



Universidade de Brasília – UnB

Instituto De Letras

Departamento de Linguística, Português e Línguas Clássicas

Programa de Pós-Graduação em Linguística

**INDEXAÇÃO E CLASSIFICAÇÃO NOMINAL DE ARGUMENTOS
VERBAIS EM LÍNGUAS INDÍGENAS DA AMÉRICA DO SUL**

Rodrigo do Prado Sateles

Brasília,

2019

Rodrigo do Prado Sateles

**INDEXAÇÃO E CLASSIFICAÇÃO NOMINAL DE ARGUMENTOS
VERBAIS EM LÍNGUAS INDÍGENAS DA AMÉRICA DO SUL**

Dissertação apresentada ao Programa de Pós-Graduação em Linguística do Departamento de Linguística, Português e Línguas Clássicas – LIP do Instituto de Letras - IL da Universidade de Brasília – UnB como requisito parcial para a obtenção do título de Mestre em Linguística.

Orientador: Prof. Dr. Thiago Costa Chacon

Brasília,

2019

Ficha catalográfica elaborada automaticamente,
com os dados fornecidos pelo(a) autor(a)

DR696i DO PRADO SATELES, RODRIGO
INDEXAÇÃO E CLASSIFICAÇÃO NOMINAL DE ARGUMENTOS VERBAIS
EM LÍNGUAS INDÍGENAS DA AMÉRICA DO SUL / RODRIGO DO PRADO
SATELES; orientador THIAGO COSTA CHACON. -- Brasília, 2019.
81 p.

Dissertação (Mestrado - Mestrado em Linguística) --
Universidade de Brasília, 2019.

1. Línguas Indígenas da América do Sul. 2. Indexação de
argumentos. I. COSTA CHACON, THIAGO, orient. II. Título.

Rodrigo do Prado Sateles

**INDEXAÇÃO E CLASSIFICAÇÃO NOMINAL DE ARGUMENTOS VERBAIS
EM LÍNGUAS INDÍGENAS DA AMÉRICA DO SUL**

Dissertação apresentada ao Programa de Pós-Graduação em Linguística do Departamento de Linguística, Português e Línguas Clássicas – LIP do Instituto de - IL Letras da Universidade de Brasília – UnB como requisito parcial para a obtenção do título de Mestre em Linguística.

Prof. Dr. Thiago Costa Chacon – LIP/UnB
(Orientador e Presidente da Banca)

Prof. Dra. María Alejandra Regúnaga
(Membro Externo/ Universidad
Nacional de la Pampa - Argentina)

Prof. Dr. Dionei Moreira Gomes
(Membro Interno/ PPGL - UnB)

Prof. Dr. Antônio Augusto Souza Melo
(Membro Suplente/ PPGL - UnB)

Dedico este trabalho inteiramente à minha avó Francisca (in memoriam)

Rerum omnium magister usus

AGRADECIMENTOS

Agradeço em primeiro lugar ao meu orientador Thiago Chacon, pelos quatro anos de muita paciência comigo, e pelo imenso apoio na minha formação como linguista. Por ser, sem dúvida, umas das principais fontes de inspiração na minha carreira acadêmica e ter sido amigo, mesmo do outro lado do Atlântico, nos momentos difíceis que quase me fizeram desacreditar no valor da pesquisa e da vida acadêmica.

À minha mãe, por me compreender e apoiar incondicionalmente em todas as minhas decisões. Um luto, um câncer, uma depressão e uma dissertação ao mesmo tempo mostraram indubitavelmente que Deus está nos fortalecendo a cada dia. E muito já estamos conseguindo superar, graças ao amor, a fé e a perseverança que nos une nessa vida.

Agradeço as minhas tias Maria, Cica, Cacá e Maninha por compreenderem minha intensa dedicação a esse trabalho; por não terem deixado faltar nada na mesa de casa, e terem cuidado bem da minha mãe enquanto eu estava muito ocupado com essa pesquisa e trabalho. Vibramos por essa conclusão.

Dedico um agradecimento especial aos professores Walkíria Praça, Dionei Gomes e Augusto Mello, por me passarem tanto de seus conhecimentos e se mostrarem tão humanos e atenciosos.

Ao valioso suporte financeiro da CAPES no ano de 2018, pois sem esse majestoso fomento a pesquisa científica no nosso país não teria a mesma qualidade, e portanto, reconhecimento.

Aos meus colegas de PBSL e mestrado, que me mostraram ser amigos de verdade: Miguel Barros (me presenteou com esse notebook com o qual escrevi 100% da dissertação), Gabriel Gabito, Murilo Barros, Eduardo Melo e Kaoru.

Ao meu amigo de sempre Marcos Vinícius, que há 10 anos me fez conhecer o sentido da verdadeira amizade. Foi suporte antes, durante e o será após o mestrado.

Aos colegas Baniwa Arthur e Augusto, por me fazerem perceber a beleza e importância da diversidade cultural do nosso Brasil. Desejo muito poder revê-los em breve.

Aos meus professores colegas de trabalho do UnB Idiomas que mostraram explicitamente apoio a mim nessa fase tensa de pós-graduação: Solange, Norma, Lorena, Thays, Luciene, Lidia, Filippo Ferrari e Andrea (il pistolero).

Dedico finalmente aos professores da época de graduação no LEA-MSI, sobretudo prof. Cláudio Menezes e Claudine. Meu sucesso acadêmico passa diretamente por vocês.

RESUMO

Este trabalho dissertativo discute sob a ótica da Tipologia Linguística os fenômenos de indexação e categorização nominal nos verbos de um *corpus* representativo de línguas indígenas faladas na América do Sul. A dissertação está organizada em três capítulos, sendo o primeiro dedicado a estabelecer objetivos, justificativa, metodologia e as diretrizes que fundamentaram previamente o início da pesquisa calcada em uma análise comparativa de línguas. A partir de uma discussão teórica sobre o vasto entendimento acerca dos sistemas de classificação nominal (SCN) e de indexação argumental proposta por Haspelmath (2013) no segundo capítulo, uma amostragem filogeneticamente diversificada de línguas sul-americanas foi levantada. Analisamos quais os tipos de padrões estruturais e semânticos encontramos e discutimos de que maneira eles se aproximam ou se distanciam do que a tipologia linguística vem definindo como sistemas de classificação nominal, como gênero/ classes nominais e classificadores. Seleccionamos no terço e último capítulo um conjunto de 23 línguas que expressam no sintagma verbal alguma informação de classificação nominal. A última etapa dessa pesquisa concentrou-se em reduzir o *corpus* a apenas três línguas, das quais foi possível explorar exemplos concretos sobre classes e indexação. Contrastamos finalmente os resultados nesse *corpus* específico.

Palavras-chave: classificação nominal, indexação, gênero, classificador.

ABSTRACT

This Master's thesis discusses the phenomena of indexing and nominal categorization in the verbs of a *corpus* representative of indigenous languages spoken in South America from the perspective of linguistic typology. The dissertation is organized in three chapters, being the first dedicated to establish goals, justification, methodology and guidelines that previously grounded the beginning of the research based on a comparative analysis of languages. From a theoretical discussion on the vast understanding of the nominal classification systems (SCN) and argument indexation proposed by Haspelmath (2013) in the second chapter, a phylogenetically diversified sample of South American languages has been considered. We analyze what types of structural and semantic patterns we find and discuss how they approach or distance themselves from what linguistic typology has been defining as nominal classification systems, such as gender/nominal class and classifiers. We selected in the third and last chapter a set of 23 languages expressing in their verbal syntagma some nominal classification information. The last stage of this research focused on reducing the *corpus* to only three languages, from which it was possible to explore concrete examples of classes and indexation. We finally contrast the results in this specific *corpus*.

Key-words: nominal classification, indexation, gender, classifier

SUMÁRIO

LISTA DE ABREVIATURAS.....	11
LISTA DE TABELAS.....	13
LISTA DE FIGURAS	14
1 - INTRODUÇÃO, REFERENCIAL TEÓRICO E METODOLÓGICO.....	15
1.1 Objetivos e Metodologia da pesquisa.....	15
1.2 Justificativa	16
1.3 Referencial teórico	17
1.4 Resumo dos capítulos.....	21
2 - INDEXAÇÃO E CLASSIFICAÇÃO NOMINAL EM LÍNGUAS INDÍGENAS DA AMÉRICA DO SUL	22
1.1 O que é classificação nominal?.....	22
2.1.1 Gênero	22
2.1.2 As classes nominais.....	27
2.1.3 Os classificadores.....	27
2.1.4 Sistemas menos prototípicos	29
2.1.3 Panorama dos SCN em línguas indígenas da América do Sul	31
2.2 Indexação	34
2.2.1. A concordância em Corbett.....	35
2.3 Classificação nominal e indexação verbal nas línguas da América do Sul	39
3- ANÁLISE DE INDEXAÇÃO E CLASSIFICAÇÃO NOMINAL EM UM <i>CORPUS</i> SELECIONADO	43
3.1 Procedimento para a obtenção do <i>corpus</i>	43
3.1.1 O SCN fora da palavra verbal	44
3.1.2 O SCN nos índices verbais.....	45
3.2 Análise de um corpus trilíngue (Mundurukú, Kubeo e Tehuelche).....	50
3.2.1 Mundurukú.....	50
3.2.2 Kubeo	57
3.3 Considerações contrastivas dos SCNs no corpus trilíngue	71
CONSIDERAÇÕES FINAIS	74
REFERÊNCIAS BIBLIOGRÁFICAS	75
APÊNDICE.....	78

LISTA DE ABREVIATURAS

- 3 – terceira pessoa
- AC – acusativo
- AL – alemão
- AN – animado
- COL – coletivo
- CL - classificador
- CLV – classificador verbal
- CNV – converbo
- CONT - contável
- DET – determinante
- DUR – duração
- EXC – exclusivo
- E.P. – especificador do predicado
- F – feminino
- FF – forma completa (full form)
- FR – francês
- Func. - funcional
- Gen- geral
- IN – inanimado
- INC – inclusivo
- INT – partícula interrogativa
- IPRF – aspecto imperfeito
- LT – latim
- LONG – objeto longo com
ponta fina
- M – masculino
- MAV – marcação/marcador do argumento verbal
- M.R. – modo real
- N - neutro
- NC – classificadores nominais (*noun classifiers*)
- NFC – Nome com Função Classificadora NM – nominativo
- NMZ - nominalizador

NT(R) – neutro OBL- caso obliquo

PL - plural

PP – particípio passado PS – passado simples PSS - possessivo

PST – passado

R – relativo

R1 – relativo de contiguidade

R2 – relativo de não-contiguidade

SCN – sistema de classificação nominal S - substantivo

SG – singular

LISTA DE TABELAS

Tabela 2.1 – Atribuição de gênero em quatro línguas indo-europeias.....	23
Tabela 2.2 - Gênero vs. sistemas de classificador.....	27
Tabela 2.4 – O classificador prototípico nas línguas sul-americanas	33
Tabela 2.5 – Configurações possíveis do MAV	40
Tabela 3.1 – A indexação verbal em um <i>corpus</i> selecionado	46
Tabela 3.2 – Alinhamento em sistemas de gênero.....	49
Tabela 3.4 – Os NFC no sintagma nominal	57
Tabela 3.5 - Indexadores verbais em Mundurukú.....	57
Tabela 3.6 – Pronomes pessoais em Kubeo.....	59
Tabela 3.7 – Demonstrativos em Kubeo.....	60
Tabela 3.8 – Determinantes Possessivos em Kubeo	61
Tabela 3.9 – Flexão dos determinantes possessivos	62
Tabela 3.10 – Marcação da classificação nominal no SN do Kubeo	63
Tabela 3.11 – Paradigmas verbais de flexão para evidencialidade não-marcada (CLASSE I).....	63
Tabela 3.12 – Paradigmas verbais de flexão para evidencialidade não-marcada (CLASSE II)	64
Tabela 3.13 – Indexadores verbais em Tehuelche	71
Tabela 3.14 – Quadro contrastivo dos SCNs no <i>corpus</i> trilíngue.....	73

LISTA DE FIGURAS

Figura 1 - Diferentes sistemas da categorização linguística	26
Figura 2 – Hierarquia da classificação nominal em Kubeo	58
Figura 3 – Zonas dialetais do Tehuelche e outras línguas da Patagônia.....	65

1 - INTRODUÇÃO, REFERENCIAL TEÓRICO E METODOLÓGICO

1.1 Objetivos e Metodologia da pesquisa

A pergunta-problema planteada neste estudo é: *Como as línguas indígenas faladas na América do Sul organizam o seu sistema verbal integrando nele a marcação de argumentos verbais a partir de morfemas que classificam semanticamente esses argumentos para além da dimensão de pessoa e número?* Assim, o objetivo geral e os objetivos específicos que nos propusemos nesta pesquisa são:

- Analisar o processo da marcação dos Sistemas de Classificação Nominal (SCN) e Indexação em línguas indígenas da América do Sul sob ponto de vista tipológico.
- Realizar um mapeamento geográfico das línguas e famílias linguísticas que apresentam classificação nominal de argumentos verbais como parte da morfologia dos verbos.
 - Analisar como a classificação nominal se correlaciona com Pessoa, Número e Função Sintática do argumento.
- Fazer generalizações descritivas sobre os dados de modo a buscar correlações entre classificação nominal nos verbos e em outras categorias morfossintáticas, como nos nomes, pronomes, demonstrativos, adjetivos e possessivos.

A pesquisa nasceu com o objetivo de investigar a marcação de gênero em morfemas verbais correferentes aos argumentos do predicado em línguas da América do Sul. Isso iria complementar trabalhos mais gerais sobre o tema da classificação nominal em línguas do continente, como Regúnaga (2012), Krasnoukhova (2012), Birchall (2014), Derbyshire e Payne (1990), Payne (1987), Aikhenvald (1994), Seifart e Payne (2007).

A noção de gênero foi tomada inicialmente de modo genérico, i.e., como correspondendo à categorização semântica de argumentos verbais para além de “pessoa” e “número”. Logo foi percebido como a definição do que é gênero é problemática sob o ponto de vista tipológico, pois alguns autores entendem gênero num sentido mais específico (e.g. sistema que envolva distinção de sexo e no máximo três classes) e outros num sentido mais amplo morfossintático a partir da manifestação da concordância nominal. Gênero também é usado em oposição a sistemas de classificadores em algumas tipologias e, como gostaríamos de também incluir esses sistemas em nossa pesquisa, não nos pareceu que gênero seria um termo adequado se estivermos interessados na pergunta mais ampla, a

qual nos atemos neste trabalho, sobre como os argumentos verbais são categorizados semanticamente (ver capítulo 2).

Como metodologia para a execução da presente pesquisa, foi realizada primeiramente uma revisão bibliográfica, o que forneceu uma fundamentação da teoria tipológica da classificação nominal e indexação (ver capítulo 2). Adotou-se aqui uma abordagem tanto quantitativa como qualitativa, pois acreditamos que uma coleta de cunho tipológico requer um número relativamente expressivo de línguas que apresentam Sistemas de Classificação Nominal (SCN) (sobretudo no verbo) e, além disso, será altamente prezado o modo como essas representações podem ser descritas com base na teoria linguística.

A partir dos autores acima citados e outros, uma garimpagem de línguas sul-americanas foi realizada. À medida que foram encontrados exemplos de línguas que possuíssem marcação de gênero e/ou classificadores, uma relação foi sendo construída, de modo a poder constituir o *corpus* comparativo de trabalho. Em um momento posterior, fez-se uma verificação de quais línguas elencadas possuíam marcação SCN na palavra verbal. No Anexo apresentamos uma lista de línguas para as quais encontramos informação na literatura que indicam a presença de índices verbais que categorizam argumentos de um predicado.

Diante de uma impalpável gama de expressões linguísticas existentes no mundo, é de grande importância que para esta investigação se faça um recorte geográfico a fim de facilitar e dar mais consistência à coleta dos dados para uma análise futura. A fim de tornar o trabalho mais direcionado, selecionamos três línguas (Mundurukú, Kubeo e Tehuelche) para o procedimento analítico detalhado do SCN. Vale notar que essas línguas, por estarem na América do Sul, poderiam apresentar certos graus de convergência areal, sobretudo indiretamente. Parte das razões pela seleção de nossa amostra foi o fato de essas línguas não serem geneticamente aparentadas, estarem em regiões geográficas distantes entre si e sem relação de contato no presente e no passado até onde sabemos, de modo que nossa amostra – ainda que pequena – possa refletir um grau efetivo de variação do fenômeno sendo estudado (cf. Croft 2003, p. 19-28).

1.2 Justificativa

A escolha do tema partiu da curiosidade ao perceber que em línguas mais conhecidas e acessíveis para mim, como o inglês, o francês, o mandarim, etc. não se

percebe a categoria verbal marcada pelo traço de gênero nos verbos, salvo algumas exceções, como no caso de certas formas não-finitas, sobretudo no particípio passado (exemplo em português: A tarefa foi feita/ O bolo está assado).

A literatura que tive acesso durante minha formação na graduação apresenta a presença do traço de gênero na formação de verbos como algo bastante raro nas línguas do mundo, e nas línguas da América do Sul, a família que hipoteticamente apresenta esse aspecto com mais veemência é a Aruak.

A importância do presente estudo, no qual será realizada sobretudo uma verificação da classificação nominal em verbos se dá sobretudo devido à falta de informação e sistematização dos dados sob uma perspectiva tipológica e comparativa. Sabe-se que muito já se documentou sobre as línguas sul-americanas de um ponto de vista individual de descrição das línguas, mas pouca informação parece haver sobre essa característica categorial de maneira mais precisa, e esta pesquisa de cunho comparativo toma um caráter essencial de estabelecer critérios para reconhecer padrões, distinções e possíveis limitações no que tange à marcação especificamente nos verbos em um grupo geograficamente fechado de línguas.

1.3 Referencial teórico

O *Curso de Linguística Geral*, de Ferdinand de Saussure, publicado no início do século XX é considerado até os dias de hoje como um marco inicial e um divisor de águas na história da Linguística como ciência autônoma da linguagem humana. Contudo, muito se havia feito em precedência a tal evento, e talvez o mais relevante para nosso início de discussão é o fato da descoberta de grandes similaridades gramaticais e lexicais do sânscrito, a língua sagrada dos hindus, com outras línguas europeias de prestígio no século XVIII, como o grego, o latim, o alemão e o persa. A gramática comparativa das línguas teve um grande período de auge durante o século XIX, sobretudo com os trabalhos de Frantz Bopp (LEROY, 1985).

O método histórico-comparativo do século XIX teve notadamente um resultado de cunho genealógico. Buscou-se desvendar relações de parentesco entre os idiomas de modo a organizá-los em grupos filo-genéticos, além de tentar chegar a uma língua indo-europeia supostamente inicial, o Proto-Indoeuropeu, da qual as outras fossem derivadas (PERRET, 2008).

