



ON MALE AND FEMALE SPEECH AND MORE: CATEGORICAL GENDER INDEXICALITY IN INDIGENOUS SOUTH AMERICAN LANGUAGES

F. Rose

► To cite this version:

F. Rose. ON MALE AND FEMALE SPEECH AND MORE: CATEGORICAL GENDER INDEXICALITY IN INDIGENOUS SOUTH AMERICAN LANGUAGES. *International Journal of American Linguistics*, 2015, 81 (4), pp.495-537. <10.1086/683158>. <hal-01216653>

HAL Id: hal-01216653

<https://hal.science/hal-01216653v1>

Submitted on 19 Oct 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

ON MALE AND FEMALE SPEECH AND MORE: CATEGORICAL GENDER INDEXICALITY IN INDIGENOUS SOUTH AMERICAN LANGUAGES¹

FRANÇOISE ROSE

DYNAMIQUE DU LANGAGE (CNRS AND UNIVERSITÉ DE LYON)

Abstract

Indexing the gender of the speaker or the addressee within any type of sentence is often considered as sociolinguistic variation rather than as a gender-exclusive rule. This paper presents a survey of categorical (rather than statistical) gender indexicality in grammar with the greatest number of languages to date. It also offers a data-informed typology of categorical gender indexicality based on 41 indigenous South American languages, aimed at encouraging and facilitating research on genderlects. Examples are classified according to which speech act participants have their gender indexed and in which area of the grammar (lexicon, discourse markers, phonology, morphology). The main findings are first, that categorical gender indexicality in the grammar is more frequent than thought, and second, that the weight of gender indexicality within a language is correlated to the domain: it is limited in the lexicon, more present with discourse markers, and pervasive at the phonological and morphological levels.

Keywords

gender indexicality, genderlect, grammatical gender, male/female speech, South America

Paper

1. Introduction

Some languages show grammatical variation depending on the gender of the speech act participants.² That is, the gender of the speaker and/or the addressee determines some aspect of the phonology, lexicon, or morphology of a language (Adam 1879, Frazer 1900, Haas 1944, Sapir 1963[1929], Bodine 1975). “Particular language usages are said to belong to the realms of men’s vs. women’s speech, appropriate variations in saying otherwise ‘the same thing’ indexing gender identities in the speech situation” (Silverstein 1985:223). This phenomenon is labeled gender indexicality (Levinson 1983:89, Silverstein 1985:233). Most often, it is the gender of the speaker that is indexed, and this is commonly referred to as ‘male/female speech’.

Gender indexicality functions on the pragmatic level. Like other kinds of social indexicality, “it does not matter what is being said, nor whom or what is being referred to; the indexical forms mark something about the context in which they are used.” (Silverstein 1985:233). Gender indexicality thus refers to the fact that a word can have its basic denotation and moreover point to (index) the gender of one or several speech act participants (the speaker, the addressee, or both).³ These speech act participants do not have to be involved as participants in the event or state expressed by the utterance. In the Garifuna examples in (1)–(3), the same meanings are expressed with completely distinct word forms according to the gender of the speaker, symbolized in the translations with the biological symbols ♀ and ♂.⁴

- (1) würinauga / gúñaru Garifuna (de Pury 2003)
 ‘yesterday ♀ / yesterday ♂’⁵

(2) uá / inó

‘no ♀ / no ♂’

(3) hiñaru / würi

‘woman ♀ / woman ♂’

In Garifuna, the locus of gender indexicality is the lexicon, but it can be the phonology or morphology in other languages.

Gender indexicality is quite distinct from grammatical gender, which indicates the gender of a referent. Grammatical gender is a grammatical classification of nouns visible in agreement patterns (Corbett 1999). It is generally based on the social gender of the referent when human, but the gender categories can also extend to non-humans. For instance, Garifuna shows grammatical gender besides gender indexicality, as in (4) and (5).

(4) wáirit-**i** aunli **lé** Garifuna (de Pury 2003)

big-3M dog 3M.DEICT

‘This (male) dog is big.’

(5) wáirit-**un** aunli **tó**

big-3F dog 3F.DEICT

‘This (female) dog is big.’

Grammatical gender is not visible in the noun, but modifiers such as adjectives and demonstratives agree in gender with the head noun.

