog'

Another example of *xai* is as follows: very often, when I pronounced a Pirahã sentence, the speakers would repeat it back to me starting with *xai* (58) [Source: Salles fieldwork, 2014].

(58)   RS: tioi miihai koihio?

ball   red   small

'Small red ball?'

Pirahã speaker: xai   nai tioi miihai koihi-aaga

COPULA DEM ball   red   small-LOCATIVE

'Yeah, this (thing) here (is a) small red ball'

We analyze *xai* in such cases as a particle, a speech act modifier (cf. Heim et al, 2014). It might be that the speaker is confirming to me that my proposition was correct, although the sentence expressing it was incomplete. Hence, he repeats it back the way he would actually say it, adding *aaga* in the end. In Portuguese, there is a similar construction. See (59), where grammatical correction is added in the answer.

(59)   Speaker A (foreigner):   Essa bola vermelho?

this ball red.MASCULINE

'Is this ball red?'

Speaker B (native):   É,   essa bola   é   vermelha

COPULA this ball COPULA red.FEMININE

'Yeah, this ball is red'

Notice that *é* in (59) is the same morphological form used as a copula be in Portuguese, as in (60).

(60)   Ela   **é**   bonita.
        3FEM.SG COPULA beautiful
        'She is beautiful'

Thus, it is plausible that in Pirahã, too, the copula also functions as a speech act modifier.

Regarding "have" predications, I found the form *xao*, used as in (61)-(65) [Source: Salles fieldwork, 2014]. The occurrence of *xao* is usually accompanied by the locative *aaga*[10]. This is expected, since, according to Freeze (1992), a possessor is widely semantically understood as a location[11].

(61)   ti tioi miisai hoihio xao   aaga
        1 ball red   two   have LOCATIVE
        'I have two red balls here'

(62)   ti ahoai   toio-koi   xao aaga
        1 shorts old-INTENSIFIER have LOCATIVE
        'I have very old shorts here'

(63)   xi   xao   kas-aaga   ioai
        3FEM have name-LOCATIVE Ioai
        'She has (a) name here: Ioai'

---

[10]It is worth noting the phonetic similarity of *xao* and *xai*. It is possible, then, that *xao* is the equivalent of *xai* with an incorporated postposition. Freeze (1992) and Kayne (1993) notice that the incorporation of "be" with a postposition delivers a possessive verbal "have". We will not pursue this analysis here, but we leave it open as a possibility.

[11]Existential predicates, possessive predicates and locative predicates seem to be somehow semantically associated, as noticed by Hornstein, Rosen & Uriagereka (1996). (i), for example, is ambiguous between the possessive reading in (ii a) and the locative reading in (ii b).

(i)   There is a Ford T engine in my Saab.

(ii)   a. My Saab has a Ford T engine.
       b. (Located) in my Saab, is a Ford T engine.

(64)   ti xao hois-aaga

    1 have son-LOCATIVE

    'I have (a) son here'

(65)   kopohi xao xaaga ahoahai tabo apo

    cup   have LOCATIVE blue   table on

    'There is a blue cup on the table'

Everett (1983; 1986) describes "have" predication cases, but under his analysis, *xao* appears as a possessive affix to *xaaga* which is taken to be the Pirahã counterpart to "have". This analysis, however, seems incorrect because Pirahã is a language in which affixes appear suffixed to their hosts (see for instance chapter 1, (section 1.2) examples from (5)-(11)). Therefore, asssuming Everett's analysis, we would in principle expect the form *aaga-xao* instead of *xao-aaga*. In addition, as data from my fieldwork show (66)-(67), *xao* can occur alone, expressing possession [Source: Salles fieldwork, 2014].

(66)   ti xao moitohoi

    1 have motor

    'I have (a) motor'

(67)   RS: niai ibaisi xao   aaga   maosai?

                     2 spouse have LOCATIVE clothes

                '(Does) your wife have clothes?'

          Pirahã speaker: xao

                             have

                            '(She) has'

We have been glossing *aaga* (or *aagaha*) as a locative, throughout this chapter, disagreeing with Everett's analysis that it would be a copula "be/have". In the data we collected, *aaga* (or *aagaha*) is used in contexts such as (68) and (69) [Source: Salles fieldwork, 2014]. In these situations and very often, Pirahãs would point to

the things they were referring to when using *aagaha*. It seems very likely, thus, that it is a locative postposition, indicating an existential reading.

(68)   RS: xi go kasi naihi?
3 WH name DEM
'What (is the) name (of) it?"

       Speaker: topagahai **aagaha** topagahai
recorder LOCATIVE recorder
'(It is a) recorder here, recorder'

(69)   RS: kaoi naihi?
who DEM
'Who (is) that?'

       Speaker: ti hoisai **aaga**
1   son LOCATIVE
'My son here'

An interesting example is (70), because it has both *xai* and *aagaha* in it. I was working with colored circles made of paper over a table and the informants should describe the size and color they saw. In this example, there was a green circle closer to the speaker and a red circle closer to me, both of the same size. He said [Source: Salles fieldwork, 2014]:

(70)   kapiiga tioi kapiiga **xai** ahoasai igihio **xaagaha** gai piai miisai
paper ball paper COPULA green close LOCATIVE DEM also red
'(The) paper ball closer to here is green. That too, (is a) red (paper ball)'

Other examples are (71) and (72) [Source: Salles fieldwork, 2014].

(71)   xai   tioi   kapiga xai   ahoasai

        COPULA ball   paper COPULA green

        'Yeah, (the) paper ball is green'

(72)   xai   tioi   kapiga xai   koihihi xaagaha

        COPULA ball   paper COPULA small   LOCATIVE

        'Yeah, (the) paper ball here is small'

In sum, under the analysis suggested here, *xai* is an overt copula, meaning "be/have", *aaga(ha)* is a locative and *xao* is a possessive verbal predicate similar to "have"[12], usually followed by the locative *aaga*.

Having described the overt copula in Pirahã, we may now turn to its covert occurrence in this language. Is it possible to infer from verballess constructions such as (73) and (74) a covert verb meaning "be" or "have"? At this point of our research, it is not possible. In (73)-(74), for example, to the extent that these sentences involve a covert verbal predicate, one can recover either a "be" meaning or a "have" meaning from them [Source: Salles fieldwork, 2014].

(73)   giopai xis sabi-koi

        dog 3ANIMAL angry/anger(?).INTENSIFIER

        '(The) dog, it is very angry'

        '(The) dog has a lot of anger'

(74)   ti hoisi naihi. nai kaba

        1 son DEM. DEM NEG

        'That is my son. That (one) no'

        'I have that son. Not that (one)'

Let us go back to possessive constructions, such as (75) [Source: Salles fieldwork, 2014]. For (75) we can posit three different structures. The first one (76a) would be a structure with two levels of DP self-embedding, and an adjective *xogi*

---

[12]Under Freeze's (1992) analysis, "have" would also be included under the label *locatives*. However, it is out of the scope of this thesis to discuss the structure of locatives in Pirahã.

modifying the less embedded NP *moitohoi*, as in (76b)[13]. The second one is also a structure with two levels of DP self-embedding, but it would involve a copular covert verb, which would project the VP in (77b). The third possible structure would involve a possessive verbal predicate, as shown in (78). This predicate would be the Pirahã null counterpart of English "to have". (78) differs from (76) and (77) in that it does not involve two levels of self-embedding within the DP, but only one.