É comum para os trabalhos em Linguística que, ao ser identificada uma certa língua,

se queira também informar a família à qual essa língua pertença. De modo clássico, dizemos que o inglês e o alemão são línguas da mesma família, denominada germânica, descendentes de um grande tronco indo-europeu.

De acordo com Luraghi (2006, p.25), além da classificação genética que, como vimos, considera uma língua ancestral de um grupo em comum, há também a possibilidade de classificar as línguas do mundo por outros dois métodos: o areal e o tipológico. Em uma classificação areal, consideram-se as línguas em situação de contato. A língua Tariana, por exemplo, é a única da família Aruak falada na região do Uaupés, no Alto Rio Negro, onde se encontra em situação de contato com várias línguas da família Tukano, dentre elas o Kubeo, o Juriti e o Tukano. Devido a esta condição espacial, o inevitável contato fez com que traços estruturais de ambas famílias compusessem o Tariana. Nesta perspectiva, o seu sistema de caso, comparado com outras línguas Aruak, aponta uma das evidências do resultado do seu contato com línguas Tukano (AIKHENVALD, 2003).

Considerando uma análise tipológica, as línguas são normalmente tratadas a partir de determinados traços estruturais específicos. Diferentemente da classificação genética, quando se agrupam as línguas do mundo sob o aspecto tipológico não prevalece a preocupação etimológica ou histórica, por exemplo, ao identificá-las nos chamados “tipos”. Uma interpretação tipológica clássica diria que é preciso prosseguir com um exame comparativo, considerando a sua estrutura em um nível de análise linguística holística (fonologia, morfologia ou sintaxe). Já trabalhos atuais (esta pesquisa é um exemplo) não se preocupam necessariamente em tipificar as línguas em grupos. Adota-se então como foco de investigação pelo menos um traço categorial compartilhado dentro de um *corpus*, ao que se convencionou chamar de *cross-linguistics*, i.e., presença da nasalidade nos segmentos vocálicos. Se considerarmos o traço de as línguas apresentarem ou não um pronome expletivo, sem valor semântico, com verbos que não possuem argumentos (e.g. *chover*, *estar calor*, etc.), diríamos que o francês e o português, apesar de pertencerem à mesma família das línguas românicas (classificação genética), estariam distribuídas em tipos diferentes apenas nesse quesito sintático restrito (classificação tipológica), já que em francês o sujeito expletivo é obrigatório, como em *il pleut* (chove) ou *il fait chaud* (faz calor).

Fundamentalmente, é importante resgatar que a Tipologia (pelo próprio termo) busca em sua essência ainda com autores recentes, como é o caso de Moraycsik (2012), abordar a noção de “tipo”. Vejamos abaixo como o autor traz a definição e uso do termo na teoria

linguística.

[...] The term “type” in everyday usage is synonymous with “kind”: it refers to a subclass of a class of entities. In this broad sense, two languages belong to the same type if they have at least one characteristic in common regardless of whether this shared characteristic is due to shared inheritance or borrowing or similar environmental conditions. In actual linguistic usage, however, two languages are generally said to belong to the same type if their similarities hold across various genetic, areal, and cultural groups. (MORAYCSIK, 2012, p. 9 e 10)¹

A primeira tentativa sistemática de se classificar as línguas em tipos da qual temos notícia remonta ao século XIX com os irmãos Friedrich e August von Schlegel. Tomando em consideração a estrutura interna das palavras (morfologia), esses autores estabeleceram três padrões estruturais de línguas que ainda hoje são canônicos na classificação tipológica: *aglutinantes*, *isolantes* e *flexionais* (tais macrotipos denominam-se *categorias holísticas*). Mais adiante uma nova classificação começou a integrar esse grupo: as línguas *polissintéticas* (RUIZ ANTÓN, 1998, p. 9).

É igualmente de ressaltar considerar que já no fim do século XVIII houve também um estudo inicial propriamente linguístico do gênero gramatical, realizada pelo alemão Johann Christoph Adelung (1783), quando coletou uma amostra de 30 línguas e comparou a presença/ausência da distinção de gênero nos artigos e adjetivos (REGÚNAGA, 2012, p. 20).

Foi em meados dos anos 1960 que o linguista norte-americano Joseph Greenberg trouxe uma contribuição muito mais concreta para a investigação tipológica das línguas. Ele colaborou para a configuração de bases de dados adequadas para o trabalho tipológico, elaborando também estratégia de medição numérica do grau como dos tipos de morfologia possíveis nas línguas, e mostrando assim que não necessariamente as línguas se encaixam em tipos muito restritos (REGÚNAGA, 2012, p. 23), como haviam proposto os irmãos von Schlegel. O trabalho de Greenberg esteve muito atrelado ao que chamamos *universais*

¹ [...] O termo “tipo” no uso cotidiano se refere a uma subclasse de uma classe de entidades. Nesse sentido amplo, duas línguas pertencem ao mesmo tipo se elas possuem ao menos uma característica em comum independentemente de essa característica seja devido a questões de parentesco ou empréstimo ou se estão em condições de ambiente similares. No uso linguístico atual, todavia, duas línguas são consideradas do mesmo tipo se suas semelhanças não se restringem a grupos de mesma classificação genética, areal ou cultural. (MORAYCSIK, 2012, p. 9 e 10, tradução nossa).

linguísticos, sob a proposta de que todas as línguas humanas possuem algumas características que são absolutamente universais (i.e, todas possuem vogais e consoantes), outras características são implicacionais (i.e, línguas que apresentam a ordem verbo-objeto têm preposição), além de também tratar de tendências universais e implicacionais, por meio de estatísticas. (RUIZ ANTÓN, 1998, p. 14). Como veremos mais adiante (cf. p. 24), a contribuição de Greenberg para a tipologia se refletirá claramente no presente trabalho, pois também nos será possível realizar interpretações análogas à medida que se procedam as análises do nosso *corpus* de línguas sul-americanas.²

Para além da perspectiva de definir tipos de línguas ou de sistemas, como vimos anteriormente, estamos interessados numa tipologia de construções (Bickell 2010). Nessa perspectiva, tipologia é o estudo da variação translinguística (*cross-linguistics*), que se ocupa de traços gramaticais ou fonológicos e da busca por explicações sobre estes fenômenos. Posto de outro modo, estamos mais interessados em descrever tipos de propriedades recorrentes entre as línguas do que em rotular as línguas com base em certas propriedades.

Ao compreender que há uma intensa relação do modo recente de se fazer pesquisa tipológica com a teoria linguística sistêmico-funcional, o presente estudo tem foco em analisar o funcionamento do Sistema de Classificação Nominal em três línguas geneticamente distintas faladas na América do Sul, levando em consideração também a total ausência de contato entre elas. A noção de classificação nominal será melhor discutida no próximo capítulo, quando levantaremos as discussões relevantes aos termos *gênero*, *classificador* e *classe nominal*. Assim que, partindo de uma perspectiva inicial de Corbett (1991), que considera ‘gênero’ e ‘classe nominal’ como termos intercambiáveis, é possível afirmarmos que a atribuição de um gênero gramatical em uma dada língua possui estreita relação com a ideia de *concordância*, apesar de um requisito semântico sempre estar presente. Outros autores, como é o caso de Regúnaga (2012), entendem que não apenas a concordância é fator para a atribuição de gênero, mas incluem também a ideia clara do sexo biológico em alguma instância.

Atribuiremos, inicialmente, ao gênero gramatical, uma compreensão bem mais ampla e complexa de sua noção tradicional, não se restringindo apenas ao binarismo masculino-

² Em seu estudo, esse autor coletou dados de 30 línguas de diferentes famílias, do qual formulou 45 universais.

feminino, como é de rotina ser subdividido, pois ao se proceder com a análise de línguas indígenas, há de ficar clara a existência de outros paradigmas idiossincráticos e culturais que abarcam na noção semântico-pragmática de gênero como, por exemplo, humano/inumano, pequeno/grande, animado/inanimado, forma física, etc. Será perceptível a abertura conceitual deste fenômeno levando-nos a falar em gênero com menos frequência, dando lugar então para o que chamaremos de Sistema de Classe Nominal.

1.4 Resumo dos capítulos

O presente trabalho foi dividido em três capítulos, sendo este, o primeiro deles, voltado a considerar uma introdução ao tema que apresentaremos a partir de agora, bem como os autores de maior relevo no tema ‘classificação nominal e indexação verbal’, além da metodologia adotada para a sua realização.

O segundo capítulo foi dedicado a uma discussão inicial, e não menos exaustiva, do que vem a ser considerado o gênero gramatical, o Sistema de Classificação Nominal e a concordância em línguas indígenas sul-americanas, baseando em especial no trabalho de Regúnaga (2012). Fez-se relevante também abordar a noção de *indexação* de Haspelmath (2013), como uma tentativa de evitar embates terminológicos, além de estar mais coeso com a análise de línguas com sistemas tão distintos. É ainda neste capítulo que falaremos também de indexação argumental na palavra verbal.

No terceiro e último capítulo apresentaremos brevemente o *corpus* de análise, inicialmente composto de 23 línguas sul-americanas e finalizado com apenas 3, para um estudo analítico prático dos fenômenos vistos no capítulo 2 com foco na argumentação indexada ao verbo. Trataremos de modo específico das seguintes línguas sul-americanas: Mundurukú, Kubeo e Tehuelche. Ainda neste capítulo será feita uma revisitação dos dados expostos nas seções anteriores, de modo a discutirmos comparativamente as análises feitas e chegarmos às conclusões do estudo tipológico.

2 - INDEXAÇÃO E CLASSIFICAÇÃO NOMINAL EM LÍNGUAS INDÍGENAS DA AMÉRICA DO SUL

O objetivo inicial deste capítulo é fundamentar a discussão sobre os Sistemas de Classificação Nominal (SCN), apresentando as noções de gênero, classes nominais e classificadores. Veremos que a classificação nominal, na visão de muitos autores, é, de certa forma, uma ampliação da visão restrita de gênero, que ainda assim não conta com uma compreensão unânime sobre seu conceito na literatura consultada. Partiremos de um tratamento do gênero gramatical em línguas indo-europeias conhecidas de um público não especializado para, em seguida, abranger a análise de fenômenos semanticamente relacionados, como é o caso dos classificadores, proporcionando a inserção dessas duas categorias em uma compreensão global da classificação nominal. Apresentaremos posteriormente um breve panorama geral da classificação nominal em línguas faladas na América do Sul. Veremos mais especificamente as noções da classificação nominal na categoria dos verbos, posto que para este estudo é o real foco de análise nas línguas selecionadas. Confrontaremos o fenômeno da concordância (agreement) dada por Corbett (2006) e a indexação proposta por Haspelmath (2013). Concluiremos com o trabalho de Birchall (2014) para uma apresentação da marcação do argumento verbal em línguas sul-americanas.

1.1 O que é classificação nominal?

2.1.1 Gênero

Em línguas como o português, o árabe e o Apurinã, observamos a presença dos gêneros masculino e feminino; o inglês moderno e o turco não apresentam marca de gênero gramaticalmente explícito (exceto para os pronomes no caso do primeiro); o latim, o grego e o alemão, além do masculino e feminino, possuem um gênero tradicionalmente conhecido como ‘neutro’, ao qual Corbett (1991) chama de *residual*.

A concepção canônica de gênero gramatical – e por canônica entende-se o que a tradição gramatical europeia nos legou durante séculos – se aproxima substancialmente da ideia de gênero sexual ou biológico, cuja transparência semântica de atribuição fica a depender da língua. No caso das línguas dravidianas como o Tamil, falada no extremo sudeste da Índia, é clara a distinção dos nomes classificados como masculinos (seres racionais machos), femininos (seres racionais fêmeas), e neutro (seres irracionais) através

de um estrito critério semântico. Bloomfield (1933, p. 280) apud (Corbett, 1991) atenta para a ocorrência do gênero no alemão, francês e latim e os critérios incertos da designação do gênero gramatical. Observemos a tradução da frase “*A raposa comeu a galinha e o galo*” em alemão, francês e latim:

Fonte: elaborado pelo autor

(1)	Der Fuchs	aß	das	Huhn	und den	Hahn.	[AL]
	DET.M.S	raposa.M.S	comer.PS	DET.AC.NT.S	galinha.AC.S	CJ	DET.AC.M.S galo.AC.S
(2)	Le renard	a mangé	la poule	et le coq.			[FR]
	DET.M.S	raposa.M.S	AUX-comer.PP	DET.F.S	galinha	CJ	DET.M.S galo
(3)	Vulpes	gallin-am	et	gall-um	ed-it.		[LT]
	raposa.NM.F.S	galinha-AC.F.S	CJ	galo-AC.M.S	comer-3S.PS		

Ao selecionar os três nomes - raposa, galinha e galo – constatamos que há, de fato, uma variação arbitrária de gênero, maior para *raposa*, mediana para *galinha*, e nenhuma variação para *galo*. Poderíamos hipotetizar neste caso que raposa seja um animal pouco distinguível biologicamente quanto à identificação dos sexos macho e fêmea, daí uma maior abertura para a arbitrariedade da atribuição do gênero gramatical - quando apenas um pode ser usado - enquanto a percepção visual se torna mais clara na identificação de um galo com relação a uma galinha. Todavia, é interessante notar que, apesar dessa suposta evidencialidade do sexo, em alemão ‘galinha’, por ser introduzida pelo determinante *das*, se apresenta como uma palavra de gênero neutro³: *das Huhn*. A tabela seguinte reúne de forma esquematizada as palavras mencionadas acima com suas respectivas variações de gênero a nível translinguístico.

Tabela 2.1 – Atribuição de gênero em quatro línguas indo-europeias

Palavra	Português	Alemão	Francês	Latim
<i>raposa</i>	Feminino	masculino	masculino	feminino
<i>galinha</i>	Feminino	neutro	feminino	feminino
<i>galo</i>	Masculino	masculino	masculino	masculino

Fonte: elaborado pelo autor

³ Esta palavra não aparece isolada na língua alemã. Um exemplo clássico que também pode servir-nos de ilustração para explorar a não-obrigatoriedade do gênero gramatical com o referente biológico é a palavra *Mädchen*, menina, que também opera como neutro, portanto, *das Mädchen*.

Os três nomes que expusemos no exemplo acima mostram uma certa tendência lógico-semântica na atribuição do gênero devido ao fato de se tratar de seres animados sexuados. No entanto, seria uma tarefa árdua ou mesmo impossível estabelecer uma explicação coerente a fim de sistematizar semanticamente a atribuição do gênero gramatical em palavras como *fogo*, *riqueza* ou *tapete* em português, por exemplo.

Nas línguas indo-europeias encontramos majoritariamente a marcação do gênero em substantivos, adjetivos, pronomes e artigos. Porém, ocorre um fenômeno bastante interessante nas línguas eslavas, que é a marcação morfológica de gênero em determinados verbos no pretérito, como nos mostra Corbett (2013, p.90) com o verbo [byt'], *ser/estar*, em russo.

(1) **Žurnal by-l zdes'.**
jornal estar-PST.[M] aqui.
'O jornal estava aqui.' (e agora foi-se)

(2) **Kniga by-l-a zdes'.**
livro estar-PST-F aqui.
'O livro estava aqui.'

(3) **Pis'mo by-l-o zdes'.**
carta estar-PST-N aqui.
'A carta estava aqui.'

Ao que possa parecer raro que um verbo explicita o gênero concordando com o argumento em masculino (1), feminino (2) ou neutro (3), como vemos em russo, em árabe a marcação de gênero também pode ser encontrada nos paradigmas verbais das segundas e terceiras pessoas, tanto do plural como do singular: *takūnu* (você é [masculino]) *takūnīna* (você é [feminino]).

Tal fenômeno fora já previsto por Greenberg (1964), ao traçar alguns Universais Linguísticos (restritos) que fazem referência ao acordo entre o verbo e o gênero gramatical. Levando inicialmente em consideração que:

Universal 30: Se um verbo possui categorias de pessoa-número ou gênero, ele sempre terá categorias de tempo e modo

Universal 31: Se um verbo concorda com o sujeito ou objeto em gênero, então o adjetivo também sempre concorda com o nome em gênero

Universal 32: Sempre que um verbo concorda com o sujeito ou objeto em gênero, ele também concorda em número.

Greenberg (1963, p.93) ainda nota a seguinte tendência estatística: “concordância de gênero entre um nome (em geral, o sujeito) e um verbo é bem menos frequente do que concordância em pessoa e número”.

Autores como Corbett (1991), Greenberg (1978), Singer (2016) e outros assumem, do ponto de vista estrutural, a noção de gênero gramatical como algo intrínseco à relação de **concordância** dentro de um domínio oracional. Do ponto de vista semântico, entenderemos o gênero como parte de um conjunto de conceitos associados que denominaremos sistema de **classificação nominal**, de acordo com Aikhenvald (2000).

Notamos até aqui a relação inerente entre o gênero gramatical e o fenômeno da concordância, segundo Corbett (1991). No entanto, faz-se necessária uma explanação acerca do que entendemos por ‘concordância’ neste trabalho. Ainda de acordo com esse mesmo autor, pela obra *Agreement*, de 2006, são confrontadas as visões de Bloomfield (1933) e Greenberg (1978), em que nos mostra que existe um debate terminológico que envolve as noções de *agreement* vs. *concord*.

Para Bloomfield (1933, p.191 apud Corbett 2006, p.6), o *agreement* – ao que chamaremos de concordância – se refere a uma extensão maior da relação gramatical que se estabelece entre as partes envolvidas, onde há um *controlador*, i.e. aquele que possui as características que serão transmitidas a outro(s) elemento(s) no *domínio* da concordância, esses elementos são o *target*, i.e. que geralmente recebe as marcas características do controlador em sua forma. Bloomfield defende a tese de que a concordância se divide em três classes: *concord* ou *congruence* (tipo de concordância dentro do sintagma nominal ou nos verbos do predicado); *government* (por exemplo em línguas de casos, em que o papel sintático determina a forma); e *cross-reference* (o argumento aqui é recuperado na estrutura do predicado por meio de um pronome). Apesar das possíveis divergências que possam ocorrer com relação a certos autores, este último conceito será estendido para também o que chamaremos de **anáfora**.

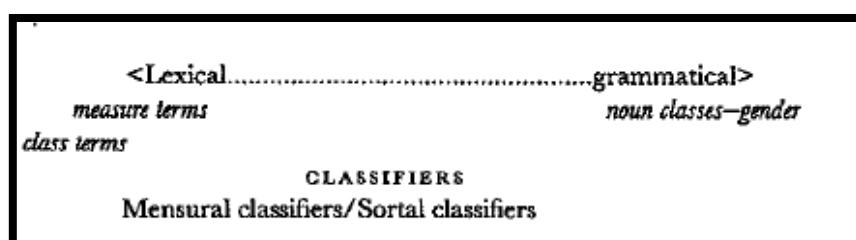
Greenberg (1978, p.50 apud Corbett 2006, p.6), contrariamente a Bloomfield, considera o termo *concord* como sendo o mais abrangente que *agreement*, e a distinção é baseada no tipo de particularidades lexicais envolvidas; e o *agreement* estaria ligado à ideia de uso de afixos, como ocorre na marcação de casos em línguas que possuem essa

característica. De todos modos, assim como Corbett (2006, p. 6-7), opta-se aqui por não estender a discussão terminológica acerca da concordância. Quanto a este ponto, nos aproximamos da tendência bloomfieldiana no que diz respeito a enxergar o *agreement* como um fenômeno generalizado, desde que marcado morfologicamente.