Linguistic distinctions depending exclusively on the gender of the speech act participants are considered to be rare and mostly limited to the Americas (Fleming 2012:297). The largest study on the issue (Fleming 2012) lists only 17 known cases of gender indexicality in the Americas, 11 from North America and 6 from South America, and three outliers from the rest of the world.⁶ Some of the best-known examples from around the world are Basque (Alberdi 1995), Chukchee (Dunn 2000), and Japanese (Shibatani 1990).⁷ In South America more specifically, gender indexicality is also supposedly rare, even according to specialists in the area. For instance, the most recent areal volume on Amazonian languages (Aikhenvald 2012) mentions only two languages with this peculiarity. Gender indexicality is worth studying, not only because it is rare and “exotic”, but also because its relation with grammatical gender and social gender is a crucial aspect of the language-culture nexus: as Silverstein (1985) puts it, it is at the intersection of structure, usage, and ideology. Papers on gender indexicality usually favor discussion of the social correlates of gender indexicality (among others McConnell-Ginet 1988, Ochs 1992, Romaine 2003) or its possible genesis (Taylor and Hoff 1980, Dunn 2000). While very interesting, the social use of this phenomenon as well as its origins are beyond the scope of this paper.

The first aim of the present study is to survey genderlects, i.e. linguistic variation according to the gender of the speech act participants, in South America. Only a few specific studies have been undertaken on this topic, though it is often mentioned in individual grammar descriptions. One important finding of my research is that this phenomenon, though marginal, is less rare than usually thought: the survey shows that some instantiation of this phenomenon is (or was) found in at least 41 South American languages belonging to 13 different stocks and one isolate. The present paper will show how diverse these 41 languages are in terms of quality of description, geographical distribution, genetic affiliation, degree of indexicality, and locus of indexicality.

A second objective is to provide a data-informed typology of such systems, to serve as a resource to spur on more (and more accurate) accounts of gender indexicality. The paper thus builds an inventory of loci of gender indexicality (the domains of grammar in which gender of the speech act participants is indexed), and classifies languages according to the loci where they index gender. A significant finding in that respect is that the pervasiveness of gender indexicality within a language is correlated to the domain of indexicality. For instance, lexical distinctions are usually limited to a few items. In contrast, the distinction is very robust when instantiated at the phonological or morphological level.

After a general introduction to the diversity of gender indexicality in Section 2, this paper will present in Section 3 a survey of categorical gender indexicality in South America. Section 4 will classify the languages of the survey in terms of which speech act participants have their gender indexed (the speaker, the addressee, or both). Section 5 will then offer a typology of the loci of gender-of-speaker indexicality and make some generalizations on the loci and the degree of gender indexicality. Section 6 will explain some difficulties in surveying gender indexicality systems, summarize the results, and underline the major findings of the paper.

2. The parameters of gender indexicality

In this section, I categorize the great diversity of phenomena referred to as gender indexicality, with the help of four parameters:⁸

- i) gender of the speaker vs. the addressee vs. both
- ii) locus of gender indexicality
- iii) categorical vs. statistical gender indexicality
- iv) non-referential vs. referential gender indexicality.

The first parameter of diversity within gender indexicality relates to which participants of the speech situation have their gender indexed. The typology given in Table 1 has been suggested by Haas (1944).

Table 1

Type 1 gender indexicality, i.e. indexicality of the gender of the speaker, has been exemplified in (4) with examples from Garifuna. An additional example taken from outside the Americas is the Thai polite particles, *kháʔ* or *khâʔ* for women, *kh ráb* for men. Type 2 gender indexicality, i.e. indexicality of the gender of the addressee, can be exemplified by Basque (6), “where the addressee is encoded in the verb form even when it is not an argument in the sentence” when speaking to a familiar addressee (Alberdi 1995:276).⁹

(6)	diagok	/ diagon	/ dago	Basque (Alberdi 1995:276)
	3S.stay.ADD♂	/ 3S.stay.ADD♀	/ 3S.stay	
	‘he/she/it stays (male familiar / female familiar / formal addressee)’			

There is no such distinction in formal speech.

Type 3 gender indexicality is the indexicality of relational gender, that is to say of both the speaker and the addressee. This type is much rarer. It has for instance been described in Yana (Sapir 1963[1929]), where lexical items have two phonologically different forms. One is used to speak among males (like *sika·ka* ‘quail’, *yuna* ‘acorn’), the other to address females or for females to address males (*sika·k^hA* ‘quail’, *yuh* ‘acorn’). Most of the literature accounts for Type 1, which seems by far the most frequent type and is frequently labeled ‘male/female speech’.¹⁰ This term unhappily suggests a diglossic variation covering all domains of the language. The term ‘male/female speech’, as well as the less frequent term ‘genderlect’, can also ambiguously refer to statistical indexicality (see below).