(75)   Kapoogo agaoa moitohoi xogi
         Kapoogo canoe   motor   big

(76)   a.   'Kapoogo's canoe's big motor'

         b.
```
                          DP
                    ┌──────┴──────┐
                   D°             XP
                         ┌─────────┴─────────┐
                        DP                   X'
                      ┌──┴──┐             ┌───┴───┐
                     D°    XP            X°       NP
                        ┌───┴───┐             ┌────┴────┐
                       DP       X'           NP        AdjP
                       │      ┌──┴──┐         │          │
                   Kapoogo   X°    NP      moitohoi    xogi
                                   │
                                 agaoa
```

---
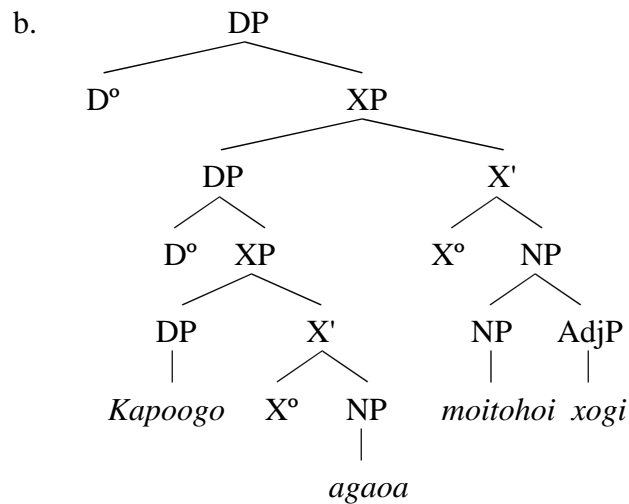
[13]We are assuming here that possessive DPs are spelled out in the specifier position of a functional projection (call it XP) between the NP and the DP. For an introductory view of this functional category, see Adger (2003).

(77)   a.   'Kapoogo's canoe's motor is big'

b.
```
                        VP
                  ┌──────┴──────┐
                 DP             V'
            ┌─────┴─────┐    ┌───┴───┐
           D°          XP   V°      AdjP
                   ┌────┴────┐  │      │
                  DP        X'  "to be" xogi
               ┌───┴───┐  ┌──┴──┐
              D°      XP X°    NP
                   ┌───┴───┐    │
                  DP      X'   motohoi
                   │    ┌──┴──┐
                Kapoogo X°   NP
                              │
                            agaoa
```
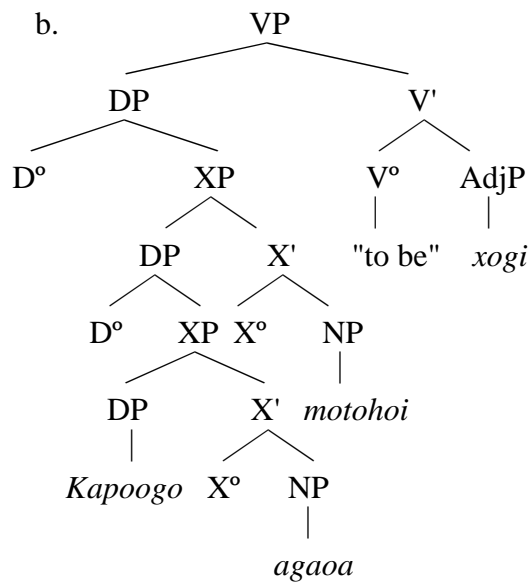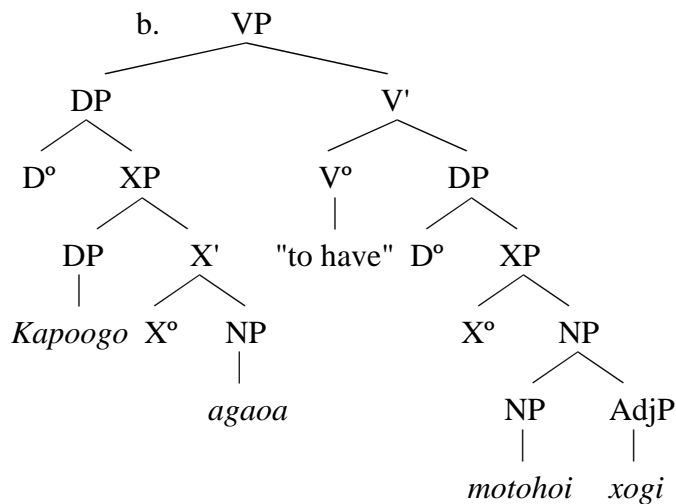
(78)   a.   'Kapoogo's canoe has a big motor'

b.
```
                   VP
            ┌───────┴───────┐
           DP               V'
        ┌───┴───┐      ┌─────┴─────┐
       D°      XP     V°          DP
           ┌───┴───┐  │       ┌────┴────┐
          DP      X' "to have" D°       XP
           │    ┌──┴──┐             ┌────┴────┐
        Kapoogo X°   NP            X°        NP
                      │                  ┌────┴────┐
                    agaoa               NP       AdjP
                                         │         │
                                      motohoi    xogi
```

We did not test these three analysis given the shallow understanding we have on the Pirahã grammar so far. However, for our discussion on the availability of self-embedding in Pirahã, we need to consider the possible structure in (78) more carefully. Both (76) and (77) involve two levels of self-embedding within the DP. (78), on the other hand, involves only one level. Thus, (78) differently from (76) and (77) is compatible with Everett's claim that the language does not allow more than one level of recursion/self-embedding within the nominal domain. Nevertheless, the possible availability of (78) as a structure for possessives in Pirahã,

however, does not exclude self-embedding in this language. Evidence for this assertion comes from possessive nominal expressions involving a sequence of four nouns, as in (79)-(80) [Source: Salles fieldwork, 2014][14].

(79)   Pronoun + kinship term + kinship term + personal belonging

         ti   ibaisi   maixi   maosai kopaixai
         1 husband mother clothes black

(80)   Noun + kinship term + kinship term + personal belonging

         migixoi   ibaisi   maixi maosai kopaixai
         Migixoi husband mother clothes black

Assuming the structure in (78), (79) and (80) are to be translated as (81a&b), respectively.

(81)   a.   'My husband's mother has black clothes'

         b.   'Migixoi's husband's mother has black clothes'

As the interpretations in (81) make transparent, we can think that in (79) and (80) one of the possessive relations is intermediated by a verbal possessive predicate, breaking, thus, a self-embedding structure, as in (78). However, if so, in (79) and (80) we still have two levels of self-embedding within the DP. In order to avoid DP self-embedding completely in these constructions, one needs to postulate two verbal possessive predicates in this structure.

(82)   a.   'My husband has mother has black clothes'

         b.   'Migixoi's husband has mother has black clothes'

---

[14]These examples were elicited using photo 9 in the appendix.

As difficult as it may be to recover the semantics of these constructions, let us just observe that if (82) were possible in Pirahã, it would involve self-embedding of VPs, which is something that Everett (2005) also denies. Therefore, examples like (79) and (80) are hard to be accounted for without postulating self-embedding.

At any rate, "have" predicates have been analyzed as complex predicates involving covert structure (see Benveniste, 1971; Freeze, 1992; Kayne, 1993; Hornstein et al, 1996). Kayne, for example, suggests that the English sentence *John has a sister* has an underlying structure in which a copular verb "be" selects for a possessive DP[15]. Thus, under Kayne's analysis, the structure of (78), a case of two possessive relations, would be more complex, as the DP *Kapoogo agaoa* would have been generated within the DP sitting in the complement position of "have", as shown in (83). In sum, under Kayne's analysis of possessive "have", there is no escape from self-embedding within the DP domain in Pirahã[16].

[15] Kayne's analysis is based on the fact that many languages (e.g. Bengali, Hindi, Punjabi, Guajarati, Marathi and Sindhi (see Bhatt, 1997)) use "be" in possessive constructions.

[16] We would still want to consider the possibility in (i), though. In this possible translation for (80), we would have four juxtaposed clauses.

(i) a. Migixoi has (a) husband.
     b. (He) has (a) mother.'
     c. (She) has clothes.
     d. (They) are black.

However, two points would have to be considered. First, is Pirahã a null subject language? There is no evidence for or against this so far. Second, if null subjects are licensed and the translation in (i) is indeed possible, then we might need to investigate what juxtaposition is and why it should be opposed to clausal embedding. For example, looking at the clauses in (i) one sees the semantic (and probably syntactic) dependency among them. In (ib), for instance, the dropped pronoun is obviously referring to the husband mentioned in (ia). It is hard to tell whether semantically dependent juxtaposed clauses involve embedding or not, though. A possible way to conclude whether juxtaposition encodes dependent inter-clausal relations is by analyzing if the clauses are prosodically dependent (cf. Palancar, 2012). We have not conducted this test so far.