Adotaremos neste estudo a classificação nominal como uma manifestação gramatical de um processo generalizado de categorização semântica que ocorre nas línguas e que está intimamente ligada ao modo como os seus falantes percebem o mundo à sua volta. De uma forma geral, a nossa capacidade cognitiva nos beneficia com essa capacidade criativa de relacionar cada elemento do mundo a outro elemento. Por exemplo, reunimos o morango, a pera e a maçã em uma categoria ampla denominada “fruta”; do mesmo modo associamos uma bola, uma laranja e a lua numa categoria que poderíamos definir como “esférico”. Assim, as línguas do mundo são naturalmente capazes de transpor essas realizações categóricas ao seu sistema gramatical ou lexical. São bastante recorrentes as categorias referentes a humanidade, animacidade, sexo, forma, consistência e propriedades funcionais (Aikhenvald, 2000, p.3).

Grinevald (2000) distribui a classificação nominal primeiramente em Lexical e Gramatical, sendo esta a que realmente nos interessa no presente estudo. Vale observar que a maioria dos autores consideram esses dois extremos dentro de um continuum. Observe-se o esquema abaixo.

Figura 1 - Diferentes sistemas da categorização linguística



Fonte: Grinevald (2000)

A classificação nominal gramatical, como mostrado na figura acima, se ramifica essencialmente em três níveis: *classificadores*, *classes nominais* e *gênero*, que em conjunto formam os mecanismos gramaticais de classificação nominal.

2.1.2 As classes nominais

Alguns autores tratam gênero e classe nominal como o mesmo tipo de sistema de classificação nominal. Aikhevald (2000) tende a tratá-los como *classes nominais*. Corbett (1991) considera os termos “classes nominais” e “gênero” como intercambiáveis: “The use of 'gender' or 'noun class' is also more a matter of tradition than of substance (...) The choice is not important; for consistency, we shall normally use the term 'gender'” (CORBETT, 1991, p. 10). Para este autor, sempre haverá uma motivação semântica na atribuição das categorias nominais, mesmo que não haja manifestação formal.

De acordo com Regúnaga (2012, p.19), é necessário estabelecer uma diferenciação entre os termos gênero e classe nominal, deixando reservado ao gênero uma especificação da classe nominal quando possui um número reduzido de classes (entre 2 e 3) e presença da ideia de sexo natural, ainda que sejam substantivos inanimados.

2.1.3 Os classificadores

O *continuum* da classificação nominal da Figura 1 revela que os classificadores estão mais suscetíveis a uma tendência menos gramatical, enquanto o gênero e classes nominais são analisados como sendo sistemas mais gramaticalizados. Existe um conjunto de critérios que embasam tal análise; dentre eles, o mais importante para nós, é o fato de que na tipologia de Grinevald (2000, 2015) os classificadores são morfemas que apenas ocorrem em determinadas posições morfossintáticas, mas não são marcas de concordância, enquanto gênero e classes nominais são manifestados essencialmente pelo sistema de concordância.

Grinevald (2000), baseando-se em Dixon (1986), lista outros elementos que ajudam a distinguir classificadores de gênero/classes nominais. Vejamos que esse conjunto de categorias procura definir sistemas prototípicos, mas de fato encontramos várias línguas cujos sistemas de classificação nominal são menos prototípicos:

Tabela 2.2 - Gênero vs. sistemas de classificador

Sistemas de classes nominais-gênero	Sistemas de classificador
1. classifica todos os nomes	não classifica todos os nomes
2. tendência a um pequeno número de classes	número maior de classes
3. sistema fechado	sistema aberto
4. fusionado com outras categorias	constituente independente

gramaticais	
5. pode estar marcado no nome	não afixado a um nome
6. realizado com padrões de concordância	marcado uma única vez
7. O nome pertence a uma só classe sem variação do falante	pode pertencer a várias classes
8. sem variação de registro	dependendo do falante usos formais e informais

Fonte: Dixon (1986) apud Grinevald (2000)

A título de ilustração, vejamos abaixo alguns tipos de classificadores com seus respectivos exemplos, retirados de Aikhenvald (1998).

- **Classificadores numerais:** morfemas especiais que aparecem próximo do numeral ou quantificador. Eles categorizam o nome em termos de animacidade, forma, dimensionalidade, acomodamento ou outras propriedades. Os classificadores numerais são relativamente frequentes em línguas isolantes do sudeste da Ásia e nas línguas do norte amazônico da América do Sul⁴.

-
- **Classificadores relativos:** morfema especial geralmente dentro de orações possessivas que categoriza a forma que o nome possuído está para o possuidor. As ilustrações (1) e (2) são do Fijiano, uma língua austronésia (LICHTENBERK 1983, p. 157-58 apud AIKHENVALD 1998, p.431).

(a) na me-qu yaqona (Fijian)

ART CL:bebível-minha kava
'minha kava (que eu pretendo beber)'

(b) na no-qu yaqona (Fijian)

ART CL:GERAL-minha kava
'minha kava (que eu plantei, ou que eu vou vender)'

-
- **Classificadores verbais:** são afixos ou incorporação nos verbos que categorizam um nome, normalmente na função S/P, indicando forma, consistência e animacidade. Podem ocorrer antes de uma raiz verbal ou depois, como no exemplo do Waris, uma língua da Papua Nova Guiné (Brown 1981, p.96).

⁴ A autora não forneceu exemplo.

(c) sa ka-m put-ra-ho-o (Waris)
côco 1SG-para VCL:REDONDO-ter-BENE-IMP
'me dê um côco (lit., 'côco para mim redondo um dê)

2.1.4 Sistemas menos prototípicos

Uma situação intermediária entre classificador e gênero pode ser vista na língua Mawng, falada no norte da Austrália. Segundo Singer (2016), existem 5 gêneros: Masculino, Feminino, Vegetação, Comestível e Geografia. Além de um núcleo de palavras com associação semântica clara a cada gênero, algumas classes de palavras parecem ser “arbitrariamente” associadas a um gênero particular. Concordam com nomes em gênero os pronomes, demonstrativos, artigos, modificadores atributivos e verbos. Em verbos intransitivos, prefixos indexam o argumento S em um dos cinco gêneros, enquanto verbos transitivos indexam o argumento A em apenas duas categorias: masculino e não-masculino, este último abrangendo todos os demais gêneros. O mais interessante, no entanto, é que um nome pode controlar diferentes tipos de concordância com um gênero dependendo do *status* de seu referente, como quando o nome para ferro que se transforma de matéria bruta em telhados de uma casa, mudando o seu gênero de Geografia (como todos os metais e minerais) para Comestível (telhados, casa, objetos domésticos são parte do gênero Comestíveis). Em termos semânticos, essa recategorização do gênero de um nome com base em transformações dos referentes nominais é típica de sistemas de classificadores, como Singer mesmo conclui (2016, p.73):

It is only when we look at how gender is used in context that nominal classification in Mawng starts to strain the boundaries of the usual conceptions of gender (...). The role of gender in these instances seems to be more like that of noun classifiers found in other Australian languages than gender systems we are familiar with.⁵

Vemos em Mawng um exemplo que vai contra noções apriorísticas que assumem que gênero estaria ligado mais à classificação do nome e à sintaxe, enquanto o classificador estaria ligado mais à classificação do referente e ao discurso. Em Mawng, atribuições

⁵ Somente quando analisamos como o gênero é usado no contexto é que a classificação nominal em Mawng começa a forçar os limites das concepções usuais de gênero (...). O papel do gênero nesses casos parece ser mais semelhante ao dos classificadores de substantivos encontrados em outras línguas australianas do que nos sistemas de gênero com os quais estamos familiarizados.

típicas de classificadores, como “instanciação de referência” e desambiguação de sentido de verbos aparecem na categoria gênero/classes nominais.

Isso justifica nossa opção tipológica por não fazer uma distinção entre tipos de sistemas de classificação nominal, considerando como parte do mesmo fenômeno sendo investigado. Para Aikhenvald (2000) e Krasnoukhova (2012), o caso do Mawng se encaixaria na categoria de classificadores múltiplos ou multifuncionais. Para Grinevald e Seifart (2004) é um tipo menos gramaticalizado de classificação nominal, em que há um equilíbrio da natureza mais lexical dos classificadores e mais gramatical dos gêneros/classes nominais.

É irrelevante para nossos propósitos se os índices podem ser analisados como classificadores ou marcas de gênero ou classes nominais ou mesmo pronomes. Como vimos nas seções anteriores, gênero e classes nominais já vêm sendo tratados indistintamente pela literatura com relação a sua função em concordância. Expandindo essa perspectiva, podemos incluir outros sistemas de classificação nominal na nossa tipologia, sobretudo os chamados classificadores verbais. Segundo Aikhenvald (2000), os *verbal classifiers* (classificadores verbais) possuem as seguintes propriedades:

-
- (i) Aparecem no verbo e caracterizam o referente do seu argumento (forma, consistência, tamanho, animacidade, etc.);
 - (ii) Normalmente se referem aos argumentos S e/ou O⁶;
 - (iii) Sua escolha é mais condicionada a uma seleção lexical que um acordo gramatical;
 - (iv) Os nomes podem não estar associados a um classificador, assim como podem se associar a mais de um.
 - (v) Pode ser limitado a um certo número de verbos, de acordo com a semântica destes.
 - (vi) Se dividem em três tipos: incorporação nominal classificatória, classificadores nominais por afixação e classificadores supletivos.
-

As propriedades (i) e (ii) indicam que esses classificadores podem ser analisados como indexadores cuja correferência com argumentos verbais se dá a partir de propriedades

⁶ Majoritariamente neste trabalho apresentamos os macro-papeis sintáticos de argumento como S (sujeito de uma oração intransitiva), A (sujeito de uma oração transitiva) e P (paciente). No entanto, este último componente parece variar livremente entre os autores que o trazem como O (objeto). Birchall (2014) foi a principal referência que adotamos para optar por P, contudo Aikhenvald (2000) é um dos exemplos de autores que optam por O.

semânticas típicas da categorização nominal. As propriedades de (iii) a (v) mostram como esses classificadores se distinguem dos sistemas de gênero, não por sua semântica, mas pelo fato de a relação estabelecida entre indexador e nome em função de argumento se basear numa maior liberdade semântica e na não obrigatoriedade gramatical da indexação.

2.1.3 Panorama dos SCN em línguas indígenas da América do Sul

Dado o exposto acima acerca da teorização referente aos SCN, podemos tão pronto observar como se apresenta tal fenômeno com foco nas línguas indígenas da América do Sul. Para esta seção, serão tomados como base os trabalhos de Krasnoukhova (2012) e Regúnaga (2012).

Como expusemos no início deste capítulo, aos linguistas foi necessário um certo afastamento da noção limitada de gênero atrelada às línguas indo-europeias para que se chegasse enfim a uma compreensão mais holística do SCN. Línguas nativas australianas, as da família Banto, o chinês, o japonês e outras já apontavam indícios da variação desse fenômeno, já que seus sistemas trazem classificadores. Ao passo que pesquisadores passaram a documentar as línguas indígenas faladas no continente americano, foram-se abrindo novos caminhos para a investigação e foi então possível a ampliação do *corpus* (REGÚNAGA, 2012).

Krasnoukhova (2012), assim como outros autores (DIXON 1982, 1986, Grinevald 2000, p.56; Corbett 1991, p.5) opta por não fazer uma distinção substancial entre os termos gênero e classe nominal, pois leva em conta que ambos dividem os nomes em classes rígidas e possuem a propriedade da concordância. Apesar de seu foco não ser o padrão de marcação argumental, a autora analisou dados de 55 línguas indígenas faladas na América do Sul, duas das quais, inclusive, serão foco da nossa análise no próximo capítulo deste estudo (Kubeo e Tehuelche).

Através do trabalho de Krasnoukhova, observamos que fica patente a proposta do continuum da classificação nominal visto no capítulo anterior, após ter acesso a um *corpus* de línguas mais abrangente. Notaremos nesta seção que as línguas podem apresentar um SCN muito bem definido e passível da categorização tipológica dada por Dixon (1982), como é o caso do Tehuelche (Chon), que distingue masculino, feminino e neutro morfologicamente por meio de sufixo no modificador (Krasnoukhova, 2012, p. 197). Outras línguas, como no caso do Mundurukú (Tupí), podemos apontar uma problematização tipológica, já que a língua se encontra em uma posição intermediária entre

um sistema de gênero/classe nominal, classificadores ou até mesmo termos de classe (Gomes, 2006). Portanto, é fundamental que adotemos a noção de prototipicidade, de modo que se uma dada língua se encaixa firmemente nos padrões de uma determinada categoria - seja gênero ou classificador -, diremos que essa o possui como protótipo⁷. Com relação ao gênero prototípico, tomemos como exemplo a língua Chamacoco (Zamucoan), falada no Paraguai, que possui um sistema binário de gênero que distingue masculino/feminino. Ocorre marcação em ambos gêneros por meio de sufixo, e a concordância é explícita tanto no nome como no(s) seu(s) modificador(es). Em (9a), podemos observar uma clara concordância sufixal em {-t}, que sintetiza informação não só de gênero, mas também número e forma completa (FF), após o numeral ‘um’. O exemplo (9b) mostra uma concordância diversificada, pois o modificador numeral apresenta a informação do gênero na raiz da palavra (supleção), operando em acordo com o núcleo do sintagma nominal, marcado com o sufixo {-e}, indicando o feminino em Chamacoco (Krasnoukhova, 2012, p. 198).

(1) Chamacoco (Zamucoan)

(a) <i>kuchi-t</i>	<i>nohme-t</i>	(b) <i>hm-e</i>	<i>otiyer</i>
coisa-M.SG.FF	um-M.SG.FF	mão-F.PL	dois.F
‘ <i>uma coisa</i> ’		‘ <i>duas mãos</i> ’	

Krasnoukhova comparou 17 línguas sul-americanas que possuem marcação prototípica de gênero gramatical. Do ponto de vista semântico, há algumas que, além do clássico masculino/feminino, também reconhecem o gênero neutro ou inanimado; outras línguas distinguem animado/inanimado e, no caso do Jarawara, tem-se o par feminino/não-feminino. Do ponto de vista morfossintático, o gênero pode ser realizado normalmente por meio de afixos que são expressos por um mecanismo de concordância, cujo núcleo é um nome – podendo ou não receber a marcação de gênero – e os modificadores (determinantes, possessivos, adjetivos, etc.). Na tabela 2.3, temos a organização do gênero gramatical na amostra da autora.

⁷ Na definição de Dubois (2007: 388), um protótipo é uma instância típica de uma categoria; os outros elementos são assimilados pela semelhança associada ao protótipo: uma maçã tem uma melhor representação de fruta que a amêndoa.

Tabela 2.3 – O gênero prototípico nas línguas sul-americanas

Línguas	Gênero realizado apenas nos modificadores	Gênero realizado no nome + modificadores
Línguas com duas classes de gênero		
Mosetén, Apurinã, Baure, Jarawara, Mocoví, Pilagá, Tiryó, Hixkaryana, Panare	SIM	-
Chamacoco	-	SIM
Línguas com três gêneros		
Wari', Tehuelche, Movima	SIM	-
Miraña, Kubeo, Desano, Tariana	-	SIM

Fonte: Krasnoukhova (2012, p.197)

Com relação aos classificadores prototípicos segundo Dixon (1982), podemos exemplificá-los com a língua Tsafiki (Barbacoa) em (10). É notável, primeiramente, que o classificador aparece como uma parte do numeral (*peman, palu*), mas tem o seu referente no nome, dando-lhe uma especificação. Nos casos apresentados, observam-se as informações semânticas complementares de gênero, comprimento, rigidez e flexibilidade. Diferentemente do que ocorre no gênero, o classificador prototípico não é hierarquizado por um sistema de concordância, além de se aproximar das palavras de categorias abertas; e a extensa variedade de classificação semântica é um fator categórico a ser levado em conta.

(10) Tsafiki (Barbacoan)

- | | | | |
|--|-----------------------------|---|----------------------|
| (a) <i>peman-ka</i>
três-CLF:gen
‘três mulheres’ | <i>sona=la</i>
mulher=PL | (b) <i>palu-de</i>
dois-CLF:longo.rígido
‘duas bananas’ | <i>ano</i>
banana |
|--|-----------------------------|---|----------------------|

A língua Tsafiki, como vimos, possui classificadores numerais. Porém, há outras modalidades de classificadores presentes nas línguas sul-americanas, como mostra a tabela 2.4. (Krasnoukhova, 2012, p. 204):

Tabela 2.4 – O classificador prototípico nas línguas sul-americanas

Classificadores numerais	Itonama, Tsafiki
Classificadores nominais	Dâw, Koro, Mocoví, Pilagá
Classificadores possessivos	Bororo, Panare, Wichí
Classificadores verbais	Itonama

Regúnaga (2012) também reconhece o SCN do ponto de vista de características prototípicas. Em seu estudo, são analisadas algumas línguas que apresentam não apenas um dos sistemas de classificação, no caso o gênero, como também a combinação deste com classes nominais e/ou classificadores. Segundo a autora, temos os seguintes sistemas nas línguas indígenas sul-americanas:

Sistemas de Gênero:*Mosetén**Jarawara**Guajiro**Wari'**Tehuelche**Guayabero***Sistemas de Gênero e Classes nominais:***Piroa**Tatuyo**Miraña**Baure***Sistemas de Gênero e Classificadores***Matsigenka**Achagua**Mamaindê***Sistemas de Gênero, Classe Nominal e Classificadores***Andokê*

Os critérios utilizados pela autora para analisar qual SCN cada língua possui foram baseados no número de classes de cada sistema (gênero tendo no máximo três classes e os demais sistemas, mas de três), semântico (gênero deve possuir uma oposição semântica que [em parte] aborde a noção de sexo biológico) e gramatical (gênero e classes nominais marcam concordância, classificadores não).

2.2 Indexação

Como discutimos anteriormente ao abordar a noção de SCN sob a ótica de diferentes

autores na literatura, é relevante retomarmos a discussão acerca da concordância e discuti-la com um grau maior de profundidade. Nesta seção fazemos sumariamente um contraponto entre a visão de concordância (*agreement*) defendida por Corbett (2006) e a indexação de Haspelmath (2013). Veremos que as relações entre índice, controlador e domínio mostram que é possível uma abrangência da análise tipológica por parte de Haspelmath, razão pela qual adotaremos futuramente a análise via a noção de indexação desse autor.