(83)   (= (78) under Kayne (1993) analysis)
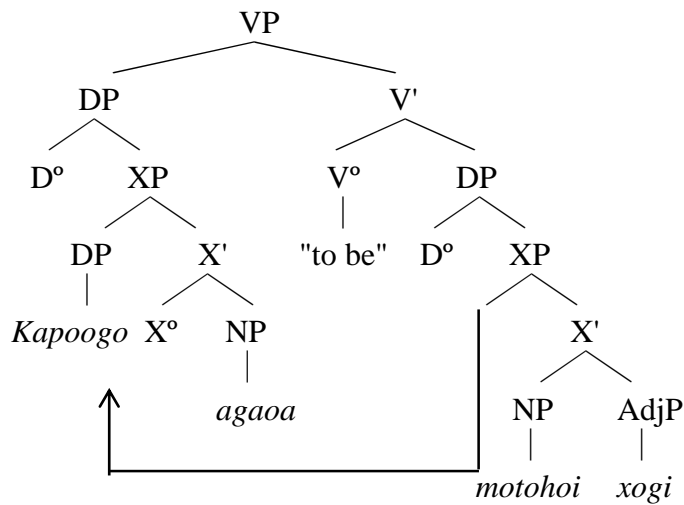
Finally, let us also observe that possessive "have" can be overtly expressed in these constructions, as shown in (84) [Source: Salles fieldwork, 2014]. This case, however, might involve topicalization of the possessive DP, as we will discuss in the next section.

(84)   [TopP migixoi ibaihai mai] DP xi] xao   xaaga   maohai kopaixai
        Migixoi spouse parent   3   have LOCATIVE *clothes   black*
   'Migixoi's husband's mother, she has black clothes'

Concluding, thus this section, we say that Pirahã is a language with self-embedding within the nominal domain. A verbal predicate analysis does not exclude self-embedding in this language, given that three possessive relations (maybe more) can be expressed within the same DP. In addition, assuming an analysis for verbal possessive constructions *à la* Kayne (1993), even DPs involving two possessive relations would require self-embedding.

3.3.2 A topic phrase analysis

Another possibility for possessive DPs involving two, three or more possessive relations, would be postulating that the first DP is outside the nominal

argument structure, hanging on a topic position, arguably on top of the DP[17]. Under this analysis, the example (52), repeated here as (85) would have the structure sketched in (86).

(85)   Kapoogo agaoa moitohoi xogi
        Kapoogo canoe   motor   big

(86)
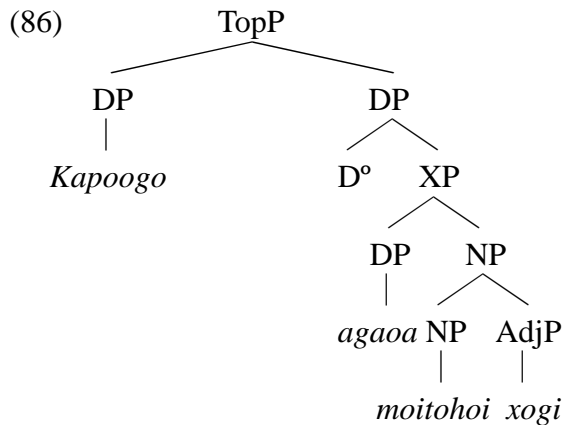
Clearly (86) does not involve two levels of self-embedding, as the first DP is outside the possessive DP. Therefore, this analysis, similarly to the analysis discussed in the last section is compatible with Everett's claim. (86), at first sight, can accommodate cases of three possessive relations (e.g. (79)-(80)) without invoking self-embedding, as sketched in (87).

(87)   [TopP [DP [XP [DP ti] [NP ibaisi]]] [DP [XP [DP maixi [NP [maosai]
                1   spouse   mother   clothes
        [AdjP kopaixai]]]]]]
                black

(87), however, does not make it clear how the topicalized DP is recovered as the possessor of the following noun. One could postulate that it is recovered by a null pronoun. But, if this is right, we would go back to our starting point, which is self-embedding, as the null pronoun would be a DP in [Spec, XP], within the DP, as in (88).

---

[17]Thanks to Luciana Storto (p.c) for having brought this alternative analysis to my attention (Abralin Congress/2014).

(88)   [TopP [DP [XP [DP ti] [NP ibaisi]]] [**DP** [XP [**DP** [XP [**DP** PRO]
[NP maixi]]] [NP [maosai] [AdjP kopaixai]]]]]

A stronger counter evidence against the analysis in (87) comes from the fact that this type of topicalization in Pirahã seems to involve an overt pronoun doubling the topic, that is a case of prenominal possessor doubling, as discussed in Grohmann & Haegeman (2003). See (89)-(90), which were elicited as the answer to the question "What is that", pointing to Poai's finger/head [Source: Salles fieldwork, 2014].

(89)   Poai hi ooi
         Poai 3 finger
         'Poai, his finger'

(90)   Poai hi apai
         Poai 3 head
         'Poai, his head'

The proper name *Poai* is topicalized, and the co-referential pronoun *hi* occupies the unmarked position within the possessive DP. In (91), we have a more complex structure, in which a possessive DP is topicalized, being doubled by the pronoun *hi* which functions as the possessor of the noun *ooi* [Source: Salles fieldwork, 2014].

(91)   ti hoisai hi ooi hi ooi
         1 son   3 finger 3 finger
         'My son, his finger, his finger'

The doubling phenomenon is also observed in the clausal domain in Pirahã, as shown in (92) [Source: Salles fieldwork, 2014].

(92)   higao hi pioahai ogai. ti piai
         Verão 3 guaraná want. 1 also
         'Verão, he wants guaraná. Me too'

(93) is a more interesting case, where the proper noun *Ioihoi* is doubled by a possessor pronoun within a DP in the sentential subject position[18] [Source: Salles fieldwork, 2014].

(93)   ioihoi xi hoihai xiahoikoi higao
         Ioihoi 3 son   leave Verão
         'Ioihoi, her son left, Verão'

Grohmann and Haegeman (2003) offer an analysis of the Germanic prenominal possessor doubling, with particular attention to Norwegian and West Flemish. See (94) from Norwegian [Source: Fiva, 1984: 2] and (95) from West Flemish [Source: Haegeman, 1998: 1].

(94)   Per   sin   bil
         Per REFLEXIVE car
         'Per's car'

(95)   Marie euren vent
         Marie her husband
         'Marie's husband'

Leaving the details of this proposal aside, the authors' point is that prenominal possessor doubling constructions are complex DPs formed by a DP in topic position being doubled by a pronoun in [Spec, PossP], which corresponds to our XP. Therefore, their analysis is the same as ours for Pirahã.

Unfortunately, it is out of the scope of this thesis to discuss the triggers for possessor doubling. More data is necessary to a deeper investigation of possessor doubling in Pirahã.

In sum, the topic phrase analysis, similarly to a hidden verbal predicate analysis (section 3.3.1), does not account for possessive constructions in Pirahã, while dispensing with self-embedding within the DP. Topicalization in this language seems to involve pronominal doubling.

---

[18]In this sentence, the last term *Higao* "Verão", seems to be an afterthought indicating that Ioisoi's son is Verão.

In this chapter, we have shown that the parameter active in possessive constructions in Pirahã follows cross-linguistic tendencies, being similar to other SOV languages, such as Kobon and Kotiria. This was attested in many different semantic types of possessor-possessum relations. Crucially, this discusses cases of recursive self-embedding demonstrating that, unlike German, Pirahã has no ban on self-embedding within DPs.

4
Conclusion

It has been more than a decade since Hauser, Chomsky & Fitch (2002) proposed a design for the Faculty of Human Language in which the capacity of recursion is responsible for discrete infinity, a distinctive property between language and animal communicative systems. Although the concept of recursion has played a central role in the field of formal linguistics, it has been widely misunderstood. The Pirahã debate, started by Everett's (2005) claim that the language lacked embedding, is the clearest demonstration of a misconception of recursion. Thus, this thesis had two main goals: (a) set straight the concept of recursion within formal linguistics and (b) investigate the availability of self-embedding of possessive DPs in Pirahã. These two goals are to be understood as theoretically independent. As we stated in chapter 2, absence or presence of self-embedding in Pirahã is not a proof of the claim that Pirahã is a non-recursive Grammar.

First, we tried to clarify how recursive functions can confer Grammar with the power to generate infinite sentences out of finite means. We did that in chapter 2. Under a minimalist model of Grammar framework, we saw how discrete infinity can be achieved through the operation Merge, which is recursive in itself since it can apply over its own outputs. However, Merge does not take into consideration the labels of the categories it concatenates to form syntactic objects. That is to say that a specific kind of Merge, that of two tokens of the same type, is not to be constrained within syntax in itself. Specific Merge, the so-called self-embedding, must then be restricted outside syntax, more specifically when phases are delivered to the interface systems. Thus, the absence of self-embedding in Pirahã or any other language would not be a counter evidence for Merge, rather revealing interface constraints on the structures delivered by syntax.

Everett (2005), for instance, says nothing about Merge in Pirahã. In fact, in his 1986 paper, he offers examples of complex morphological words, which indicates the existence of a combinatorial system in the language. See (1)-(6) [Source: Everett, 1986: 322: (477)-(482)].

(1)   xabagi + soixaoxoisai → xabagisoixaoxoisai 'saw'
toucan   beak

(2)   xapai + toii → xapaitoii 'ladder'
foot   handle

(3)   hóii + hoi → hóiihoi 'bowstring'
bow   vine

(4)   xapai + soi → xapaísoí 'shoe'
foot   leather

(5)   pi + gáía → pigáía 'scissors'
thorn crooked

(6)   kao + ogiái → kaogiái [type of bass (fish)]
mouth big

Therefore like any other grammar, Pirahã has Merge. In order to prove that a language does not have Merge, one has to prove that the language is linear. Everett never proved it for Pirahã.

Regarding the unavailability of self-embedding in Pirahã at any domain, we proved in chapter 3 that Everett (2005) is empirically incorrect. We showed many instances of self-embedding of possessive DPs. Even if we propose a verbal predicate analysis or a topic phrase analysis to the examples provided, self-embedding cannot be discarded with to account for the empirical evidence. The word order parameter active in Pirahã possessive constructions is actually expected, under the assumption that cross-linguistically, SOV languages follow some word order pattern tendencies, such as the order possessor>possessum. Thus, this language, as any other language that we know, exhibit universal patterns.

Having said so, let us now compare our data and conclusions to the state of affairs within Pirahã research. Our conclusion that there is no ban on self-embedding in this language within the nominal domain (at least for possessive DPs) goes towards the points made by Amaral et al. (forthcoming) and Rodrigues et al.

(forthcoming). Amaral et al. conducted a pilot study on Pirahã postpositional phrases and found evidence for self-embedding within this domain. The example they elicited is (7) [Source: Amaral et al., forthcoming: 12: (22)].

(7) tabo apo tiapapati apo kapiiga apo gigohoi
     board on   chair   on   paper   on coin
     'The coin on the paper on the chair on the board'

Rodrigues et al. (forthcoming) also found evidence for self-embedding in Pirahã. Their analysis involve self-embedding of a bare VP inside a matrix VP in obligatory control constructions. Their example is given in (8) [Source: Rodrigues et al, forthcoming: 6: (14a)].

(8) ti kapiiga kagakai ogabagai
     I paper study   want
     'I want to study'

Most intriguing, however, are cases of sentential self-embedding. That is, the embedding of a CP inside another CP. Everett (1986) presents (9) [Source: Everett, 1986: 263: (231)] as the Pirahã strategy to embed an infinitive complement to a matrix clause. The nominalization, thus, would be used as an strategy for subordination in such cases.

(9) koxóí   soxóá   xibíib-i-haí   tiobáhai   biío
     Koxóí   already   order-PROX-RELATIVE CERT child   grass
     kai-sai
     do-NOMINALIZER
     'Koxóí already ordered the child to cut the grass'

Reducing the valency of verbs (i.e transforming them in nominal expressions through a nominalizer) as an strategy to embed them into other clauses is a very common strategy cross-linguistically (e.g. the language Quechua, cf. Lefebvre & Muysken, 1987). However, there is an interesting case in Pirahã, which involves

the nominalization *gai-sai* 'say-NOMINALIZER', used in reported speech[1]. See (10) [Source: Salles fieldwork, 2014].

(10) hi   gai-sai   ti pioahai ogai
      3   say-NOMINALIZER 1   guaraná want
      'He said: I want guaraná'

Given that in (9), the nominalization process targets the embedded predicate, (10) is rather unexpected, as the nominalization process targets the main predicate. Everett (1986) offers a translation for *hi gaisai* constructions as 'His saying', possessive construction. This is a plausible analysis, assuming the main verb to be nominalized. In 2005, Everett maintains the translation of *gai-sai* as a possessive construction, but argues that the following clause is not a complement of the nominalized predicate, being rather syntactically juxtaposed to it. Thus, in his recent analysis, the author takes these to be cases of parataxis[2]. If so, (10) is a case of direct speech, as the translation indicates.

Data from our fieldwork, however, suggests that this issue needs to be more carefully investigated. In (10), the referent of the pronoun *ti* '1st person' is not the narrator, who uttered the sentence, but the pronoun *hi* '3rd person', to whom the speech is attributed. This suggests that (10) is indeed a case of direct speech. Direct speech *gai-sai* is produced both by young and old speakers in our fieldwork. See, however, (11) [Source: Salles fieldwork, 2014].

(11) RS: Ioihoi gai-sai   maaga   pioahai xogai
               Ioihoi say-NOMINALIZER Maaga   guaraná   want
               'Ioihoi said Maaga wants guaraná'

      Maaga: Ioihoi xi   gai-sai   ti   pioahai xogai
               Ioihoi 3FEM   say-NOMINALIZER 1   guaraná want
               'Ioihoi said I want guaraná'

---

[1]Everett (2009) denies his own work (1983; 1986), rather analyzing *-sai* as a marker of old information.

[2] As mentioned in chapter 3 (footnote 16), juxtaposition is not always synonymous with parataxis, or lack of embedding. See, for instance, Palancar (2012) for clausal juxtaposition and subordination in Otomi.

The dialogue in (11), however, indicates that with *gai-sai* constructions, indirect speech is also possible. Notice that, when repeating back the sentence to me, Maaga Pirahã, a male 20 year old speaker, substitutes *Maaga* for *ti*, leaving no doubt that the referent of *ti* is the reporter of the utterance, i.e. himself. Another example is given in (12) [Source: Salles fieldwork, 2014], where the pronoun *xi*, in the second clause, refers back to the subject of the matrix clause. This co-reference would not be possible in direct speech.

(12)   ioihoi   xi   gai-hai   xi   pioahai   xogai
        Ioihoi   3   say-NOMINALIZER 3   guaraná   want
        'Ioihoi said she wants guaraná'

In our fieldwork, we found out, therefore, that *gai-sai* constructions can trigger direct and indirect speech interpretations. These two interpretations, however, seem to be available for young speakers only as adults and elders resisted the indirect interpretation. Therefore, there might be an intergenerational ongoing grammatical change in Pirahã. In (13), for example, produced by Kobio Pirahã, a male 50-year-old speaker, the possessive pronoun *ti* '1st person' is interpreted by the speaker as obligatorily referring back to *Migixoi* [Source: Salles fieldwork, 2014].

(13)   migixoi   xi   gai-sai   maxaha   ti   maosai   kobiai apisai
        Migixoi   3   say-NOMINALIZER beautiful   1   clothes   white   shirt
        'Migixoi, her saying: "My white clothing, shirt, (is) beautiful".'