2.2.1. A concordância em Corbett

Segundo Corbett (2006), a noção essencial do *agreement* é a covariação entre um constituinte sintático (verbo no nosso caso) em função de outro constituinte (um argumento). Deste modo, é um aspecto sistemático de uma dada língua, entendendo-se aqui uma prototipicidade, quando dentro de um domínio (e.g. sentença) há um elemento controlador, isto é, que determina a concordância, e um ou mais elementos *target*, que irão se ligar a esse controlador e assumir certas características funcionais e semânticas que ele informa, tais como gênero, número, pessoa e outros.

Observamos abaixo um exemplo do italiano, dado por Corbett (2006, p.9).

Italiano (Pierluigi Cuzzolin, personal communication)

(11) **il** **nuov-o** **quadr-o**
 DEF.M.SG. novo-M.SG quadro(M)-SG
 ‘o novo quadro’.

(12) **i** **nuov-i** **quadr-i**
 DEF.M.PL novo-M.PL quadro(M)-PL

(13) **la** **nuov-a** **tel-a**
 DEF.F.SG. novo-F.SG pintura(F)-SG
 ‘o novo quadro’.

(14) **le** **nuov-e** **tel-e**
 DEF.F.PL novo-F.PL pintura(F)-PL

Com um sistema de concordância muito semelhante ao do português, no italiano opera o que Corbett chama de *Canonical Agreement*, pois obedece a alguns critérios morfossintáticos que o autor estabelece e que serão retomados na seção 2.2.3:

- i. controlador e alvo estão no mesmo domínio;
- ii. concordância obrigatória;
- iii. todos os nomes podem controlar concordância;
- iv. todos os alvos refletem o padrão de concordância de um controlador.

2.2.2 A indexação em Haspelmath

Por indexação, adotamos aqui a ideia de Haspelmath (2013), que a apresenta como um conceito muito mais pertinente à análise tipológica de línguas que os termos clássicos como *pronome* ou *marcas de concordância*. E isso se dá sobretudo pelo fato do distanciamento da perspectiva focal nas línguas europeias, i.e. inglês, francês, português. Desse modo, a indexação como fenômeno linguístico mais amplo se concretiza através de elementos funcionais aos quais chamaremos *índice* ou *index* (na sua acepção em inglês). Assim, chamaremos de indexação do argumento a marcação deste elemento sintático em uma posição de correferenciação dentro junto à palavra verbal. De acordo com Haspelmath, há três configurações morfossintáticas do *index*, sendo eles *gramm-index*, *cross-index* e *pro-index*. Vejamo-los nos exemplos abaixo.

(15) Inglês - exemplo de *gramm-index*

John sleep-s

John dormir-3.sg ‘John dorme’

*sleep-s

*John sleep

(16) Espanhol - exemplo de *cross-index*

Los chicos camina-n

Camina-n

*Los chicos camina

(17) Árabe - exemplo de *pro-index*

a. ra’ay-tu l-kalb-a

ver.PRF-1SG.SUBJ DEF-cachorro-ACC

‘Eu vi o cachorro’.

b. ra’ay-tu-hu (*ra’aytuhu lkalba)

ver.PRF-1SG.SUBJ-3SG.M.OBJ

‘Eu o vi.’

Observamos que em inglês há uma expressão conjunta do argumento sujeito *John* em forma nominal completa (*full nominal*) e a marcação no verbo da terceira pessoa do singular em *-s* (*index*), resultando na concordância morfossintática. Esse tipo de indexação, quando a coocorrência do argumento e do índice é obrigatória no parâmetro gramatical da língua, Haspelmath chama de *gramm-index*.

No caso do espanhol, é possível que o argumento sujeito não apareça, porém é necessária a manifestação da marca verbal *-n* que o identifique gramaticalmente no verbo. Havendo essa facultatividade, Haspelmath identifica esse tipo de indexação como *cross-index*.

O último modelo de indexação é o *pro-index*, que encontramos no exemplo do árabe. Em (17a), o objeto *l-kalb-a* é expresso. Em (17b), o mesmo foi substituído pelo índice *-hu*, sendo impossível que esses dois elementos ocorram simultaneamente. Assim, *-hu* se manifesta como um *pro-index*, já que o índice e o argumento nominal não aparecem juntos no mesmo domínio.

2.2.3. A contraposição das propostas de Corbett e Haspelmath

Ao fazer uma associação aos três tipos de indexação apresentadas por Haspelmath, é possível aproximarmos a noção do *canonical agreement* com a do *gramm-index*. Observar-se-á que em ambos encontramos o controlador presente e com expressão explícita de seus traços categoriais; o target tem uma expressão de concordância obrigatória e fechada, i.e., a concordância se baseia em um conjunto limitado de traços gramaticais presentes no controlador. Quanto ao domínio, este apresenta o gênero do adjetivo dependente daquele do nome; é local; é apenas um entre vários outros domínios (Corbett, 2006, p. 9).

Quando o controlador não está presente no mesmo domínio gramatical do target (oração é o domínio para a concordância entre verbo e sujeito, ou o sintagma nominal é o domínio para a concordância entre nome e adjetivo, por exemplo), Corbett (2006) propõe que se trata de um tipo de concordância não-canônica. A vantagem da abordagem de Haspelmath (2013) é que ele tipifica a concordância não-canônica em dois tipos: o *pro-index* e o *cross-index*.

As propriedades semânticas e gramaticais de verbos e adjetivos determinam o número de argumentos que eles devem tomar. Verbos intransitivos como *andar*, por exemplo, tomam apenas um argumento. Se pensarmos que a expressão prototípica dos argumentos se dá pelos nomes, podemos então questionar qual a função dos índices nos diferentes tipos de indexação que propõe Haspelmath (2013). No caso do *gramm-index*, em que o nome e o índice são obrigatórios, vemos que o índice não acrescenta um argumento nem substitui o nome na expressão do argumento. Podemos entender, então, que o índice é uma marca de concordância (canônica, como diria Corbett). Por outro lado, no *pro-index*, onde vemos que ou o índice ou nome em função de argumento podem estar presentes (mas nunca ambos), vemos que o índice possui a função de expressar o argumento verbal, i.e. ele não meramente concorda com o nome, mas o substitui, tendo assim a função de argumento do verbo, semelhante à de um pronome anafórico que substitui um sintagma nominal em um contexto sintático diverso.

A situação menos clara é a do *cross-index*, quando o índice deve estar presente no verbo enquanto o nome em função de argumento é opcional. O fato de os dois poderem estar presentes no mesmo domínio nos levaria a analisar uma situação de concordância; por outro lado, vemos que o índice pode satisfazer a função de argumento, substituindo o nome como no caso do *pro-index*. No entanto, o fato de o argumento poder ser representado tanto pelo nome quanto pelo índice no mesmo domínio (ou como diria Corbett [2006, p.106] o fato de que o argumento está multirepresentado na mesma oração) nos leva a interpretar ambos como elementos correferentes, possuindo a mesma função sintática na oração. A função do índice seria, portanto, a de correferenciar, i.e., marcar por meio de uma espécie de cópia, na palavra verbal, um argumento nominal.

Com a finalidade de esquematizar a nossa associação teórica dos dois autores que expusemos acima, propomos abaixo uma tabela resumitiva, seguida de exemplos concretos de alguns outros dados, dentre os quais, de línguas que mais à frente servirão para uma análise mais detalhada.

Tabela 2.4 - Relação entre índice e concordância

Haspelmath (2013)	Corbett (2006)	Distribuição do Índice	Função do Índice
--------------------------	-----------------------	-------------------------------	-------------------------

gramm-index	Concordância Canônica	índice e controlador/argumento nominal devem estar presentes no mesmo domínio	Expressão de Concordância com o argumento
cross-index	Concordância Não-canônica	índice deve estar presente no target e o controlador/argumento nominal pode ou não estar presente no mesmo domínio	Expressão de Correferência com o argumento
pro-index	Concordância Não-canônica	ou o índice ou controlador/argumento nominal devem estar presentes no mesmo domínio	Expressão do Argumento

2.3 Classificação nominal e indexação verbal nas línguas da América do Sul

Viu-se até o momento que a indexação é um recurso morfossintático muito comum nas línguas para expressar um argumento verbal, trazendo ao predicado (verbo) traços categoriais como *pessoa*, *número* e *gênero*. Nesta seção trataremos especialmente da indexação verbal nas línguas da América do Sul com atenção especial na informação do gênero gramatical, e para tal será essencial o trabalho de Birchall (2014), que se engaja na perspectiva de Haspelmath (2013) de evitar os termos *concordância* e *referência cruzada*, optando por *indexação*, de modo a considerar a marcação argumental na morfologia do verbo como um fenômeno de difícil compatibilidade na definição terminológica até então, sendo, portanto, algo *sui generis* (Birchall, 2014, p.42).

Observemos no exemplo seguinte como é realizada uma marcação do argumento verbal da língua Yurakar'e (van Gijn 2006, 186 apud Birchall, 2014:42):

Yurakar'e (isolada)

- (16) **ma-bëbë-y** **petche=w**
3pl-search-1sg fish=pl
'Eu procurei peixes.'

O verbo {-bëbë-} é preenchido com prefixo e sufixo, marcando respectivamente o objeto {ma-} 'peixes' e o sujeito transitivo na primeira pessoa {-y}. Nota-se, portanto, que em Yurakar'e há duas posições morfológicas, i.e., duas realizações possíveis de indexação verbal. Esses marcadores, ao mesmo tempo que expressam informação de pessoa, também expressam número, por fusão de traços em um único marcador (index).

Seguindo a proposta de Birchall (2014), para compreendermos a natureza do marcador de argumento verbal (MAV)⁵, é importante considerar seis fatores:

- a. Identificar qual o argumento está sendo marcado (S, A, P).
- b. Localizar o argumento na palavra verbal (prefixo ou sufixo).
- c. Identificar a característica do alinhamento sintático do marcador quanto à transitividade verbal.
- d. Verificar se há fusão de argumentos no marcador.
- e. Verificar se há fusão de característica no marcador (pessoa, número e gênero)
- f. Analisar se há obrigatoriedade do marcador.

Como resultado de uma investigação já realizada em Birchall (2014), apresentamos a tabela 2.5, que reúne as informações dos dados coletados por ele em um *corpus* de 77 línguas indígenas faladas na América do Sul apresentando indexação.

Tabela 2.5 – Configurações possíveis do MAV⁸

Pessoa	Número	Gênero
S, A, P (prefixo, sufixo, variável, não-marcado)	S, A, P (prefixo independente, sufixo independente, fusão com pessoa, não marcado)	S, A, P (prefixo independente, sufixo independente, fusão com pessoa, não marcado)

Fonte: Birchall (2014)

Em quase todas as línguas que possuem nomes com gênero gramatical analisadas por Birchall (2014), essa categoria vai se refletir em algum argumento do verbo, tendo como base de atribuição a semântica do SCN. Notou-se ainda que as informações de gênero são fusionadas com pessoa e número. Inclusive, 77% dos índices verbais fundem o gênero com as noções semânticas de número e pessoa, porém não se tem registro da vinculação do gênero com número excluindo a pessoa (Birchall, 2014, p.65).

Quanto às possíveis localizações dos marcadores, vimos na seção 2.2.3 que Lazard (1998) identifica nas línguas indo-europeias uma maior tendência dos verbos de indexarem seus argumentos por meio de sufixos e que o índice poderia se manifestar como prefixos

⁸ Os termos apresentados na tabela 2.5 foram retirados do autor-fonte.

em línguas indígenas sul-americanas. Na amostra de Birchall (2014) a posição prefixal da marcação de argumentos, em todos os macro-papéis temáticos (S, A e P) tem se mostrado predominante. Ocorre a indexação prefixada em muitas línguas dos grupos Tupí e Caribe. Já os sufixos são frequentes em línguas andinas e do oeste amazônico; e os circunfixos foram encontrados nas línguas da família Guaycuriana, e.g. Mocoví (p.55).

Nota-se que os argumentos expressos por meio de circunfixos são diferentes dos que Birchall chama de *mixed set*. Enquanto os sistemas com circunfixos se realizam por meio de afixação descontínua, o *mixed set* é um mecanismo de indexação não coordenado, ou seja, ora são prefixos ora são sufixos. É o caso da língua Ika (Chibchan) (Birchall, 2014, p.57).

Quanto aos possíveis alinhamentos morfossintáticos, o *corpus* de Birchall mostrou que há línguas que indexam no mesmo paradigma tanto os argumentos S e A, reservando o argumento P para um *set* diferente, apresentando um padrão nominativo- acusativo. Essa configuração é a mais abundante do *corpus* e corresponde a 30 das 66 línguas como a pessoa indexada no núcleo do argumento (Birchall, 2014, p.58).

O alinhamento ergativo-absolutivo ocorre quando o argumento S e P se encontram no mesmo paradigma, sendo que A é marcado em um *set* à parte. Ocorre em algumas línguas Tupí faladas na Amazônia meridional, acima do Rio Madeira, como no caso do Mekens (p.59), do Katukina-Kanamari e do Xavante (Macro-Jê).

Além dos dois alinhamentos morfossintáticos apresentados acima, Birchall também encontrou outros dois, de marcações que poderíamos considerar mais extremas: o alinhamento tripartite, que corresponde a uma marcação em *sets* individuais para cada um dos três papéis argumentais, e.g. Wichí (família Matacoano), falado no Chaco argentino (p. 61); e o alinhamento neutral, em que todos os três argumentos compartilham o mesmo índice.

Revisado o arcabouço teórico que fundamentou a discussão de nossos principais pontos conceituais, passaremos no seguinte capítulo a uma verificação de como o fenômeno dos SCN e a indexação podem ser vistos em algumas línguas específicas para, em seguida, procedermos com uma análise comparativa do nosso *corpus* selecionado.

Neste capítulo foi possível fornecer um conciso aparato teórico referente à discussão fundamental dos sistemas de classificação nominal (SCN) e processos de indexação de

argumentos verbais. Vimos que “gênero” é um conceito que transcende muitas vezes, e mesmo não de forma unânime na literatura, a ligação intrínseca da ideia de sexo biológico ou ao que a este se remeta em alguma instância semântica e cultural. Foi possível posteriormente observar a configuração estrutural dos SCN em conjunto com outros elementos da sentença e notou-se com clareza que o fenômeno da concordância entre os autores tende a direcionar uma bipartição do fenômeno, entre o que é classe nominal/gênero e classificador. No entanto, conforme foi argumentado nas seções 2.2.2 e 2.3 com base em novas descrições de línguas australianas e sul-americanas, nem sempre é possível identificar sistemas prototípicos, i.e., que seguem à risca a tabela proposta por Dixon (1982) e apresentado por Grinevald (2000). Posto que no presente estudo nos focamos em discutir e analisar a argumentação verbal expressa no núcleo do predicado⁹, decidiu-se por uma aproximação mais acentuada à teoria de indexação proposta por Haspelmath (2013).

⁹ Na definição de Tallerman (2015, p. 40), um PREDICADO expressa um ‘evento’ na sentença, que pode ser de fato um evento (como um *colapsar* ou *explodir*), mas pode ainda se referir a ações, processos, situações, estados e outros.

3- ANÁLISE DE INDEXAÇÃO E CLASSIFICAÇÃO NOMINAL EM UM CORPUS SELECIONADO

No capítulo anterior foi possível fornecer um conciso aparato teórico referente à discussão fundamental dos sistemas de classificação nominal (SCN) e processos de indexação de argumentos verbais. Vimos que “gênero” é um conceito que transcende muitas vezes, e mesmo não de forma unânime na literatura, à ligação intrínseca da ideia de sexo biológico ou ao que a este se remeta em alguma instância semântica e cultural. Foi possível posteriormente observar a configuração estrutural dos SCN em conjunto com outros elementos da sentença e notou-se com clareza que o fenômeno da concordância entre os autores tende a direcionar uma bipartição do fenômeno, entre o que é classe nominal/gênero e classificador. No entanto, conforme foi argumentado nas seções 2.2.2 e 2.3 com base em novas descrições de línguas australianas e sul-americanas, nem sempre é possível identificar sistemas prototípicos, i.e., que seguem à risca a tabela proposta por Dixon (1982) e apresentado por Grinevald (2000). Posto que no presente estudo nos focamos em discutir e analisar a argumentação verbal expressa no núcleo do predicado, decidiu-se por uma aproximação mais acentuada à teoria de indexação proposta por Haspelmath (2013).

Neste terceiro capítulo, serão apresentadas primeiramente as etapas adotadas até a escolha do *corpus* final de análise das nossas três línguas faladas na América do Sul (Mundurukú, Kubeo e Tehuelche). Posteriormente, cada uma dessas línguas será analisada individualmente na sua tipologia de SCN, considerando os contextos morfossintáticos, semânticos e funcionais, antecedidas sempre de informações introdutórias gerais, tais como localização geográfica, número de falantes e os autores de base.

3.1 Procedimento para a obtenção do *corpus*

O primeiro passo para a seleção do nosso *corpus* de análise mais minuciosa de apenas três línguas foi a elaboração de uma listagem representativa de línguas indígenas faladas na América do Sul em que constasse alguma manifestação dos SCN, i.e., seja no nome, no adjetivo, no verbo, em forma de prefixo, sufixo, clítico, em fusão com outras categorias como pessoa e número, etc.

A partir de um levantamento bibliográfico em gramáticas, teses e artigos, coletamos um total de 23 línguas, das quais identificamos as famílias e buscamos identificar as classes gramaticais que apresentavam a marcação de gênero, além da semântica nele presente e.g. masculino, feminino, animado, inanimado, etc. Também julgamos pertinente indicar quando ocorresse a marcação na forma de classificador. Para tal, foram elaboradas algumas tabelas (em apêndice) dividindo essas 23 línguas em seções correspondentes às seguintes classes gramaticais:

- Demonstrativo
- Numeral
- Nome
- Pronome
- Modificador
- Possessivo
- Verbo

Em um segundo momento, elaboramos uma tabela mais específica dos indexadores verbais, que é o foco real neste trabalho, partindo das mesmas 23 línguas. O objetivo, portanto, é que aprofundemos em explorar detalhes neste segundo momento.

3.1.1 O SCN fora da palavra verbal

A lista, organizada em ordem alfabética, contendo 23 línguas sul-americanas com presença de algum SNC - vide tabelas no apêndice - foi montada com base em diversos autores, sendo Alexandra Aikhenvald (2000) a mais notória. A intenção nesta primeira etapa é fazer considerações mais generalizadas do que podemos absolver através dos dados, observando as distintas classes gramaticais.

Das 23 línguas, representando 12 famílias e 3 isoladas, o grupo das Aruak é o que conta com o maior número (6), seguido das famílias Tukano, Arawá e Macro-Jê (todas com 2 representantes). Focaremos nossa discussão nessa seção principalmente nas línguas Aruak, mas apresentaremos alguns comentários sobre outras línguas do nosso *corpus*.