Kobio shows at least a preference for direct speech. Even when he was just repeating sentences I uttered in order to correct me, he would use direct speech. (14) is a clear example of it. I uttered (14a) and asked him if he could say that about his wife. He answered me back with (14b). We can infer from the conversational context, that he does not want to say his wife's clothes are old, thus, he attributes to me the responsibility for the utterance 'Your wife's clothes are very old'.

(14)   a.   RS: ti ibaisi maosai toio-koi

1 spouse clothes old-INTENSIFIER

'My wife's clothes (are) very old'

b.   Kobio: xai   xi   gai-sai   migixoi kobio

COPULA 3 say-NOMINALIZER Migixoi   Kobio

niai ibaisi   maosai   toióxoí-koí   xi

2   wife   clothes old-INTENSIFIER 3

maosai

clothes

'Yes, her saying, Migixoi: Kobio, your wife's clothes are very old, her clothes'

This preference for direct speech was also shown by other older speakers, such as Capixaba (Hiahoai), Dudu Pirahã, and Domingos Pirahã. What is the nature of this preference of older speakers'? Our hypothesis is that this constraint is not syntactic, but related to evidentiality. Many languages have overt morphology to express whether the information they are reporting is something they saw/hear themselves or heard from others[3]. Therefore, Pirahã older speakers could be using direct speech to mark evidentiality, avoiding indirect speech not to compromise with the truth validity of the sentences they are uttering. Thus, if there is a ban on sentential subordination in the grammar of Pirahã elders, this is an interface constraint, related to evidentiality. This needs to be carefully tested. However, if it is on the right track, our data suggests that young speakers are not subject to this constraint anymore.

At any rate, this thesis has shown that there is no general ban on self-embedding in Pirahã. Thus, the cultural principle proposed by Everett is pointless. This is an expected conclusion, given that it is not at all clear how culture could interfere in the mechanisms (e.g. Merge) internal to Grammar.

---

[3]See Aikhenvald (2003) for the typology of evidentials cross-linguistically.

5
References

Abney, S. (1987). *The English Noun Phrase in Its Sentential Aspect*, Doctoral dissertation, Massachusetts Institute of Technology, Cambridge.

Adger, D. (2003). *Core Syntax: A minimalist approach*. Oxford: Oxford University Press.

Aikhenvald, A. (2003). Evidentiality in typological perspective*. In Aikhenvald, A and R. M. W. Dixon (eds.) *Studies In Evidentiality*. John Benjamins, Amsterdam, Netherlands. 1-32.

Amaral, L.; M. Maia; C. Rodrigues; T. Roeper, F. Sandalo & G. Romling Silva (forthcoming). A pilot study on Pirahã Postpositional phrases. In: Amaral, L., M. Maia, A. Nevins & T. Roeper. *Recursion and Embedding in Brazilian Languages and Beyond*. Cambridge: Cambridge University Press.

Amoroso, M. (2009). Mura. In Instituto Socioambiental (ISA)/Povos Indígenas no Brasil. *Enciclopédia dos Povos Indígenas no Brasil*. Retrieved from: http://indios.org.br/pt/povo/mura/print

Arnauld, A. & C. Lancelot. (1975). *General and rational grammar: The Port-Royal Grammar*. (Edited and translated by Jacques Rieux and Bernard E. Rollin.) The Hague: Mouton. [Originally published, 1660, Paris: Pierre Le Petit.].

Arsenijevic, B. & W. Hinzen. (2012). On the Absence of X-within-X Recursion in Human Grammar. *Linguistic Inquiry*, 43, 423-440.

Axiomatic method. P.S. Novikov (originator), *Encyclopedia of Mathematics.* URL:
http://www.encyclopediaofmath.org/index.php?title=Axiomatic_method&oldid=17770

Bach, E. (1967). Have and be in English syntax. *Language*, 43, 462-85.

Bar-Hillel, Y. (1953). On Recursive Definitions in Empirical Science. *11th International Congress of Philosophy*, vol. 5, 160–165.

Benveniste, E. (1971). *Problems in general linguistics*. Translated by Mary Elizabeth Meek, Coral Cables, Florida, University of Miami Press.

Bhatt, R. (1997). Obligation and possession. In: H. Harley. (ed.). *The proceedings of the MIT roundtable on argument structure,* MITWPL, Cambridge, MA.

Bhatt, R. (2003). Locality in correlatives. *Natural Language and Linguistic Theory,* 21, 485–541.

Bolhuis, J.J.; I. Tattersall; N. Chomsky; & R. C. Berwick. (2014). How could language have evolved? *PLoS Biology*, 12, issue 8.

Boškoviċ, Ž. (2008). What will you have DP or NP?. In: *Proceedings of NELS 37*. 1015114. GLSA, University of Massachusetts, Amherst.

Cable, S. (2009). The Syntax of the Tibetan Correlative. In: Liptak, A. (ed) *Correlatives Cross Linguistically*. John Benjamins: Amsterdam.

Cascudo, L. C. (2002). *Geografia dos mitos brasileiros*. Third edition. Global: São Paulo.

Cardinaletti, A. (2004). Towards a Cartography of Subject Positions. In: Rizzi, L. (ed.). *The structure of IP and CP*. The Cartography of Syntactic Structures, vol. 2, New York: Oxford University Press, 115-165.

Chomsky, N. & M. Halle. (1968). *The sound pattern of English.* New York: Harper and Row. xiv, 470 pages. Reprinted 1991, Boston: MIT Press.

Chomsky, N. (1975[1955]). *The Logical Structure of Linguistic Theory*. The MIT Press, Cambridge, Massachusetts.

Chomsky, N. (1956). Three models for the description of language. *IRE Transactions of Information Theory* IT-2, vol. 3, 113–124.

Chomsky, N. (1957). *Syntactic Structures*, The Hague/Paris: Mouton

Chomsky, N. (1965). *Aspects of the Theory of Syntax*, Cambridge, Massachusetts: MIT Press

Chomsky, N. (1970a). Remarks on Nominalization. In Jacobs, R. A. and Rosenbaum, P. S. (eds.). *Readings in English Transformational Grammar*. Boston: Ginn. 184-221.

Chomsky, N. (1970b). Deep structure, surface structure and semantic interpretation. In: Jakobson, R. & S. Kawamoto (orgs.). *Studies in general and oriental linguistics presented to Shiro Hattori on the ocasion of his sixthiest birthday*. Tokyo, TEC Co. Ltd.

Chomsky, N. (1981). *Lectures on Government and Binding.* Dordrecht: Foris.

Chomsky, N. (1986). *Knowledge of Language: Its Nature, Origin, and Use*. New York: Praeger.

Chomsky, N. (1993). A minimalist program for linguistic theory. In: Hale, K. & S. J. Keyser (eds.). *The view from Building 20.* MIT Press, Cambridge, MA, 1-52.

Chomsky, N. (1995). *The Minimalist Program*, Current studies in linguistics 28, MIT Press, Cambridge, MA.

Chomsky, N. (2000). Minimalist inquiries. In: Martin, R.; D. Michaels & J. Uriagereka (eds.). *Step by step.* Cambridge: MIT Press. 89-155.

Chomsky, N. (2001). Derivation by phase. In: Kenstowicz, M. (ed.) *Ken Hale: A life in language.* Cambridge: MIT Press. 1-52.

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

Chomsky, N. (2007). Approaching UG from below. In: Sauerland, U. & H. Gärtner (eds.). *Interfaces + Recursion = Languages?* Chomsky's minimalism and the view from syntax-semantics. Berlin: Mouton de Gruyter. 1-30.

Church, A. (1936). An unsolvable problem of elementary number theory. *American Journal of Mathematics* 58, 345–363.

Cinque, G. (2010). *Towards an unified theory of relative clauses.* Talk given at University of Essex.

Comrie, B. (1989). *Language universals and linguistic typology.* 2nd edition. Chicago: University of Chicago.

Corballis, M. C. (2007). Recursion, language, and starlings. *Cognitive Science*, 31, 697–704.

Culicover, P. & R. Jackendoff. (2005). *Simpler Syntax.* Oxford University Press: Oxford. 301-338, 473-491.