Demonstrativo: apenas duas línguas do *corpus* (Mosetén e Yurakaré) não apresentaram SCN. Nas línguas da família Aruak percebe-se um sistema de atribuição que coincide com as suas posições areais: é predominante na região do oeste amazônico que as variedades

dessa família façam distinção entre feminino e não-feminino, como no caso do Apurinã, Baniwa, Tariana e Baure. Já em Lokono e Palikur, que fazem parte do grupo Aruak localizado no nordeste da América do Sul, vemos que em Lokono uma atribuição curiosamente contrária à tendência do outro grupo, i.e., o masculino e o não- masculino se contrastam, além de constar ainda o gênero não-humano; em Palikur há uma distribuição semântica entre masculino vs. feminino vs. neutro. Em Kubeo (Tukano) e Sikuani (Guahiban) nota-se um sistema duplo de gênero e classificadores, e em Siona (Tukano) faz-se a atribuição por anáfora.

Numeral: os classificadores numerais são presentes na maior parte das línguas Aruak, porém chama a atenção o caso do Lokono que não marca qualquer SCN. Apesar de que em Rikbatsa (Macro-Jê) e Chipaya (Uru-Chipaya) não foi possível identificar se há ou não SCN, em grande parte do *corpus* é presente.

Nome: é a classe gramatical controladora do SCN. A única língua Aruak que não consta de marcação é o Palikur, e as outras 5 seguem o mesmo modelo de atribuição dado para os demonstrativos, com exceção do Baure, que possui aqui um sistema de classificadores nominais.

Pronome: no grupo Aruak, o Tariana é a única língua com sistema de classificadores. O Miraña (Bora) se destaca por marcar feminino vs. masculino apenas no dual, para todas as pessoas, e na 3ª pessoa singular.

Modificador descritivo: é uma função sintática que se associa normalmente com os adjetivos e verbos nominalizados. O Baure é um exemplo de língua Aruak sem essa marcação, e o Baniwa possui dois sistemas de classificadores. Lokono, Apurinã, Palikur e Paumari atribuem condições de marcação, além de outros critérios, à semântica dos verbos. Percebe-se muito frequente em verbos estativos.

Possessivo: em Yurakaré (isolada) é a única classe que recebe marcação, e em situações referentes a animais de estimação. Curiosamente, todas as línguas mostraram ocorrência de atribuição de SCN, sendo que em Miraña a marcação é opcional no objeto possuído.

3.1.2 O SCN nos índices verbais

Estivemos restritos a no máximo dois padrões distintos de organização do paradigma de indexadores em cada língua, limitando nossa observação a predicados verbais finitos,

na voz ativa, no modo declarativo e com valores semânticos e formais não marcados no que tange modalidade deôntica e epistêmica. A partir da seleção do *corpus*, fizemos uma nova tabela, com enfoque mais detalhado do SCN no verbo. Informamos primeiramente a quantidade de sistemas que a língua apresenta, seguido da semântica envolvida na atribuição, os argumentos indexados, pessoa, número e se há fusão dessas categorias com o SCN. O resultado pode ser conferido na tabela 3.1.

Os índices dos argumentos foram classificados conforme os seguintes parâmetros:

- **Classificação Semântica:** masculino, feminino, neutro, animado, inanimado, humano, não-humano, propriedades físicas e funcionalidade.
- **Pessoa:** 1a, 2a, e 3a pessoa
- **Número:** singular, dual, plural ou não-singular
- **Macro-papel semântico do argumento:** A (agente de verbo transitivo), O (paciente de verbo transitivo), S (argumento de verbo intransitivo), Sa (argumento de verbo intransitivo ativo) e Sp (argumento de verbo intransitivo estativo).
- **Forma:** proclíticos, prefixos, sufixos, enclíticos, formas livres e incorporação nominal

Tabela 3.1 – A indexação verbal em um *corpus* selecionado

Língua	Família	n° de SCN	Semântica	Argum.	Pessoa	Número	Fusão	Marcação	Fonte
Apurinã	Aruak	2	Não-fem. vs. fem / Propriedades físicas	A/Sa/P/Sp P	3	SG	Pessoa	Sufixo	Fagundes 2000
Baniwa	Aruak	1	Não-fem vs. fem (AN).	A/Sa Sp/P	3	SG	Pessoa/ número	Próclise Ênclise	Aikhenvald 2007
Baure	Aruak	2	Não-fem. vs fem. / Propriedades físicas	A/Sa/P/Sp P	3	SG	Pessoa	Próclise Ênclise	Aikhenvald 2007

Chipaya	Uru-Chipaya	1	Masc. vs. fem. (focalização opcional)	S/A	3	-	-	Clítico	Danielson 2007 Krasnokhova 2012
Jarawara	Arawá	2	Anim vs. inan. Masc. vs fem.	S/A/P	1/2/3	Não SG	- Modo verbal	Prefixo Sufixo	Dixon 2004
Kubeo	Tukáno	1	Masc (AN), fem (AN) Inanimado,	S/A	1/3	SG/PL	Pessoa e TAM	Sufixo	Chacon 2012
Kwazá	Isolada	1	Propriedades físicas	-	3	-	-	-	Van der Voort 2004
Lokono	Aruak	1	Masc. vs. fem.	A/Sa P/Sp	3	-	Pessoa	Prefixo Sufixo	Pet 2011
Miraña	Bora	1	Propriedades físicas	S/A	1/2/3	-	-	-	Seifart 2005
Mosetén	Mosetén - Chimané	1	Masc. vs. fem.	S A/P	1 3	SG/PL	não	Sufixo	Sakel 2004 Kranoukhova 2012
Movima	Isolada	2	Masc vs. fem vs. neu / Propriedades físicas	S/A/P	1/3	SG	Pessoa	Sufixo	Haude 2006 Grinevald 2000
Mundurukú	Tupí	1	Propriedades físicas	S/P Sa e Sp	3	-	-	R-CL-Verbo	Gomes 2006
Palikur	Aruak	3	Masc vs. fem vs. neu; Fem. vs. não-fem; Propriedades físicas	Sa/A Sp/P	3	-	-	Próclise	Aikhenvald e Green 1998
Paumari	Arawá	2	Masc vs. fem / Propriedades físicas	S/A/P S/P	3	SG	-	- Prefixo ka-	Aikhenvald 2010 Salzer e Chapman 1998
Rikbatsa	Macro-Jê	1	Propriedades físicas	-	3	-	-	-	Silva 2011
Saliba	Piaroa-Saliban	2	Não-fem vs. fem/ Propriedades	Sp/P Sa/A	3	-	-	Ênclise Prefixo -	Estrada Ramirez 1996 Labrada 2015

			físicas						
Siona	Tukáno	1	Não-fem vs. fem	S/A	1/3	SG/PL	P/N e TAM	Sufixo	Wheeler 1987
Tariana	Aruák	2	Não-fem vs. fem/ Propriedades físicas	S/Sa/A/P	3	-	Pessoa	Prefixo e Sufixos	Aikhenvald 2003
Tehuelche	Chon	1	Masc/fem vs. neu	P/Sp e nominalizações	3	-	Pessoa	Sufixo	Garay 1998 Krasnoukhova 2012
Wari	Chapakuran	1	Neu vs. masc/fem. no passado real/presente)	A/S/P	3	-	Pessoa	Sufixo	Everett & Kern 1997

Das 23 línguas apresentadas na seção anterior, em 20 delas encontramos SCN no predicado. Não entraram na lista acima o Kariri (Rodrigues 1997), o Sikuni (Queixalós 1998) e o Yurakare' (van Gijn 2006), cujos sistemas não se referem às relações gramaticais entre argumentos e os núcleos de uma oração.

Constatamos que a maior parte do *corpus* é representado por línguas com apenas 1 (um) SCN marcado no predicado, e a maioria delas mostra um sistema de gênero que contrasta a semântica masculino-feminino. As línguas de família Aruak seguem um padrão muito coerente com as informações que obtivemos de marcação nas outras classes gramaticais com o par não-feminino vs. feminino, mas o Palikur é um caso raro que se destaca por ter o maior número de SCN: três na indexação verbal que também se encontram no SN, mais dois que apenas se encontram no SN (Aikhenvald e Green 1998). Quanto às classes nominais, elas ocorrem com muita frequência em praticamente todas as famílias, variando consideravelmente na quantidade de classes representadas, sendo o Paumari de menor número (2 - Aikhenvald 2010), e o de maior número atribuído ao Kwazá, com 150 (Van der Voort 2004), sendo ambos sistemas relativos à semântica de 'propriedade física'. Sobre a categoria de classificadores verbais, Palikur e Rikbatsa são as únicas línguas que os apresentaram.

Ao fazermos uma comparação das nossas listas anteriormente analisadas no tocante às diferentes atribuições dos SCN, é possível estabelecer aqui duas novas generalizações tipológicas, lembrando dos universais de Greenberg:

- Línguas que têm um SCN no predicado, sempre terão também no SN.
- Todos os SCN encontrados no predicado serão também encontrados no SN

(excetuando-se os classificadores verbais).

A propósito dos alinhamentos morfossintáticos, nota-se que o grupo tipológico que mais se destaca em quantidade são as línguas de sistemas ativo-estativo (7 línguas), abrangendo a totalidade das línguas Aruak, o Baure e o Sáliba.

Tomando como referência apenas o sistema de gênero, podemos organizar o nosso *corpus* com os alinhamentos na tabela 3.2.

Entre os 18 sistemas de gênero constatados, 16 indexam gênero do argumento A (além de outros), enquanto 14 indexam do argumento P (além de outros). Os classificadores verbais apenas classificam os argumentos S/P, e os CNs tendem muito a expressar os argumentos S/P, já que em 8 dos 9 casos analisados, somente o Miraña expressa os argumentos A/S.

Tabela 3.2 – Alinhamento em sistemas de gênero

Alinhamentos		Línguas
Nominativo	A/S	4 - Chipaya, Kubeo, Palikur, Siona
Nominativo-Acusativo	A/S vs. P	1 - Mosestén
Acusativo	P	1 - Wari'
Direto	A/S/P	4 - Jarawara, Movima, Paumari, Wari
Ativo-Estativo	A/Sa vs. P/Sp	7 - Apurinã, Baniwa, Baure, Lokono, Palikur, Saliba, Tariana
Estativo	P/Sp	1 - Tehuelche

Palikur e Jarawara são línguas que possuem dois sistemas de gênero no predicado e usam o mesmo sistema de alinhamento para ambos. Já as línguas com um sistema de gênero e outro de CN ou CLV possuem um padrão de distribuição complementar.

Ainda que não tenhamos exposto por completo as informações sobre pessoa e número na tabela 3.1, é possível generalizar que a marcação da classificação nominal de um argumento na terceira pessoa é presente na totalidade de línguas do corpus, seguido da primeira, e só assim a segunda. Podemos, portanto, defender que:

- Sempre que houver indexação da 2ª pessoa, haverá também da 1ª. E sempre que houver indexação da 1ª pessoa, haverá também da 3ª.

Nota-se que a maior parte dos índices se fundem com a categoria de pessoa, quase sempre por meio de sufixação, o que se verificou em 9 línguas. Muito recorrente também são os prefixos e clíticos.

Retomando as duas línguas do nosso corpus que apresentaram CLV (Rikbatsa e Palikur), e combinando as suas informações vistas anteriormente através de Silva (2011) e Aikhenvald e Green (1998) com as três primeiras proposições dadas por Aikhenvald (2000) (seção 2.1.2) – as que mais nos interessam neste estudo - podemos ratificar que em ambas o CLV aparece no verbo expressando categorização do referente. Em Rikbatsa faz alusão à ideia de ‘forma’, enquanto que em Palikur se remete à ‘propriedade física’. Quanto aos argumentos envolvidos, em ambas há apenas classificação dos argumentos S/P. Se baseiam ainda, como vimos acima, em propriedades físicas do referente, o que sugere uma maior tendência de classificação lexical em detrimento de um fenômeno mais gramaticalizado e, por consequência, requerer concordância.

3.2 Análise de um corpus trilíngue (Mundurukú, Kubeo e Tehuelche)

Nesta seção faremos uma análise individual dos SCN em cada uma das três línguas sul-americanas do nosso corpus específico (Mundurukú, Kubeo e Tehuelche), utilizando as informações iniciais tratadas na tabela comparativa de 23 línguas, porém tratando de identificar aspectos mais pontuais dos sistemas encontrados e suas características tipológicas, com base nos exemplos retirados de suas respectivas fontes.

A análise consistiu primeiramente em fazer uma introdução geral das línguas, escolhidas com base nos pressupostos mencionados no capítulo inicial deste trabalho, logo uma exploração dos SCN dentro dos sintagmas nominais, e então seguimos seis etapas de verificação similares às utilizadas no início deste capítulo para a identificação precisa dos indexadores verbais para estruturar novas tabelas. Os parâmetros de verificação foram os que citamos acima, na seção 3.1.2.

3.2.1 Mundurukú

As informações desta seção sobre o Mundurukú estão fundamentadas nos trabalhos de Dionei Gomes (2006; 2009). Trata-se de uma língua do tronco Tupí, que faz parte especificamente de uma família como o mesmo nome, sendo atualmente a única ainda

falada¹⁰. Conta com cerca de 10.000 falantes em mais de 80 aldeias localizadas majoritariamente no Vale de Tapajós e seus afluentes, no estado do Pará. São encontrados Mundurukú ainda no Amazonas, especificamente na Bacia do Rio Madeira e no município de Juara-MT (Gomes 2006). Do ponto de vista social, existe um certo esforço por parte da comunidade Mundurukú no que tange à preservação da sua língua materna, ainda que o bilinguismo e o câmbio linguístico para o português seja uma realidade cada vez mais presente, dado que é um povo com forte ativismo político, que luta pela demarcação de suas terras e pela sobrevivência da tradição cultural.

Do ponto de vista tipológico sobre a indexação, podemos iniciar a discussão sobre o Mundurukú com o que já apuramos anteriormente. Vimos que nessa língua não existe um sistema de gênero (restrito) como ocorria na maior parte do nosso *corpus* (cf. 2.1.1), cujo controlador é o núcleo do SN e os alvos (*targets*) variam segundo o mecanismo da língua. O sistema do Mundurukú consta de várias categorias com semântica de ‘propriedade física’, expressas por meio de *Nomes em Função Classificadora (NFC)* na definição de Gomes (2006). De acordo com o autor, os NFCs não são parte de uma classe específica de nomes classificadores, e, sim, são nomes inalienáveis que operam com a função classificadora, como no caso de partes do corpo humano, partes do corpo animal, partes do corpo vegetal e partes de alguns objetos (como a corda de um arco).¹¹

Esses NFCs possuem um sentido primitivo altamente lexical e independente, mas também operam na gramática do Mundurukú como um determinador de classe. Por exemplo, a raiz *bu* no sentido primitivo significa ‘dedo’, como em *Biboy bu* ‘o dedo de Biboy’, mas como NFC assume uma característica semântica do referente que ele classifica, que é a propriedade de ser ‘cilíndrico e flexível’ (como um dedo). Em Mundurukú *ixi bu* significa ‘cipó’, trata-se de uma composição de um elemento núcleo do significado ao qual se adiciona um classificador que se remete à forma física do objeto.

No quadro 3.3, apresentado por Gomes (2009), podemos esquematizar alguns exemplos dos nomes *bu* ‘dedo’ e *dup/tup* ‘folha’ que podem ou não exercer a função de classificador (A = NFC, B = nome dependente); cada uma divisão desses nomes em duas

¹⁰ Informação recente em comunicação pessoal com Dionei Gomes, pois nas publicações ainda consta como viva a língua Karuaya.

¹¹ “Há um conjunto de nomes que revelam uma função classificadora em Mundurukú. Todos eles são nomes de partes e, portanto, de posse inalienável. Como classificadores, dizem respeito à forma do objeto designado.” (GOMES, 2006, p. 178)

classes, que se organizam com base na necessidade do emprego dos relacionais. De fato, os NFCs são sistematicamente precedidos de um prefixo relacional, sendo ele R1 (para o referente que expressa contiguidade com relação ao seu núcleo) ou R2 (para o referente que expressa não-contiguidade). É importante salientar que esses prefixos relacionais, que atuam dentro ou/e fora da palavra verbal, podem ser considerados bases de relações sintáticas, porém não expressam em absoluto qualquer ideia ou função de categorização nominal. Recordemos que em Mundurukú os nomes que desempenham a função classificadora são sempre inalienáveis, e nos exemplos que seguem, vemos que não existe diferença morfológica que os diferencie quanto à função que exercem. Abaixo propomos uma tabela que ilustra o contraste entre esses dois usos que apresentamos:

Tabela 3.3 – Comparação entre o uso classificatório e não-classificatório de um nome

	Class e I			Classe II			
A	(a) puy (b) i-bu	@-b	‘cobra’ ‘cobra’	(a) warepup u (b) tup	du p	‘borboleta’ ‘borboleta’	
B	(a) Biboy (b) i-bu	@-bu	‘dedo do Biboy’ ‘dedo dele’	(a) ako dup (b) tup		‘folha bananeira , ‘folha dela’	d e

Fonte: Gomes (2009: 10)

Vemos, então, que os NFCs se diferenciam dos nomes por uma questão semântica. Uma outra dimensão que os diferencia é sintática e diz respeito ao tipo de *incorporação nominal* a que esses morfemas estão sujeitos. A incorporação nominal em Mundurukú possui as seguintes funções, segundo Gomes (2006, p.448).

- i. identificar o argumento absolutivo e evitar ambiguidade;
- ii. tornar mais específico o significado do verbo;
- iii. promover um rearranjo sintático e mudar o foco da informação.

Vamos analisar abaixo algumas estruturas da oração em Mundurukú e os NFCs.

- (13) (a) "e)n tu puy @-bu e=su-bu-aoka?"
você INT cobra R1-NFC 2=R2-NFC-matar.PRF
- (b) "õn i-bu @-aoka.ka-m"
eu R2-NFC R1-matar.DUR-IPRF
- "Você matou a cobra?" "Eu a estou matando."

Os dados acima mostram como é estreita a relação entre os relacionais R1 e R2, representados respectivamente pelos prefixos @-, *su-* e *i-*, com o NFC *bu*, cujo referencial está associado à sua forma cilíndrica e flexível (no caso a cobra). Observamos em (13a) que o NFC *-bu-* se associa ao verbo *-oaka* ‘matar’, resultando em uma incorporação nominal, ou seja, na própria palavra verbal ocorre um nome, sintaticamente exercendo a função P. Esse NFC ainda em (13a) pode ser constatado anteriormente ao sintagma verbal, junto ao radical *puy* (cobra). Como o NFC faz referência ao SN que está fora da palavra verbal, usa-se o R2 *-su-*, o relacional de não contiguidade. Já em (13b), o SN de objeto possui o R2 *i-*, que está sendo usado como um recurso anafórico, informando que o NFC é referente em outro domínio sintático, no caso uma oração precedente. O R1 @- que aparece no verbo em (13b) reflete que o SN em função de objeto precede imediatamente o verbo.

Os relacionais R1 e R2 condicionam o tipo de indexação em referência a um elemento de 3ª pessoa singular, uma vez que servem de base para tal. Assim, R1 sucede diretamente o elemento referente – *puy* - e se liga ao NFC, fazendo uma espécie de intermediação entre eles. O R1 marca uma contiguidade sintática entre os dois constituintes. Podemos interpretar o R1 como uma espécie de *gramm-index*, pois o mesmo marca que dois constituintes gramaticalmente relacionados se encontram dentro de um mesmo domínio sintático, i.e., o SN em (a). Já o R2 aparece junto ao NFC incorporado ao verbo em (a) e marca que o argumento a que ele se refere encontra-se em outro domínio sintático, i.e., no SN e não no predicado. Em (b) o NFC está no SN e com R2, pois ele retoma anaforicamente o nome do SN em (a). Logo, com base em (a) e (b), podemos interpretar o R2 como um *pro-index* na terminologia de Haspelmath.

É de se notar, ainda no nosso exemplo, que *puy*, que significa puramente ‘cobra’, forma um único argumento P quando “completada” pelo elemento relacional seguido do NFC. Semanticamente, poderíamos compreender que há um primeiro item do sintagma

que expressa uma espécie de núcleo do significado, talvez com uma carga do referente mais específica e de uma classe gramatical mais aberta. Já o NFC concede uma carga de menor lexicalização, cuja função, além de indicar a forma do referente, organiza no seu próprio sistema os substantivos da língua em classes semânticas, que no exemplo se refere a algo cilíndrico e flexível, ou seja, um grau de especificidade dentro da macro-categoria de NFC que em Mundurukú se remete a forma física.

Observamos ainda em (13) que o NFC se repete, deixando uma espécie de cópia, porém desassociada da palavra semântica nuclear. O fato dessa desassociação dentro do próprio domínio sintático requer o uso obrigatório do R2, expressando descontiguidade. Quanto à repetição do classificador em posição pré-verbal (“incorporação por subida”) e junto ao R2, poderíamos inclusive considerar um potencial caso de concordância, porém segundo Gomes (2006, p.216) os NFC em Mundurukú não estariam em processo de concordância, pois a presença destes na estrutura verbal é compreendida como fenômeno de *incorporação* nominal e a ocorrência com o sintagmas precedidos de um R2 ou em verbos como em *e=su-bu-aoka* seria um caso de retomada anafórica. Vemos isso mais especificamente no exemplo em (13b).

Vejamos agora o exemplo (14) retirado de Gomes (2006, p.88), que nos mostra como um NFC também pode ser o argumento S, na oração ‘a borboleta caiu’:

- (14) **warepupu** **@ -dup** **o'=tup-'at**
 borboleta R1-NFC 3S=R2.NFC-cair.PRF
 'A borboleta caiu.'

Algo semelhante ao que vimos anteriormente ocorre também com *warepupu* (borboleta) em função de sujeito de uma oração intransitiva, quando o NFC *-dup* se agrega ao relacional e classifica o núcleo do nominal na categoria ‘foliforme’. O NFC *tup* (alomorfe de *dup*) é repetido anaforicamente no predicado por um processo de incorporação nominal, posicionando-se entre o clítico e a raiz verbal, formando um composto (GOMES, 2006).

Na seção 3.1.1 tratamos sobre o SCN fora da palavra verbal. Propomos então aqui uma demonstração resumida de como ocorre a sua marcação em alguns outros contextos gramaticais que envolvem o sintagma nominal.

Demonstrativo – assim como ocorre com os quantificadores, vemos que o demonstrativo pode anteceder uma retomada anafórica contendo o NFC, como em (15), em que *sapokay @-dap* se repete após o demonstrativo *ija*, sem o seu núcleo *sapokay*. Os pronomes em Mundurukú não ocorrem com relacionais (Gomes, 2006, p.120).

(15) **sapokay @-dap ija @-dap tap-rut**

galinha R1-pena esta R1-pena R2.pena-ser.branco

'As penas desta galinha são brancas.'

(lit. 'As penas da galinha, penas desta, são penas brancas')

Numeral – O NFC 'a 'arredondado' se refere a uma casa, elemento citado na narrativa anterior, que sucede o quantificador *pug*).

(16) **pug) @-'a @-be ma ade jijã wuyjuyu)**

um R1-NFC R1-em mesmo muitos muito gente

'Em uma casa (lit. em uma arredondada), morava muita gente.'

Modificador de nome possuído - em uma estrutura possessiva em Mundurukú, o nome possuído é classificado. Gomes (2006, p.183)

(17) **[[ako] @-dup] o'=tup-'at**

bananeira R1-folha 3S=R2.folha-cair.PRF

'A folha da bananeira caiu.'

Modificador adjetival – os nomes que assumem função de modificador adjetival são identificados no trabalho de Gomes como “nomes descritivos” (2006, p. 185). O NFC que segue o descritivo '*it'it* 'o pequeno' é o mesmo encontrado no nome modificado, '- 'a' 'redondo'.

(18) **wa'i @-'a 'it'it @-'a @-kug) o'=ju**

balde R1-NFC pequeno R1-NFC R1-com 3S=ir.PRF

'Ela foi com um baldezinho.'

(lit. 'Ela foi com a redondeza balde, a redondeza pequena')

Pronome– conforme vimos em 13b, o SN ‘*i-bu*’ opera como um pronome, exercendo a função de P, retomando por meio de recurso anafórico o sintagma nominal original ‘*puy @-bu*’ (cobra). Na sentença abaixo, observamos o mesmo fenômeno com os mesmos elementos lexicais de 13b. Nota-se que na segunda oração ‘*õn i-bu @-aoka-m*’ o NFC alude a todo o SN por uma retomada anafórica. (Gomes, 2006, p. 260)

(19) **õn puy @-bu @-jojo-m xepxep @-pu. õn i-bu @-aoka-m**
 Eu cobra R1-NFC R1-ver-IPRF dois R1-NFC eu R2-NFC R1-matar-IPRF
 ‘Estou vendo duas cobras. Vou matá-las.’

Em se tratando do sintagma nominal na língua Mundurukú, podemos traçar um percurso do que acabamos de apresentar. Na maior parte das categorias gramaticais que foram descritas no trabalho de Gomes (2006) há ocorrência de associação dos Nomes com Função Classificadora. Das sete categorias analisadas, apenas uma delas não possui NFC, a saber, o elemento possuidor. Os NFC nos nomes são os mais abundantes, e é o que vimos em (14) acima, em que a raiz ‘*warepuru*’ (borboleta) recebe o nome classificador ‘*@-dup*’, completando o sentido do primeiro elemento com uma característica física de ‘foliforme’.

O demonstrativo apresentado em (15) é um exemplo da como essa palavra (*‘ija*’) introduz uma repetição do nome inalienável que funciona como classificador em retomada anafórica. O interessante é que em casos assim percebe-se a ocorrência tripla dessa categoria (NFC), que acaba finalmente na incorporação. Também é comum, segundo Gomes (2006, p.265), que a ocorrência do SN demonstrativo esteja junto ao nome inalienável sozinho, sem que SN nome + nome inalienável esteja na oração. No caso do numeral, visto no exemplo (16), este é sucedido pelo NFC, que é introduzido pelo R1. Observamos que nesse caso o numeral (*‘pug*)’assume hipoteticamente a posição que seria de ‘*ok*’ casa.

Em uma estrutura de posse viu-se que é junto ao objeto possuído que o NFC se associa, enquanto que ao possuidor tal categoria não consta em Mundurukú. Uma situação análoga é vista com o modificador adjetival, que retoma anaforicamente o NFC do nome modificado anterior. Nota-se que não se trata de um caso de sintagma único entre o modificador e modificado, pois há uma relação apositiva entre eles.

Já quanto ao NFC ligado à função de pronome, convêm destacar que na língua há os pessoais, quantificadores e demonstrativos. Estes constituem uma classe fechada e não co-ocorrem com os relacionais. No SN quando na função de sujeito, os clíticos tendem a

estar sozinhos na oração, sem que co-ocorram os pronomes independentes (1ª Pessoa). A 3ª pessoa é expressa pelos pronomes não-pessoais e pelos nomes.

O esquema proposto abaixo na tabela 3.4 resume as ocorrências no nome classificador dentro do SN em Mundurukú.

Tabela 3.4 – Os NFC no sintagma nominal

Nome	Demonstrativo	Numeral	Possuidor	Possuído	Modificador Adjetival	Pronome
Sim	Sim	Sim	Não	Sim	Sim	Sim

Para finalizar esta seção sobre o sistema de classificação nominal da língua Mundurukú, elaboramos uma tabela de indexadores verbais com as informações analisadas em 3.2.

Tabela 3.5 - Indexadores verbais em Mundurukú

Indexador (exemplos)	Pessoa	Número	Semântica	Função sintática	Tipo de indexação	Forma
<i>bu-</i>	3	sing	class. que indica “forma”: cilíndrico e flexível	S/P	<i>cross-index</i> ou <i>pro-index</i> (a depender do relacional: R1 - cross R2 - pro)	Nomes incorporados ou argumentos independentes da oração combinados a nomes
<i>up-</i>			class. que indica “forma”: foliforme			

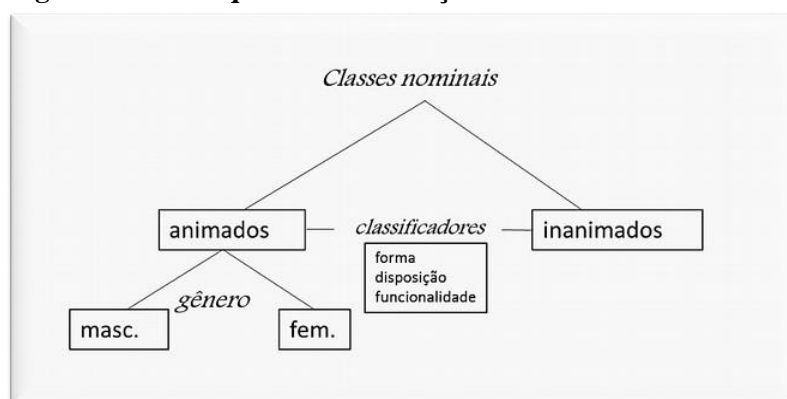
3.2.2 Kubeo

O Kubeo é uma língua pertencente à família Tukano, mais especificamente ao grupo Tukano Oriental, falada ao longo dos rios Uaupés, Cuduyarí e Querarí, localizados na parte noroeste da bacia amazônica. Conta com cerca de 6 mil usuários, distribuídos nos territórios do Brasil e da Colômbia, segundo o shite Ethnologue. Grande parte da população que fala o Kubeo também tem conhecimento de outras línguas da região, não só pelo

contato intenso, como também pela questão cultural do matrimônio obrigatório entre grupos étnicos distintos (ver Chacon 2012). A literatura que usaremos como base de análise na presente seção sobre a língua Kubeo está majoritariamente concentrada em Thiago Chacon (2012) e Morse e Maxwell (1999). Bem como fizemos na breve descrição dos sistemas de classificação nominal do Mundurukú, trataremos também aqui o mesmo procedimento, partindo dos dados mostrados na tabela 3.1.

De acordo com Chacon (2012), os nomes em Kubeo são classificados de modo hierárquico por um sistema de categorias gramaticais, com grande transparência semântica nos termos. Podemos considerar que existem dois níveis de classificação nominal: um de maior hierarquia, dividindo o léxico nas classes *animado* e *inanimado*; e um de outro de menor hierarquia subclassificando os nomes animados em gênero masculino e feminino. Tanto alguns nomes animados quanto inanimados possuem ainda classificadores nominais que indicam forma, disposição e funcionalidade (Chacon 2012, p.235)¹².

Figura 2 – hierarquia da classificação nominal em Kubeo



Fonte: criado pelo autor com base em Chacon (2012).

De acordo com Morse e Maxwell (1999), os verbos do Kubeo se dividem sintaticamente em estativos e dinâmicos, e não se combinam com classificadores, apenas com morfemas que indicam o gênero do argumento S/A na terceira pessoa do singular. As informações semânticas de pessoa, número e gênero se encontram fundidas nas mesmas marcas com outras categorias verbais como evidencialidade e tempo gramatical. As marcas de pessoa, número e gênero correferenciam o argumento em função A ou S, como ilustrado abaixo (Chacon 2012).

¹² Nota-se que o autor denomina o nível hierárquico superior de *classes nominais*, reservando o termo *gênero* para uma ramificação da classe dos animados (Chacon 2012, p. 235).

(20) **yawi yi-re há-mi**
 onça 1SG-N.NM ver-3MSC.PS
 ‘A onça me viu’

(21) **yawi makarõ-i ki-ame**
 onça floresta-LOC estar-PS.3MSC
 ‘A onça vive no mato’

A presença do argumento como um nome não é obrigatória, o que mostra que as marcas de gênero são *cross-indexes*, como ilustrado a seguir:

(22) **aruka mi paki? kirami ki-kibe**
 onde 2SG.PSS pai? casa estar-3M.INF
 ‘Onde está teu pai? está em casa’

Considerando o SN do Kubeo, encontramos o SCN marcado no pronome pessoal, no demonstrativo, no nome e no marcador descritivo, assim como mostra-se a seguir.

Pronome pessoal – Como visto na sentença (20), o Kubeo possui a mesma marcação de gênero nos pronomes pessoais livres, porém, em um dos paradigmas verbais, a língua estendeu a marcação de gênero para a primeira pessoa, fato que não ocorre com os pronomes pessoais. Nestes, apenas encontramos a marcação de gênero na 3ª pessoa: *i* para o masculino ‘ele’, e *õ* para o feminino ‘ela’:

Tabela 3.6 – Pronomes pessoais em Kubeo

Pronome pessoal	Pessoa/número/gênero
yí	1SG
mí	2SG
ĩ	3M
õ	3F
maha	1INC
ñiha	1EXC
miha	2PL
na	3PL

Demonstrativo – os demonstrativos se dividem em *proximais* e *distais*, a depender da posição do observador para com o objeto referenciado. Esta classe se aproxima muito dos

nomes em sua categorização interna, no que se refere às flexões (CHACON, 2012).

Tabela 3.7 – Demonstrativos em Kubeo

Categorias gramaticais	Formas proximais	Formas distais
INANIMADO CONTÁVEIS	i=NP _{HEAD}	ãdĩ# NP _{HEAD}
INANIMADOS MASSIVOS	i=e	ãdĩ=e
INANIMADO CONTÁVEIS GENÉRICO	‘jo	ãdõ
ANIMADO MASCULINO	‘jãĩ	ãjĩ
ANIMADO FEMININO	i-ko	ãdõ
ANIMADO PLURAL	i-dã	ãdĩ-dã ãdĩã

Enquanto os demonstrativos proximais tem sua forma básica em i=, nas formas distais a base é o afixo independente ãdĩ# ‘este’. Ambas formam o plural animado pelo sufixo -dã no lugar do classificador clítico =wi (animado coletivo), que usa o sufixo plural (!) para codificar a forma plural animada. Em algumas das construções podemos perceber fusão da base com o sufixo inanimado (‘jo) e masculino (‘jãĩ).

Do ponto de vista morfossintático, observa-se que há concordância do demonstrativo proximal com o classificador; no caso do distal há uma tendência de se manter sem acordo com o *head noun*. Observam-se os exemplos a seguir:

- (23) a. **ãdĩ biki-ki**
 este velho-MS
 ‘este homem de meia idade (30 ~ 40)’
- b. **ãjĩ biki-ki**
 ESTE.MSC velho-MS
 ‘este velho homem (60 acima)’

Segundo Chacon (2012, p.329) a raiz biki é polissêmica, podendo equivaler a ‘velho’, ‘grande’, ‘respeitoso’, ‘crescer (verbo)’ e ‘mais velho’. A concordância entre o demonstrativo e o *head noun*, como se vê em (23), leva a uma segunda interpretação frasal.

Possessivo – há variadas formas de marcação do possessivo em Kubeo, sendo a mais produtora por meio de *determinantes possessivos* (CHACON, 2012, p. 315).

O determinante =hi é a forma mais simples de se codificar uma possessão na língua. Este morfema funciona como um ‘possessivo genérico’, em que o elemento possuidor precede o possuído, com o morfema de possessivo ligado a ele. Essa é a representação básica da estrutura de posse em Kubeo:

NP_{POSSUIDOR} hi= NP_{POSSUIDO} (NOME OU CLÍTICO)

Os determinantes possessivos representam a forma primária de indicar posse em Kubeo. Encontram-se por meio de clíticos ou formas independentes, neste caso seguidas pelo símbolo #. Estes são complementados por um nome, desde que a estrutura seja composta. Já as formas que se tratam de elementos clíticos precisam ser combinados por um outro nome ou uma ênclise.

Tabela 3.8 – Determinantes Possessivos em Kubeo

	PADRÃO	EXCEÇÃO: ALGUNS TERMOS DE PARENTESCO
Primeira Pessoa Singular	‘hi#	hi=
Segunda Pessoa Singular	‘bi#	bĩ=
Terceira Pessoa Singular	NP-i-	‘hi#
Primeira Pessoa Plural Inclusivo		bāhẽ#
Primeira Pessoa Plural Exclusivo		ḵihẽ#
Segunda Pessoa Plural		bĩhẽ#
Terceira Pessoa Plural		‘dẽ#

Seguem dois exemplos de orações possessivas básicas. Em (24a) o pronome determinante possessivo de segunda pessoa singular antecede o nome, que por sua vez é acompanhado do seu classificador com semântica de ‘objeto longo com ponta aguda’. Em (24b) o pronome de primeira pessoa plural inclusivo é seguido do verbo com sufixo ‘nominalizador contável’. (CHACON, p. 315):

(24) a. **‘bĩ bĩbĩ=jo**
 seu beija-flor=CL.LONG
 ‘seu beija-flor’

b. **bāhẽ ki-dõ**
 NOSSO.INC existir-NMZ.CONT
 ‘o lugar onde nós vivemos’

- (26) **bõa=hĩ-ki** 'kihĩ-ki
 fish=DIM-MSC small-MSC
 'a small little fish'

Em resumo, segue a tabela esquemática das classes em Kubeo, respondendo à pergunta: há marcação de SCN no sintagma nominal da língua?