Darwin, C. (1882). *The descent of man, and selection in relation to sex.* London: John Murray

Davies, J. (1981). Kobon. *Lingua Descriptive Studies*, 3. Amsterdam: North-Holland Publishing Company.

Dedekind, R. J. W. (1932[1888]). Was sind und was sollen die Zahlen? reprinted in Fricke, R.; E. Noether & O. Ore (eds.). (1930–1932), *Gesammelte mathematische Werke*. Vieweg: Braunschweig; III: 335–391.

Dékány, E. & M. den Dikken. (2015). Recursion Restricted. Talk presented in *The 12th International Conference on the Structure of Hungarian (ICSH12).* Leiden Univeristy, The Netherlands, Leiden University Centre for Linguistics (LUCL).

Dikken, M. den. (2013) *The morphosyntax of (in)alienably possessed noun phrases*: The Hungarian contribution. New York: CUNY.

Di Sciullo, A.-M. (2014). Minimalism and I-Morphology. In: Kosta, P.; L. Schurcks & T. RadevBork (eds.). *Minimalism and Beyond: Radicalizing the Interfaces.* Amsterdan: John Benjamins. 267–286.

Dryer, M. (1980). The positional tendencies of sentential noun phrases in universal grammar. *Canadian Journal of Linguistics* 25. 123–95.

Dryer, M. (2007). Word Order. In: Shopen, T. (ed.). *Language Typology and Syntactic Description* (3 vols.). 2nd ed. Cambridge: Cambridge University Press.

Everett, D. (1983). *A lingua pirahã e a teoria da sintaxe.* Doctoral dissertation, State University of Campinas, Brazil.

Everett, D. (1986). Pirahã. In: Derbyshire, D. C. & G. K. Pullum (eds.). *Handbook of Amazonian languages.* Berlin: Mouton de Gruyter. 200–326.

Everett, D. **(**2005). Cultural constraints on grammar and cognition in Piraha: Another look at the Design Features of human language. *Current Anthropology* 46. 621-646.

Everett, D. (2009). Pirahã culture and grammar: A response to some criticisms. *Language* 85. 405-442.

Everett, K. M. (1998). Acoustic Correlates of Stress in Pirahã. *The Journal of Amazonian Languages*: 104-162.

Fitch, W. T. & M. HAUSER. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science* 303. 377–380

Fitch, W. T.; M. D. Hauser & N. Chomsky. (2005). The evolution of the language faculty: clarifications and implications. *Cognition* 97, 179-210.

Fiva, T. (1984). NP-Internal Chains in Norwegian. *Working Papers in Scandinavian Syntax,* 14.

Frank, M. C.; D. L. Everett; E. Fedorenko & E. Gibson. (2008). Number as a cognitive technology: Evidence from Pirahã language and cognition. *Cognition,* 108. 819–24.

Freeze, R. (1992). Existentials and other locatives. *Language,* 68. 553–595.

Galilei, G. (1632). *Dialogo di Galileo Galilei Linceo...dove ne i congressi di quattro giornate si discorre sopra i due massimi sistemi del mondo Tolemaico, e Copernicano; proponendo indeterminatemente le ragione filosofiche e naturale tanto per l'una, quanto per l'altra parte.* G.B. Landini, Florence.

Gazdar, G.; E. Klein; G. K. Pullum & I. A. Sag. (1985). *Generalized phrase structure grammar.* Cambridge, MA: Harvard University Press.

Gentner, T. Q.; K. Fenn, D. Margoliash & H. C. Nusbaum. (2006). Recursive syntactic pattern learning by songbirds. *Nature* 440. 1204–1207.

George, L. M. & J. Kornfilt. (1981). Finiteness and boundedness in Turkish. In: Heny, F. W. (ed.). *Binding and filtering.* Cambridge, MA: MIT Press. 105–127.

Gödel, K. (1931/1986). Über formal unentscheidbare sätze der Principia Mathematica und verwandter systeme. In: Feferman, S.; J. W. Dawson; S. C. Kleene; G. H. Moore; R.M. Solovay & J. Van Heijenoort. *Kurt Gödel: Collected Works*, Vol. I, Publications 1929–1936. Oxford: Oxford University Press. 144–195.

Gödel, K. (1934/1986). On undecidable propositions of formal mathematical systems. In: Feferman, S.; J. W. Dawson; S. C. Kleene; G. H. Moore; R.M. Solovay & J. Van Heijenoort. *Kurt Gödel: Collected Works*, Vol. I, Publications 1929–1936. Oxford: Oxford University Press. 346–371.

Goldberg, A. (1995). *Constructions: A construction grammar approach to argument structure.* Chicago: University of Chicago Press.

Goldberg, A. (2006). *Constructions at work: The nature of generalization in language.* Oxford: Oxford University Press.

Gonçalves, M. A. (1993). *O significado do nome: Cosmologia e nominação entre os Pirahã.* Rio de Janeiro: Sette Letras.

Gonçalves, M. A. (2000). Pirahã: *Indigenous peoples in Brazil*: ISA. Internet site: http://pib.socioambiental.org/en/povo/piraha. Accessed in August 11, 2015.

Gonçalves, M. A. (2001). *O mundo inacabado*: Ação e criação em uma cosmologia amazônica: Etnografia Pirahã. Rio de Janeiro: Editora da UFRJ.

Gordon, P. (2004). Numerical cognition without words: Evidence from Amazonia. *Science,* 306. 496–99.

Gordon, M. (2005). A perceptually-driven account of onset-sensitive stress. *Natural Language and Linguistic,* Theory 23: 595-653.

Grassmann, H. (1861). *Lehrbuch der Mathematik fur hohere Lehrenstalten,* Band 1. Berlin: Enslin.

Greenberg, J. H. (1966). *Universals of language.* 2nd edition. Cambridge, MA: MIT Press.

Grohmann, K. K. & L. Haegeman. (2003). Resuming reflexives. *Nordlyd* 31.

Haegeman, L. (1998). *Notes on the Anaphor-Agreement Effect in the DP*. Manuscript, Université de Genève.

Haegeman, L. (2001). DP Periphery and Clausal Periphery: Possessor Doubling in West Flemish? Relating Nominal periphery to Clausal Periphery. In: Adger, D.; C. de Cat & George Tsoulas (eds.). *Peripheries.* Kluwer, Dordrecht.

Haegeman, L. (2003). The External Possessor Construction in West Flemish. In: Coene, M.; d'Hulst, Y. & Tasmowski, L. (eds.). *From NP to DP, Vol. II: The Expression of Possession.* John Benjamins, Amsterdam. 221-255.

Hale, K. L. (1975). Gaps in grammar and culture. In: Kinkade, M. D.; K. L. Hale & O. Werner (eds.). *Linguistics and Anthropology: In honor of C. F. Voegelin.* Lisse: The Peter de Ridder Press. 295-315.

Hauser, M., N. Chomsky and T. Fitch. (2002) The faculty of language: What is it, who has it, and how did it evolve? *Science,* 298. 1569–1579.

Heinrichs, A. (1964). Os Fonemas de Mura-Pirahã. In: Boletim do Museu Paraense Emílio Goeldi. Vol. 21.

Hermon, G. (1985). *Syntactic modularity*. Dordrecht: Foris.

Hinzen, W. (2006). *Mind design and minimal syntax.* Oxford: Oxford University Press.

Hinzen, W. (2014). Recursion and Truth. In: Speas, M. & T. Roeper. Recursion: Complexity in Cognition. Springer International Publishing, Switzerland. 113-137.

Hockett, C.F. (1960). The origin of speech. *Scientific American*, 203. 88-96.

Hornstein, N.; S. Rosen & J. Uriagereka. (1996). Integral existentials. In: Camacho, J.; L. Choueiri & M. Watanabe (eds.). *Proceeedings of the Fourteenth West Coast Conference on Formal Linguistics*. Stanford, California: CSLI Publications. [Distributed by Cambridge University Press.]

Intituto Brasileiro de Geografia e Estatística - IBGE. (2010). *Censo demográfico 2010, Características Gerais dos Indígenas, Resultados do Universo.* Retrieved from:
http://www.teleios.com.br/wp-content/uploads/2012/08/Caracter%C3%ADsticas-Gerais-dos-Ind%C3%ADgenas-IBGE-ref-Censo-2010.pdf. Acessed in August 12, 2015.

Jackendoff, R., & S. Pinker. (2005). The nature of the language faculty and its implications for evolution of language (Reply to Fitch, Hauser & Chomsky). *Cognition*, *97*(2), 211-225.

Kaminski, J.; J. Call & J. Fischer. (2004). Word learning in a domestic dog: evidence for "fast mapping". *Science,* 304. 1682-1683.

Kayne, R. S. (1993). Toward a modular theory of auxiliary selection. *Studia Linguistica,* 47. 3-31.

Kelepir, M. (2001). *Topics in Turkish syntax: Clausal structure and scope*. Cambridge, MA: MIT dissertation.

Kennedy, H. (ed.). (1973). Selected works of Giuseppe Peano. Toronto: University of Toronto Press.

Kleene, S. (1952). *Introduction to Metamathematics.* Amsterdam: North-Holland.

Kluender, K. R., R. L. Diehl & P. Killeen. (1987). Japanese quail can learn phonetic categories. *Science,* 237. 1195–1197.

Koptjevskaja-Tamm, M. (2001). Adnominal Possession. In M. Haspesmath, E. König; W. Oesterreicher and W. Raible (eds.). *Language Typology and Language Universals* (LTLU). Berlin/New York: Walter de Gruyter. Chapter 72. 960-970.

Kornfilt, J. (2001). On the syntax andmorphology of clausal complements and adjuncts in the Turkic languages. In: Bisang, W. *Aspects of typology and universals.* Berlin: Akademie. 63–82.

Krause, C. (2000a). Anmerkungen zum pränominalen Genitiv im Deutschen. *Von der Philologie zur Grammatiktheorie: Peter Suchsland zum 65. Geburtstag.* ed. by Josef Bayer and Christine Roemer. Tübingen: Niemeyer. 79–96.

Krause, C. (2000b). On an (in-)visible property of inherent case. *North Eastern Linguistic Society,* 30. 427–442.

Kuhl, P. K. & J. D. Miller. (1975). Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. *Science,* 190. 69–72.

Kuhl, P.K. & D. M. Padden. (1982). Enhanced discriminability at the phonetic boundaries for the voicing feature in macaques. *Perception and psychophysics*, 32 (6). 542-550.

Landau, E. (1966). Foundations of Analysis. *The arithmetic of whole, rational, irrational and complex numbers.* Chelsea, New York NY. Translated from the German by F. Steinhardt.

Lefebvre,C. & P. Muysken. (1987). *Mixed categories: Nominalizations in Quechua.* (Studies in natural language and linguistic theory.) Dordrecht: Kluwer.

Liberman, A. M.; K. S. Harris; H. S. Hoffman & B. C. Griffith. (1957). *The discrimination of speech sounds within and across phoneme boundaries. Journal of Experimental Psychology,* 54. 358-368.

Liberman, A. M.; F. S. Cooper; D. P. Shankweiler & M. Studdert-Kennedy. (1967). Perception of the speech code. *Psychological Review*, 74. 431-461.

Lobeck, A. & K. Denham. *Lingusitics for Everyone.* Boston, MA. Cengage. Second Edition.

Longobardi, G. (1994). Reference and proper names. *Linguistic Inquiry*, 25. 609-665.

Longobardi, G. (2005). Towards a unified grammar of reference. *Zeitschrift fur Sprachwissenchaft*, 24, issue 1. 5-44.

Maia, M.; A. França; A. Gesualdi; A. Lage & C. Oliveira. (forthcoming). The processing of PP embedding and coordination in Karaja and in Portuguese. In: Amaral, L.; M. Maia; A. Nevins & T. Roeper. *Recursion and Embedding in Brazilian Languages and Beyond.* Cambridge: Cambridge University Press.

Nagel, E. & J. R. Newman. (1957). *Gödel's Proof.* NYU Press.

Nascimento, M. (2013). *Marcadores de evidencialidade, tempo, modo e aspecto em Kaingang.* Master thesis. Federal University of Rio de Janeiro.

Navarro, E. A. (2011). *Curso de Língua Geral (Nheengatu ou Tupi Moderno). A língua das origens da civilização amazônica.* São Paulo: edição do autor.

Neves, M. H. M. (1981). *A Teoria Lingüística em Aristóteles.* In: ALFA. São Paulo.

Neves, M. H. M. (2002). *A gramática – história, teoria e análise, ensino.* São Paulo: Editora UNESP.

Nevins, A.; D. Pesetsky & C. Rodrigues. (2009a). Piraha Exceptionality: a Reassessment. *Language,* 85(2). 355-404.

Nevins, A.; D. Pesetsky & C. Rodrigues. (2009b). Evidence and Argumentation: a Reply to Everett. *Language,* 85.3. 671-681.

Nichols, J. (1988). On alienable and inalienable possession. In: Shipley, W. (ed.). *In honor of Mary Haas.* Berlin: Mouton de Gruyter. 557-609.

Nimuendaju, C. (1948). The Mura and Pirahã. Bulletin 143: *Handbook of South American Indians* 3. Washington, DC: US Government Printing Office. 255–69.

Nimuendaju, C. (1982). As tribos do Alto Madeira. *Textos Indigenistas*. São Paulo: Loyola.

Odifreddi, P. (1989). *Classical recursion theory.* North-Holland, Amsterdan.

Odifreddi, P. & S. B. Cooper. (2012). Recursive Functions. In: *The Stanford Encyclopedia of Philosophy* (Fall 2012 Edition). Zalta, E. N. (ed.). URL = <http://plato.stanford.edu/archives/fall2012/entries/recursive-functions/>. Accessed in August 11, 2015.

Palancar, E. L. 2012. Clausal juxtaposition and subordination: New light from Otomi. In: Gast, V. & H. Diessel (eds.). *Clause linkage in cross-linguistic perspective: Data-driven approaches to cross-clausal syntax*. Berlin: De Gruyter Mouton. 37–76.

Payne, T. E. (1997). Voice and valence adjusting operations. In: Payne, T. E. *Describing Morphosyntax: a guide for field linguists*. Cambridge: Cambridge University Press. 169-222.

Peano, G. (1958[1889]). Arithmetices principia nova methodo exposita. *Opera Scelte,* 2. 20–55 (Reprinted in Unione Matematica Italiana, 1957–1959).

Peirce, C. S. (1881). On the Logic of Number. *American Journal of Mathematics,* 4, number 1. 85–95.

Pena, H. F. (1909). Fala dirigida à assembleia legislativa provincial do Amazonas. Relatórios da presidência da província do Amazonas. Rio de Janeiro: Tipologia do Jornal do Commercio.

Pepperberg, I. M. (2008). *Alex and Me: How a Scientist and a Parrot Discovered a Hidden World of Animal Intelligence — And Formed a Deep Bond in the Process.* New York, Collins.

Picallo, C. (1991). Nominals and Nominalizations in Catalan. *Probus,* 3.3. 279-316.

Pinker, S., & R. Jackendoff. (2005). The faculty of language: what's special about it? *Cognition, 95*(2). 201-236.

Plungian, V. (1995). *Dogon.* Munich: Lincom Europa.

Ritter, E. (1991). Two Functional Categories in Noun Phrases: Evidence from Modern Hebrew. In: Rothstein, S. (ed.). Perspectives on Phrase Structure. *Syntax and Semantics,* 26. Academic Press, New York. 37-6.

Rodrigues, C.; R. Salles & F. Sandalo. (forthcoming). Word order in obligatory control: evidence for self-embedding in Pirahã. In: Amaral, L., M. Maia, A. Nevins & T. Roeper. *Recursion and Embedding in Brazilian Languages and Beyond*. Cambridge: Cambridge University Press.