Tabela 3.10 – Marcação da classificação nominal no SN do Kubeo

Nome	Demonstrativo	Numeral	Possuidor	Possuído	Modificador Adjetival	Pronome
Sim	Sim	Sim	Não	Sim	Sim	Sim

Já com relação ao sintagma verbal, nas tabelas 3.11 e 3.12, mostramos um resumo do sistema Kubeo conforme seus dois paradigmas de indexadores de gênero na evidencialidade não marcada: paradigma de passado recente/presente (classe I) e o paradigma de passado distante/genérico (classe II) (a semântica do paradigma depende da classe semântica do tema verbal, (cf. Chacon 2012):

Tabela 3.11 – Paradigmas verbais de flexão para evidencialidade não-marcada (CLASSE I)

Indexador	Pessoa	Número	Semântica	Função sintática	Tipo de indexação	Forma
<i>-bi</i>	3	sing.	masculino	S, A	cross-index	sufixo
<i>-biko</i>			feminino			
<i>-ma</i>		pl.	plural animado			
<i>-wi</i>	1/2/3inanimado	sg/pl.	sincrético			

Tabela 3.12 – Paradigmas verbais de flexão para evidencialidade não-marcada (CLASSE II)

Indexador	Pessoa	Número	Semântica	Função sintática	Tipo de indexação	Forma
<i>-ame</i>	3	sing.	masculino	S, A	cross-index	sufixo
<i>-ako</i>			feminino			
<i>-ima</i>		pl.	animado			
<i>-awĩ</i>	1/2/3inanimado	sg/pl. inclusivo	sincrético			
<i>-karã</i>	1	pl. exclusivo	animado			

A diferenciação básica entre estas duas classes de índices do Kubeo se resume basicamente a que na classe I vemos que os índices possuem na sua totalidade uma configuração que referencia à presença da marcação do gênero, pessoa e número no alvo (palavra verbal), compartilhando normalmente o mesmo sufixo e gerando correferência entre a indexação e o argumento.

Ainda com relação às pessoas gramaticais envolvidas, a generalização que fizemos na seção 3.1.2 pode ser retomada, e vemos aqui o total de 8 índices com marcação da 3ª pessoa, 5 com marcação de 1ª e apenas 2 deles fazem referência à 2ª pessoa gramatical.

Nota-se por fim uma peculiaridade interessante referente às distinções de gêneros na classe II dos paradigmas verbais se comparados com as mesmas atribuições no caso dos pronomes pessoais, o que parece tratar-se de um fenômeno mais raro nas línguas. Poderíamos quanto a isso hipotetizar, e aprofundar buscas em um estudo posterior, que diacronicamente a língua possuía, mas que tenha reduzido em algum momento a expressão de gênero mais ampliada no seu sistema de pronomes.

3.2.3. *Tehuelche*

O Tehuelche é uma língua ou ainda um complexo dialetal da família Chon,

atualmente extinta do sul da Patagônia argentina¹³. Segundo Ana Fernandez Garay (1998)¹⁴, os dialetos do complexo Tehuelche abrangem historicamente, no total, 5 grupos (ver figura 3).

- *Tehuelche setentrional boreal*– corresponde ao dialeto *querandí* e desaparecido supostamente no início do século passado.
- *Tehuelche setentrional austral* – o dialeto era correspondente ao *güniina iájech*, hoje em extinção.
- *Tehuelche meridional boreal* – eram falantes do teushen; viviam entre os rios Chubut e Santa Cruz. Também desapareceram no início do século XX.
- *Tehuelche meridional austral* – os falantes deste grupo eram conhecidos igualmente como aonek'enk, patagones, chewelches, ainda tem falantes vivos que foram informantes para os dados de Garay.
- *Selknan* – grupo encontrado na Terra do Fogo. Certamente hoje não existem mais falantes; à época da coleta dos dados – ano de 1998 - os falantes eram já idosos.

Figura 3 – zonas dialetais do Tehuelche e outras línguas da Patagônia



Fonte: Garay (1998, p.35)

¹³ Segundo uma reportagem divulgada pelo site argentino *Tiempo Sur*, divulgada em 04 de janeiro de 2019, a senhora Dora Machado foi a última falante de Tehuelche que se encontrava viva até então. Depois do seu falecimento nesse mesmo dia, já não há falante representante da língua Tehuelche. Link para acesso: <https://www.tiemposur.com.ar/nota/163880-fallecio-la-ultima-descendiente-parlante-de-tehuelches>

¹⁴ Esta será a nossa referência para todas as informações referente ao Tehuelche.

Veremos nesta seção que o Tehuelche (auto-denominado *aonek'enk*) apresenta um sistema de classificação nominal bastante peculiar e ao mesmo tempo complexo, sobretudo no seu sintagma nominal, o que nos levou a dar uma atenção especial além do sintagma verbal.

Como vimos na tabela 3.1, o único tipo de SCN que encontramos na língua Tehuelche é o gênero, e este se apresenta como masculino, feminino ou neutro, majoritariamente marcado no *target*. No entanto, existe na língua uma neutralização das categorias de gênero a depender do tipo de alvo, o que gera um sincretismo - nas palavras de Garay. Em certos contextos, feminino se iguala morfológicamente ao masculino e fazem contraposição ao neutro; em outros, acontece a neutralização do feminino com o neutro, fazendo contraposição ao masculino (1998, p.190). Desta forma, para que saibamos a qual gênero pertence *p'ajXen* 'faca', é necessário que se observe a combinação deste substantivo com outros elementos da oração, já que não é explícita qualquer morfologia de referência a gênero no controlador do SN. Se para dizer 'faca afiada' acrescentamos o substantivo¹⁵ *kwese-nk* 'afiado' após *p'ajXen* (*p'ajXen kwesenk*, portanto), eliminamos a possibilidade do substantivo modificado pertencer aos gêneros feminino ou neutro pois, pelo sincretismo, o sufixo derivativo *-nk* (masculino) no modificador *kwese-nk* é suficiente para identificarmos que *p'ajXen* pertence ao masculino, contrapondo-se ao sufixo *-n(l)* indistintamente atribuído aos outros dois gêneros do Tehuelche.

Estrutura de posse - Ainda no âmbito do sintagma nominal, observemos a seguir uma situação de construção possessiva, na qual seguimos com o contraste masculino vs. feminino/neutro (Garay 1998, p.191). Note que no Tehuelche o objeto possuído concorda com o possuidor. O *target* recebe o sufixo *-e* para indicar o masculino (27), enquanto os gêneros feminino e neutro são marcados com o sufixo *-n* (28).

(27) t jatene šome²¹⁶

t – jatene some[?]

3 + S S

seu pedra boleadeira

“A pedra da boleadeira”

¹⁵ Garay (1998) reconhece como substantivo o adjetivo qualificativo.

¹⁶ O símbolo ? corresponde foneticamente a uma oclusiva glotal.

(28) T 'oteln̄ eqon

T - 'oteln̄ e – qon

3 + S 1 + S

seu + olho minha avó

“O olho da minha avó”

Outro caso que mostra ainda o sincretismo masculino vs. feminino/neutro é em relações apositivas em que um verbo estativo indicando propriedade recebe as marcas sufixais *-nK* (masc.) e *-n* (fem./neu), conforme os exemplos (29-31) (:192).

(29) kaw: č'aj-n (neutro)

S S-N

toldo o grande

“o grande toldo”

(30) le' t'a:rte-nK (masc.)

S S-M

água a suja

“a água suja”

(31) ka:rken k'ete-n (fem.)

S S-F

toldo o grande

“o grande toldo”

Uma segunda situação de relação apositiva constitui um novo sincretismo - masculino/neutro vs. feminino - ocorrendo quando um substantivo é aposto de outro. Nos substantivos alvo de gênero masculino ou neutro encontramos o sufixo *-k'o*; já aqueles de gênero feminino marcará com *-k'on* (:193).

(32) **aone-k'o** **'alen** (masc.)

S-M S

sulista homem

“homem sulista”

(33) **kenaj-k'on** **mer** **ka:rken** (fem.)

S-F D S

a que é de onde essa mulher

“de onde é essa mulher?”

(34) **welom-k'o** **kaj** (neu.)

S-N S

Tudo camada

“todas as camadas”

Como constatamos até o momento, em todos as situações trazidas por Garay (1998) evidenciam que a marcação do gênero em Tehuelche se expressa apenas no alvo do SN. No entanto, a autora identificou que existe um caso especial que mostra o controlador portando a expressão de gênero. Trata-se de uma relação linear em que o substantivo se encontra posposto ao verbo, não importando a função sintática que este exerce no predicado, i.e., a sua função argumental. Neste caso, teremos o sincretismo masculino vs. feminino/neutro. Com o sufixo *-je* é marcado o masculino, enquanto *-ne* marca feminino ou neutro (:193). Vejamos em (35) um exemplo do controlador pós- verbal *jajke[ne]* sendo marcado no gênero neutro (:193).

(35) **Xeple?** **'ejkejonškn** **jajkene**

Xeple? ? - ejkejon - š¹⁷ - kn jajke-ne (neutro)

S 3 + V + EP + MR S-N

lagarto o temer fogo

“O lagarto teme o fogo”

¹⁷ A constante aparição do afixo *-š* (especificador do sujeito) é recorrente entre um verbo e seus complementos ou adjuntos. No caso de (13), antecede o morfema de modo real.

De acordo com Garay (1998), em nível semântico o gênero também pode sofrer alteração para atender a um “emprego informativo” em contextos que envolvam a ideia de número, pluralidade ou totalidade (1998:194). Resumidamente são os seguintes:

i. Mudança de gênero para indicar oposição entre a unidade e o par ou conjunto de objetos, ou seja, quando o falante quer individualizar um elemento que naturalmente faz par ou faz parte de um conjunto de vários, essa unidade é sempre dada no masculino, i.e., pé, orelha, bota. Porém, quando se remete ao par ou conjunto, será dado no neutro.

ii. Indicar oposição entre unidade e coletivo de alguns objetos, como no caso de mosquito, couro, ovo, dedo, etc.

iii. Informar ao ouvinte sobre a oposição parte/todo. Vejamos em (27) e (28) como se opera a marcação do gênero na distinção entre ‘serra’ e ‘campo’ em Tehuelche:

(36) ošge:wT ʔaj

o - š - ge:Wt ʔaj

1 + PL + S Func.

nosso campo em

“...em nosso campo”

(36) ge:wT kawr

S S¹⁸

serra sobre

“sobre a serra”

Nota-se que o Tehuelche tem tanto em (27) como em (28) o mesmo substantivo *ge:Wt* para expressar as semânticas de ‘campo’, remetendo a uma ideia de totalidade, e ‘serra’, cuja ideia é de uma parte (do campo). A implicação semântica que acaba por alterar o gênero – neutro em (27) e masculino em (28) – desta vez não traz como alvo um nome, mas curiosamente uma partícula funcional. Assim que a presença de *ʔaj* indica que o substantivo que o precede pertence ao gênero gramatical neutro, bem como a sua versão *kawr* atribui ao seu termo antecessor *ge:wT* o gênero masculino (1998:195).

¹⁸ Acreditamos que tenha havido um problema de glosagem, quando no lugar de S (substantivo) seria possivelmente Func. (funcional), conforme o raciocínio da autora textualmente expresso na página 195.

Em Tehuelche existem alguns substantivos que apresentam uma possessão inerente, que são derivados de verbos transitivos do grupo I (veremos mais adiante nesta seção). Alguns exemplos são: j-eno (meu amigo); m-eno (teu amigo); o-k-w-eno (nosso amigo [dual]) e o-s-w-eno (nosso amigo [plural]). Esses mesmos prefixos são usados com verbos em correferência aos participantes 1 de verbos intransitivos (1998:197).

De acordo com Garay (1998, p.264), a expressão morfológica de gênero no verbo categoriza a terceira pessoa singular em masculino e feminino, que recebem a marca prefixal *k-*, e o neutro, que recebe o prefixo ʔ (uma oclusão glotal). O masculino ou feminino possuem formas distintas em outros paradigmas da língua.

Sobre a morfologia dos verbos Tehuelche, observa-se que:

- Praticamente não apresentam variação morfológica, só quando a raiz do verbo termina em consoante, e recebe uma vogal de apoio.
- São três tipos de verbos: existenciais, intransitivos (que se dividem em 2 grupos [grupo 1 e 2]), e transitivos (também divididos em 2 grupos [grupo 1 e 2]).
- Existenciais – fenômenos da natureza (sem expressão de gênero)
- Intransitivos – apenas uma valência.

Os verbos do grupo 1 (intransitivos ou transitivos) são aqueles que indexam o argumento com morfemas de gênero. Estes marcam o gênero de seus argumentos e são: (1) verbos intransitivos que atribuem uma propriedade ao seu único argumento (S) (Garay 1998: 261), conforme em (37) e (38) abaixo, e certos verbos transitivos que marcam o gênero dos objetos (P) conforme em (39) e (40) (Garay 1998: 264-6):

(37) kastersk' e wajenK

k - aster - s - k'e wajenK (masc.)
 M ser.profundo - EP - MR poço
 “o poço é profundo”

(38) ʔa:rsk'n te:m

ʔ - a:r - s - k'n te:m (neutro)
 N ser.seco - E.P. - MR terra
 “a terra está seca”

(39) ekernoskoT

e - k - erno - s - koT

1 + 3 + V + EP + T.F.M

eu o deixar

“eu vou deixá-lo”

(40) exa:t'en ? - ernosk'

e - xa:t'en ? - erno - s - k'

1 + S3 + V + E.P. + M.R.

eu comida a deixar

“a comida, eu a deixei”

Notemos como (30) mostra que a expressão nominal do argumento não é obrigatória, o que nos leva a analisar as marcas de gênero como um *cross-index*. Conforme comentamos anteriormente, os indexadores *k-* e *ʔ-* são os mesmos prefixos que marcam o possuidor em construções “possessivas” inalienáveis (Garay 1998: 197). Na tabela abaixo oferecemos um resumo dos fatos do Tehuelche.

Tabela 3.13 – Indexadores verbais em Tehuelche

Indexador	Pessoa	Número	Semântica	Função sintática	Tipo de indexação	Forma
<i>k-</i>	3	singular	masculino ou feminino	S, P	<i>cross-index</i>	prefixo
<i>ʔ-</i>			neutro			

3.3 Considerações contrastivas dos SCNs no corpus trilíngue

Para esta última seção, ainda engajados em explorar o corpus específico das três línguas que acabamos de apresentar em três etapas individuais, discorreremos, considerando novamente uma abordagem translíngüística, acerca de alguns pontos em comum e outros que divergem dos que expusemos nas páginas precedentes. Vale recordar

que as línguas Mundurukú (Tupí), Kubeo (Tukáno) e Tehuelche (Chon) pertencem a grupos genéticos distintos, porém atendem igualmente aos critérios tipológicos necessários para o nosso estudo, ou seja, todas possuem sistemas de classificação nominal que interagem com o sistema verbal da língua.

Dado que buscamos identificar as diferentes configurações dos SCNs, o primeiro ponto que devemos retomar das línguas é qual seriam os tipos de sistemas de classes a que elas apresentam. E aqui reforçamos a preferência por não distinguir gênero gramatical de classe nominal, com base em Dixon (1986) e Grinevald (2000) – ver seção 2.2.2 – apesar de reconhecer que há autores, como Krasnoukhova (2012) e Regúnaga (2012), que fazem claramente essa repartição conceitual e foram certamente imprescindíveis para compor a base conceitual de nossa exposição até o momento.

Vimos que o Mundurukú é uma língua que não explicita em sua morfologia nenhum sistema de gênero, contrariamente ao que ocorre em Kubeo e em Tehuelche. Inclusive, uma breve pesquisa em descrições de outras línguas do tronco Tupí nos permite atestar que o gênero gramatical não consta nos seus sistemas de classificação nominal. Esse dado é asseverado por Aikhenvald (2000), quando prevê uma série de famílias de línguas indígenas que também são desprovidas dessa marcação, mas que com certa frequência possuem sistema/s de classificador/es.

The Tupí, Pano, Carib, Yanomami, Maku, Tacana, Quechua, Piraha, Aymara, Jivaro and Ticuna families, isolates Aikana, Koaia, and four of the Arawak languages (Terena, Amuesha, Chamicuro, Bahwana) of South America have no genders or noun classes; however, they often have classifiers of other types. (AIKHENVALD, 2000: 80).

O Mundurukú, como observa Gomes (2006), não constitui uma classe de classificadores em si como acontece em Kubeo, mas sim possui uma série de nomes que exercem com alta recorrência o papel sintático de classificadores (NFC), o que abre um possível questionamento terminológico sobre classificá-los ou não de fato como ‘classificadores’. Esses nomes com função classificadora operam com o auxílio de elementos relacionais servindo de base para nomes. Em contraposição, os 16 classificadores do Kubeo constituem uma classe morfológica, podendo expressar semanticamente tanto seres animados como inanimados. Apesar dessas diferenças entre as duas línguas, vemos uma grande semelhança que reside na informação veiculada de propriedades físicas do objeto referente. O Kubeo além dos classificadores ainda possui o SCN de gênero, resultando num duplo sistema, e se posicionando talvez em uma posição

intermediária com relação ao Mundurukú e o Tehuelche.

Já partindo para a análise dos índices verbais, vimos que todas as três línguas indexam a 3ª pessoa, assim como ocorreu de forma padrão no corpus mais amplo com 23 línguas. Os dados do Kubeo mostraram, porém, que nos paradigmas verbais é possível encontrar, além da 3ª, flexões de 1ª pessoa singular com informações sobre o gênero do argumento. Quanto à informação de número, o singular é padrão em todas, e apenas o Kubeo marca o plural, porém gênero somente é expresso no singular nesta língua, igualmente ao que vemos para o Tehuelche.

Tanto o Mundurukú como o Tehuelche mostraram indexar sujeitos de orações intransitivas (S) e pacientes (P), enquanto que o Kubeo indexa sujeitos de orações intransitivas (S) e transitivas (A). O *cross-index* foi o tipo de indexação que encontramos com abundância no corpus.

Por fim, a forma de indexação diverge bastante entre as três línguas: o Múndurukú mostra incorporação nominal e também associa argumentos independentes da oração combinados a nomes. Já o processo de afixação ocorre nas outras duas línguas, sendo sufixo em Kubeo e prefixo em Tehuelche.

Como fizemos nas seções anteriores, propomos uma última tabela mostrando de forma esquematizada as características essenciais que encontramos na indexação verbal cruzando as informações das línguas Mundurukú, Kubeo e Tehuelche.

Tabela 3.14 – Quadro contrastivo dos SCNs no *corpus* trilíngue

Língua	Pessoa	Número	Tipo de SCN	Semântica	Função Sintática	Indexação	Forma
Mundurukú	3	sing.	NFC	Propr.física s	S/P	<i>cross-index</i>	Nomes incorporados ou argumentos independentes da oração combinados a nomes
Kubeo	3/1	Sing.	CL gênero	Propr. físicas M, F	S/A	<i>cross-index</i>	sufixo
Tehuelche	3	singular	gênero	M, F, N	S/P	<i>cross-index</i>	prefixo

CONSIDERAÇÕES FINAIS

Ao ser iniciado o trabalho de pesquisa, constatou-se que o continente americano (ou sul-americano) abriga em seu território uma vastíssima variedade de línguas que revelam uma igualmente ampla diversidade gramatical. Essa diversidade, e sua aparente diferença com relação às línguas indo-europeias, motivou a busca de novas conceituações e aparatos teórico-metodológicos, como a noção de classificação nominal e indexação, dentro da área da tipologia linguística.