Roeper, T. & W. Snyder. (2005). Language learnability and the forms of recursion. UG and external systems: Language, brain and computation In: di Sciullo, A. M. (ed.). *Linguistik aktuell,* 75. Amsterdam: John Benjamins. 155–69.

Roeper, T. (2011). The acquisition of recursion: how formalism articulates the acquisition path. *Biolinguistics,* 5(1). 56-86.

Ross. J. (1967). *Constraints on variables in syntax*. PhD dissertation, MIT.

Sandalo, F. (1989). *Aspectos da língua pirahã*. Master Thesis in Linguistics. State University of Campinas. Campinas, SP.

Sauerland, U. 2010. Experimental evidence for complex syntax in Pirahã. Ms, ZAS Berlin Retrieved from: [http://www.zas.gwz-berlin.de/574.html?&L=1]. Acessed on August 12, 2015.

Sheldon, S. N. (1974). Some Morphophonemic and Tone Pertubation Rules in Mura-Pirahã. *International Journal of American Linguistics*, vol. 40(3). 279-282.

Sheldon, S. N. (1988). Os sufixos verbais Mura-Pirahã. *Série Linguística* 9.146–75.

State University of Amazonas. (2013). Nova cartografia social da Amazônia: Hiaitihyii aba aigioo Pirahã ogiaaga bigi apatiso kaoai bogi aoaga micihiai akasaba apagaiso no Sul do Amazonas. Project coordinator: Almeida, A. W. B. Research team Araújo, J. N. [et al.]. Manaus: UEA Edições.

Stenzel, K. S. (2013). *A Reference Grammar of Kotiria (Wanano).* Lincoln: University of Nebraska Press, 2013.

Stowell, T. (1981). *Origins of phrase structure*. Cambridge, MA: MIT dissertation.

Szabolcsi, A. (1983). The possessor that ran away from home. *The Linguistic Review,* 3. 89–102.

Szabolcsi, A. (1994). The noun phrase. In: Kiefer, F. & K. É. Kiss (eds.). *The syntactic structure of Hungarian. Syntax and Semantics,* 27. San Diego: Academic Press. 179–274.

Tomalin, M. (2006). *Linguistics and the Formal Sciences: The Origins of Generative Grammar*. Cambridge University Press, Cambridge.

Topintzi, N. (2004). Moraic Onsets and WSP Partition in Pirahã. In: Daskalaki, E.; N. Katsos; M. Mavrogiorgos & M. Reeve (eds.). *CamLing 2004, Proceedings of the Second Postgraduate Conference in Language Research.* Cambridge: Cambridge Institute of Language Research. 211-218.

Turing, A. (1936). On computable numbers, with an application to the entscheidungsproblem. *Proceedings of the London Mathematical Society,* 42. 230–265.

van der Hulst, H.G. (2010). Re Recursion. In: van der Hulst, H. (ed.). *Recursion and Human Language.* Berlin: Mouton de Gruyter. i-xv.

van Heijningen, C.A.A.; J. de Visser; W. Zuidema; C. ten Cate. (2009). Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species. *Proceedings of the National Academy of Sciences*, 106, issue 48. 20538-20543.

von Humboldt, W. (1886). *Ober die Verschiedenheit des Menschlichen Sprachbaues.* Berlin.

Watumull, J.; M. D. Hauser; I. G. Roberts & N. Hornstein. (2014). On recursion. *Frontiers in Psychology,* 4, article 1017.

Appendix

Photo 1 – Map of the high Maici villages drawn by Pirahãs in the Piquiá village, with the help of cartographers from the State University of Amazonas - UEA. [Source: UEA, 2013].
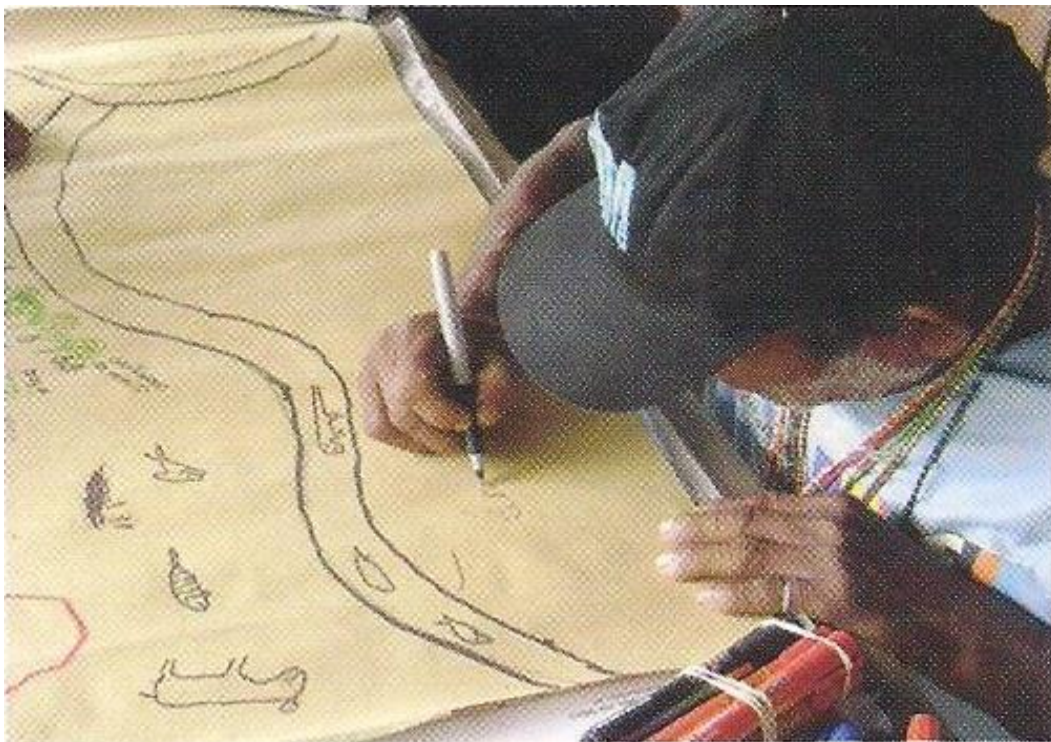

Photo 2 – A Pirahã man drawing the map in photo 1. [Source: UEA, 2013].

Photo 3 – Pirahãs from Piquiá exhibiting their map, which was the basis for the high Maici part of the map in Figure 1 (chapter 1 (section 1.1)). [Source: UEA, 2013].

Photo 4 – Pirahãs from the lower Maici exhibiting their map, the basis for the lower Maici part of the map in Figure 1 (chapter 1 (section 1.1)). [Source: UEA, 2013].

Photo 5 – A woman wearing a traditional Pirahã necklace. [Source: Salles fieldwork, 2014].

Photo 6 – Pirahã traditional dance configuration in the festivals. [Source: Salles fieldwork, 2013].

Photo 7 – Woman making a necklace with fishing line and beads. [Source: Salles fieldwork, 2014].

Photo 8 – Example of drawing used for elicitation. I drew the first man and explained to my informant it was him, Pihoio Pirahã. Then I drew a canoe and told him it was his canoe. I drew a motor on his canoe and explained that his canoe had a small motor. Then, I did the same for the drawing of Kapoogo, one of his friends, except that for Kapoogo, the canoe's motor was big. After the drawing was complete and the context set, I asked him whether (i a) was correct; he repeated the sentence back. Then I asked him (i b), inverting the word order; he corrected me, repeating (i a). Whenever I tried to invert the word order, he would repeat (i a) to me.

(i)   a.   Pihoio agaoa motohoi koihi
                Pohoio canoe motor small
                'Pihoio's canoe's motor is small'

        b.   motohoi agaoa Pihoio koihi
                motor canoe Pihoio small

Photo 9 – Example of picture used for elicitation. My mother-in-law is beside my husband, wearing black clothes. I showed the picture and asked my informants to correct my sentence if I were wrong. The elicited sentence was (i).

(i)   Migixoi ibaisi maixi maosai kopaixai
        Migixoi spouse mother clothes black
        'Migixoi's husband's mother's clothes are black'