Tomou-se como objetivo geral explorar como as línguas indígenas da América do Sul marcam o *Sistema de Classificação Nominal (SCN)* em sua estrutura. Pode-se avaliar que esse propósito foi atingido com sucesso, considerando que foi possível compreender a dinâmica do gênero gramatical, dos classificadores e classes nominais, partindo de uma visão panorâmica do fenômeno para o seu aprofundamento em um corpus específico e comparativo.

Considero que os resultados obtidos neste trabalho dissertativo devem ser levados em conta para futuras investigações no campo da linguística tipológica, pois muito ainda há que se explorar na estrutura verbal, por exemplo, a relação do seu núcleo com as informações dos gêneros e classificadores e as motivações que levariam um sistema a ser explícito morfológicamente. Algo muito relevante e que não abordamos em nenhum momento foi em relação às implicações diacrônicas e fonológicas na constituição formal dos sistemas.

Alcançamos com sucesso identificar em 23 línguas (e mais especificamente em 3) como se dá a relação do sistema de classificação nominal (SCN) com informações de outras categorias, tais como a Pessoa, Número, Função Sintática do argumento, e em alguns casos Tempo, Aspecto, Modo e Modalidade.

Considero finalmente que o árduo processo de pesquisa bibliográfica ao cabo de pouco mais de dois anos me permitiu abrir os horizontes da pesquisa profissional, o que considero que me foi inédito, e afinco meu desejo de poder compreender ainda mais como as línguas do mundo externam seus inúmeros modos de “classificar”, “generalizar” e “indexar”.

REFERÊNCIAS BIBLIOGRÁFICAS

ADELAAR, Willem F. H. 2004. *The Languages of the Andes*. Cambridge University Press.

AIKHENVALD, Alexandra Y. *Classifiers: A typology of noun categorization devices: A typology of noun categorization devices*. OUP Oxford, 2000.

AIKHENVALD, Alexandra Y., and GREEN, Diana. 1998. "Palikur and the Typology of Classifiers." *Anthropological Linguistics* 40 (3): 429–80.

AIKHENVALD, Alexandra. 2003. *A Grammar of Tariana, from Northwest Amazonia*. Cambridge Grammatical Descriptions. Cambridge University Press.
Disponível em : <http://print.google.com/print?id=n4EwdpczyP8C>.

AIKHENVALD, Alexandra. 2003. *Classifiers: A Typology of Noun Categorization Devices*. Oxford Studies in Typology and Linguistic Theory. Oxford, New York: Oxford University Press.

AIKHENVALD, Alexandra Y.; GREEN, Diana. *Palikur and the typology of classifiers*. *En Language at Large*. Brill, 2018. p. 394-450.

BICKEL, Balthasar. 2010. *Grammatical Relations Typology*. *The Oxford Handbook of Linguistic Typology*, November.

BIRCHALL, Joshua Thomas Rigo. *Argument marking patterns in South American languages*. Utrecht: LOT, 2014. Texto integral disponível em: <https://www.lotpublications.nl/argument-marking-patterns-in-south-american-languages-argument-marking-patterns-in-south-american>

BIRCHALL, Joshua. 2016. *Argument Marking (ARGEX)*. In Muysken, Pieter et al. (eds.) *South American Indian Language Structures (SAILS) Online*. Leipzig: Online Max Planck Institute of Evolutionary Anthropology. Disponível em <http://sails.cld.org>

CAMPBELL, Lyle. *Typological characteristics of South American indigenous languages*. In: *The indigenous languages of South America: A comprehensive guide*, p. 259-330, 2012.
CARDOSO, Valéria Faria; DE OLIVEIRA, Mileide Terres. *Marcação de gênero em rikbaktsa (Macro-Jê)*. *Revista ECOS*, v. 17, n. 2, 2014.

CERRÓN-PALOMINO, Rodolfo. 2006. *El chipaya o la lengua de los hombres del agua*. 1. ed. Lima: Pontificia Univ. Católica del Perú, Fondo Ed.

CHACON, Thiago Costa. *Lexical and viewpoint aspect in Kubeo*. In: *Conference on Indigenous Languages of Latin America IV*, University of Texas. 2009.

CHACON, Thiago Costa. *The Phonology and Morphology of Kubeo: The Documentation, Theory, and Description of an Amazonian Language*. Tese de Doutorado em Linguística. Hawai'i, Estados Unidos: University of Hawai'i at Manoa. 2012.

COMRIE, Bernard. *Language universals and linguistic typology: Syntax and morphology*. University of Chicago press. Chicago.1989.

CORBETT, G. *Agreement*. Cambridge University Press, 2006.

CORBETT, Greville G. (ed.). *The expression of gender*. Walter de Gruyter, 2013.

CORBETT, Greville G. *Gender*. Cambridge: Cambridge University Press. 1991.

CROFT, William. *Typology and universals*. Cambridge University Press. Manchester. 2002.

DANIELSON, Swintha. 2007. *Baure: An Aruak Language of Bolivia*. Leiden: Leiden University Press.

DELANCEY, Scott. n.d. "Toward a History of Tai Classifier Systems." Tsl.7.26del. Acessado em 4 de junho de 2019. Disponível em: <https://benjamins.com/catalog/tsl.7.26del>.

DIXON, R. M. W. *The Amazonian languages*, ed. by RMW Dixon and Alexandra Y. Aikhenvald. Cambridge: Cambridge Language Surveys. 1999.

DUBOIS, Jean. *Grand dictionnaire : linguistique & sciences du langage*. Larousse, 2007.

ESTRADA RAMIREZ, Hortensia . 1996. *La Lengua Sáliba: Clasea nominales y sistema de concordancia*. (Premios Nacionales de Cultura.) Santafé de Bogotá: Tercer Mundo Editores. xxxiii+230pp.

EVERETT, DANIEL LEONARD, and BARBARA KERN. 1997. *Wari: The Pacaas Novos Language of Western Brazil*. London; New York: Routledge.

FAGUNDES, Sidney da Silva. 2000. *The Language of the Apurinã People of Brazil (Maipure/Aruak)*. State University of New York, Buffalo.

Falleció la última descendiente parlante de Tehuelches. Disponível em: <<https://www.tiemposur.com.ar/nota/163880-fallecio-la-ultima-descendiente-parlante-de-tehuelches>>. Acesso em: 13 de out. 2019

GIVON, T. n.d. "Syntax." Z.Syn1. Acesso em 04 de junho de 2019
Disponível em: <https://benjamins.com/catalog/z.syn1>.

GOMES, Dionei Moreira. 2010. "Classificação Nominal Em Mundurukú: Forma, Função e Tipologia." LIAMES: Línguas Indígenas Americanas 9 (1): 7.
Disponível em: <https://doi.org/10.20396/liames.v9i1.1461>.

GOMES, Dionei Moreira. 2010. "Classificação Nominal Em Mundurukú: Forma, Função e Tipologia." LIAMES: Línguas Indígenas Americanas 9 (1): 7.
<https://doi.org/10.20396/liames.v9i1.1461>.

GOMES, Dionei Moreira. *Estudo morfológico e sintático da língua Mundurukú (Tupí)*. 2006. Tesis Doctoral. External Organizations.

- GREENBERG, Joseph H. (ed.). 1963. *Universals of Language*. London: MIT Press, pp.
- GRINEVALD, Colette. *A morphosyntactic typology of classifiers*. Systems of nominal classification, 2000, vol. 4, p. 50-92.
- GRINEVALD. 2015. *Linguistics of classifiers*. In James D. Wright (ed.), *International Encyclopedia of the Social and Behavioral Sciences*, 811–818. Oxford: Elsevier.
Disponível em : [https:// doi.org/10.1016/B978-0-08-097086-8.53003-7](https://doi.org/10.1016/B978-0-08-097086-8.53003-7)
- HASPELMATH, Martin. *Argument indexing: A conceptual framework for the syntactic status of bound person forms*. Languages across boundaries: Studies in memory of Anna Siewierska, 2013, p. 197-226.
Disponível em : <https://doi.org/10.1515/9783110331127.197>
- HAUDE, Katharina. 2006. “*A Grammar of Movima*.” Theses, Radboud Universiteit Nijmegen.
Disponível em : <https://doi.org/10.1093/oxfordhb/9780199281251.013.0020>.
- KRASNOUKHOVA, Olga. *The noun phrase in the languages of South America*. Netherlands Graduate School of Linguistics, 2012.
- LABRADA, Rosés, J. 2015. *The Mako language: Vitality, grammar and classification*, Thèse de Doctorat (Ph.D.), French Studies & DDL, University of Western Ontario & Université Lumière-Lyon 2.

APÊNDICE

Marcação do SCN em diferentes classes gramaticais:

DEMONSTRATIVO

Lingua	Família	Demonstrativo
Apurinã	Aruak	sim (3 nfem vs. fem)
Baniwa	Aruak	sim (1. nfem vs. fem)
Baure	Aruak	sim (nfem vs. fem)
Chipaya	Uru-Chipaya	sim
Jarawara	Arawá	sim (msc vs. fem)
Kariri	Macro-Gê	sim (3. animado vs. inanimado)
Kubeo	Tukáno	sim (1. msc vs. fem; 2. classificadores)
Kwazá	Isolada	sim (classificadores)
Lokono	Aruak	sim (3 sg msc vs. nmsc vs. nhum)
Miraña	Bora	sim
Mosetén	Mosetén-Chimane	não
Movima	Isolada	sim (1. 3 sg msc vs. fem vs. neu)
Mundurukú	Tupí	sim
Palikur	Aruak	sim (msc vs. fem vs. neu)
Paumari	Arawá	sim (msc vs. fem)
Rikbatsa	Macro-Gê	sim (msc vs. fem)
Saliba	Piaroa-Saliban	sim (1. CM)
Sikuani	Guahiban	sim (1. msc vs. fem vs. inanimado; 2. classificador)
Siona	Tukáno	sim (anafórico)
Tariana	Aruak	sim (nfem vs. fem)
Tehuelche	Chon	sim
Wari	Chapakuran	sim (msc sg/pl vs. fem sg/pl vs. neu)
Yurakaré	Isolada	não

NUMERAL

Lingua	Família	Numeral
Apurinã	Aruak	sim (3 nfem vs. fem)
Baniwa	Aruak	sim (2. classificadores)
Baure	Aruak	sim (classificadores numerais)
Chipaya	Uru-Chipaya	?
Jarawara	Arawá	não
Kariri	Macro-Gê	sim (1. prefixos classificadores)
Kubeo	Tukáno	sim (1. msc vs. fem; 2. classificadores)
Kwazá	Isolada	sim (classificadores)
Lokono	Aruak	não

Miraña	Bora	sim (opcional em numerais acima de 2)
Mosetén	Mosetén-Chimane	sim (msc vs. fem, apenas para numeral 1)
Movima	Isolada	sim (2. classificadores)
Mundurukú	Tupí	sim
Palikur	Aruak	sim (3. numeral classifiers)
Paumari	Arawá	sim (apenas para numeral 1)
Rikbatsa	Macro-Gê	?
Saliba	Piaroa-Saliban	sim (1. CM)
Sikuani	Guahiban	sim (1. msc vs. fem vs. inanimado; 2. classificador)
Siona	Tukáno	sim (anafórico)
Tariana	Aruak	sim (classificador)
Tehuelche	Chon	sim (fem vs. masc/neu)
Wari	Chapakuran	não
Yurakaré	Isolada	não

NOME

Lingua	Família	Nome
Apurinã	Aruak	sim (não em todos os nomes: 1. nfem vs. fem, 2. compostos com NCs)
Baniwa	Aruak	sim (1. nfem vs. fem, 2. classificadores)
Baure	Aruak	sim (classificadores)
Chipaya	Uru-Chipaya	sim
Jarawara	Arawá	não
Kariri	Macro-Gê	não
Kubeo	Tukáno	sim (1. msc vs. fem; 2. classificadores)
Kwazá	Isolada	sim (classificadores)
Lokono	Aruak	sim (3 sg msc vs. nmsc vs. nhum)
Miraña	Bora	sim
Mosetén	Mosetén-Chimane	?
Movima	Isolada	sim (2. classificadores)
Mundurukú	Tupí	sim
Palikur	Aruak	não
Paumari	Arawá	não
Rikbatsa	Macro-Gê	sim (msc vs. fem)
Saliba	Piaroa-Saliban	sim (1. CM)
Sikuani	Guahiban	sim (1. msc vs. fem vs. inanimado; 2. classificador)
Siona	Tukáno	sim
Tariana	Aruak	sim (nfem vs. fem; classificador)
Tehuelche	Chon	sim
Wari	Chapakuran	não
Yurakaré	Isolada	não

PRONOME

Lingua	Família	Pronome
Apurinã	Aruak	sim (3 nfem vs. fem)
Baniwa	Aruak	sim (1. nfem vs. fem)
Baure	Aruak	sim (3 sg: nfem vs. fem)
Chipaya	Uru-Chipaya	sim
Jarawara	Arawá	sim (3 sg an, 3 nsfg an, 3 in)
Kariri	Macro-Gê	não
Kubeo	Tukáno	sim (1. msc vs. fem; 2. classificadores)
Kwazá	Isolada	não
Lokono	Aruak	sim (3 sg msc vs. nmsc vs. nhum)
Miraña	Bora	sim (fem vs. msc em dual para todas as pessoas, singular apenas para 3a pessoa)
Mosetén	Mosetén-Chimane	sim (3 msc vs. fem)
Movima	Isolada	sim (1. 3 sg msc vs. fem vs. neu)
Mundurukú	Tupí	sim
Palikur	Aruak	sim (msc vs. fem vs. neu)
Paumari	Arawá	sim (msc vs. fem)
Rikbatsa	Macro-Gê	sim (msc vs. fem [neutralizado em 1PL e 2SG])
Saliba	Piaroa-Saliban	sim (2. gênero)
Sikuani	Guahiban	sim (1. msc vs. fem vs. inanimado)
Siona	Tukáno	sim
Tariana	Aruak	sim (nfem vs. fem; classificador)
Tehueleche	Chon	sim (3a pessoa)
Wari	Chapakuran	sim (3sg/pl msc vs. fem vs. neu)
Yurakaré	Isolada	não

MODIFICADOR DESCRITIVO

Lingua	Família	Modificador descritivo
Apurinã	Aruak	sim (3 nfem vs. fem; 'verbos descritivos')
Baniwa	Aruak	sim (2. classificadores)
Baure	Aruak	não
Chipaya	Uru-Chipaya	sim
Jarawara	Arawá	sim (msc vs. fem)
Kariri	Macro-Gê	sim (1. Prefixos classificadores)
Kubeo	Tukáno	sim (1. msc vs. fem; 2. classificadores)
Kwazá	Isolada	sim (classificadores)
Lokono	Aruak	sim (verbos estativos: 3 msc vs. nmsc vs. n.hum)
Miraña	Bora	sim
Mosetén	Mosetén-Chimane	sim (msc vs. fem)
Movima	Isolada	sim (2. classificadores)
Mundurukú	Tupí	sim
Palikur	Aruak	sim (verbo estativo: 3 fem vs. nfem)

Paumari	Arawá	sim (1. msc vs. fem; 2. ka- vs.∅. OBS: msc e fem não são usados quando um verbo estativo é)
Rikbatsa	Macro-Gê	sim (msc vs. fem)
Saliba	Piaroa-Saliban	sim (1. CM)
Sikuani	Guahiban	sim (1. msc vs. fem vs. inanimado; 2. classificador)
Siona	Tukáno	sim (anafórico)
Tariana	Aruak	sim (nfem vs. fem; classificador)
Tehuelche	Chon	sim
Wari	Chapakuran	sim (msc vs. fem vs. neu)
Yurakaré	Isolada	não
Wari	Chapakuran	sim (msc vs. fem vs. neu)

POSSESSIVO

Lingua	Família	Possessivo
Apurinã	Aruak	sim (Possuidor: 1. 3 nfem vs. fem)
Baniwa	Aruak	sim (possuidor: 1. nfem vs. fem; possessao predicativa 2. classificadores)
Baure	Aruak	sim (Possuidor: 1. nfem vs. fem)
Chipaya	Uru-Chipaya	sim
Jarawara	Arawá	sim (possuidor: 3 sg an, 3 nsfg an, 3 in)
Kariri	Macro-Gê	sim (2. classificadores possessivos)
Kubeo	Tukáno	sim (1. msc vs. fem; 2. classificadores)
Kwazá	Isolada	sim (possuído: classificadores)
Lokono	Aruak	sim (possessor: 3 msc vs. nmsc vs. nhum)
Miraña	Bora	sim (possuído, opcional)
Mosetén	Mosetén-Chimane	sim (possuído msc vs. fem)
Movima	Isolada	sim (possuidor: 1. 3 sg msc vs. fem vs. neu)
Mundurukú	Tupí	sim
Palikur	Aruak	sim (possuidor: msc vs. fem vs. neu; possuído: 5. classificadores possessivos)
Paumari	Arawá	sim (possuidor, mas marcas são afixadas no possuído)
Rikbatsa	Macro-Gê	sim (2. classificadores possessivos)
Saliba	Piaroa-Saliban	sim (possuidor: prefixo gênero/pessoa; possuído: CM de nomes alienáveis)
Sikuani	Guahiban	sim (possuidor: 2. classificador)
Siona	Tukáno	sim (anafórico)
Tariana	Aruak	sim (nfem vs. fem; classificador)
Tehuelche	Chon	sim (possuidor: msc vs. fem/neu)
Wari	Chapakuran	sim (possuidor: msc vs. fem vs. neu)
Yurakaré	Isolada	sim ('animal doméstico')

Línguas, Famílias e Fontes primárias de informação

Língua	Família	Fonte
Apurinã	Aruak	Facundes 2000
Baniwa	Aruak	Aikhenvald 2007
Baure	Aruak	Danielsen 2007, Krasnoukhova 2012
Chipaya	Uru-Chipaya	Cerrón palomino 2006
Jarawara	Arawá	Dixon 2004
Kariri	Macro-Gê	Rodrigues 1997
Kubeo	Tukáno	Chacon 2012
Kwazá	Isolada	Van der Voort 2004
Lokono	Aruak	Aikhenvald 2000, pet 2011
Miraña	Bora	Seifart 2005
Mosetén	Mosetén-Chimane	Sakel 2004, Krasnoukhova 2012
Movima	Isolada	Haude 2006; SAILS
Mundurukú	Tupí	Gomes 2009
Palikur	Aruak	Aikhenvald e Green 1998
Paumari	Arawá	Aikhenvald 2010, Salzer e Chapman 1998
Rikbatsa	Macro-Gê	Silva 2011
Saliba	Piaroa-Saliban	Estrada Ramirez 1996; Labrada 2015
Sikuani	Guahiban	Queixalós 1998, 2000
Siona	Tukáno	Bruil 2014
Tariana	Aruak	Aikhenvald 2000, página 384, 370
Tehuelche	Chon	Garay 1998, Krasnoukhova 2012
Wari	Chapakuran	Everett & Kern 1997
Yurakaré	Isolada	Van Gijn 2006