A second parameter is the locus of gender indexicality, i.e. the domain of the language or the use in which it is instantiated. Günthner (1996) lists the major speaking practices that can index gender: grammar (phonology, morphology, syntax, lexicon), pitch, choice of languages or varieties, communicative styles, discourse strategies, and discourse genres. This paper is concerned only with the first domain—gender indexicality in the grammar. Section 5 classifies the data in the survey with respect to four different loci within grammar (lexicon, discourse markers, phonology, morphology).

The third parameter is whether gender indexicality is categorical or statistical. Categorical indexicality implies that a linguistic form indexes exclusively one gender (of speaker or addressee), and another form exclusively indexes the other gender.¹¹ So gender is obligatorily indexed. This was the case in the Garifuna examples (1)-(3). Statistical indexicality only implies a tendency of association of one form to a gender. A well-known case study is that of Lakoff (1973) on women's language and the way women are addressed in English. Some linguistic features are primarily used by women, such as certain nouns, adjectives, and exclamations; tag questions; rising intonation at the end of declaratives; politeness forms and excuse formulae; and correct grammar; while taboo-words and swear-words are avoided. These characteristics are statistical, in that none of them is sufficient to predict the gender of the speaker or the addressee.¹²

A common viewpoint is that gender indexicality is always statistical, and thus always pertains to the interaction of social status and discourse rather than to grammar (Ochs 1992, Trechter 1995). Other studies on gender indexicality nevertheless argue that some systems can be considered to be grammaticalized, in that they are categorical, obligatory, stable, ubiquitous, and consciously assumed as gender-based in the metalinguistic discourse of the

speakers (Silverstein 1985:234, Fleming 2012, Ribeiro 2012). This paper focuses on categorical indexicality.

A fourth parameter is whether the element that indexes gender refers to speech act participants or not. So gender indexicality is further categorized into referential and non-referential gender indexicality (for a detailed discussion, see Fleming 2012, Rose 2013a). Referential gender indexicality is found when elements that are pragmatic indexes themselves, i.e. pronouns or deictics referring to 1st or 2nd person, have distinct forms for a male or a female referent. Gender is then both indexical and referential.¹³ Non-referential gender indexicality occurs in utterances where the denotational meaning does not necessarily include the speaker and/or the addressee, as in examples (1) to (3) above. This paper covers both referential and non-referential gender indexicality.

Within this diversity of gender indexicality phenomena, this paper is specifically concerned with categorical gender indexicality as found in the grammar of indigenous South American languages. It will not be concerned with cases of statistical indexicality, nor with implementation elsewhere than in the grammar. It will cover the three types of gender indexicality and both referential and non-referential gender indexicality.

3. A survey of categorical gender indexicality in South America

3.1. Methodology of the survey

The first aim of the survey is to give as comprehensive a list as possible of languages from South American families exhibiting categorical gender indexicality. Thus, it includes extinct languages and languages which no longer index gender, without implying a diachronic perspective. The criteria for including a language in the survey is that it was said to index gender in the grammar in a categorical manner, by at least one author, at one point in time.

When possible, such claims were verified by consulting additional sources or specialists. Languages were excluded from the survey when the variation was presented as statistical ("women use more frequently...") or due to an additional factor more robust than gender ("some women use X, they all come from Y region"). Therefore, the indexicality systems of the survey are based solely on social gender, except for data from two languages. In Chipaya (Cerrón-Palomino 2006), variants refer to gender, age, and intimacy (see Section 4.2). In Kadiwéu (Sandaló 2011), variants refer to the dialect of noble women, that of noble men, and that of non-noble men (see Section 5.3). I then constructed a database, indicating for each language its location (country and geographical coordinates), genetic affiliation, source of information and whether or not the source specifically focused on gender indexicality, type of gender indexicality (1-3), and locus of indexicality. A condensed version of the database is presented in the Appendix 1.

The existing surveys of gender indexicality in languages of the Americas (Fleming 2012) and languages of Brazil (Borges 2004) pointed me to nine languages: Chiquitano, also called Bésiro, Island Carib, Garifuna (Island Carib's offspring), Kamaíurá, Karajá, Kayabi, Kokama, Pirahã, and Xavante.¹⁴ These languages are all indigenous to South America, except Island Carib (previously spoken in Dominica) and Garifuna (spoken in Belize and Honduras), but these are related to the Arawak family from South America, and so are included in my survey. Starting with these and more specific surveys (Fabre 2004, Costa and Oliveira 2011), the list of South American languages showing gender indexicality grew larger after investigating as many languages as possible through posting on discussion lists,¹⁵ reading areal studies (Adelaar 2004, Aikhenvald 2012), and perusing grammars. Moreover, many colleagues, including specialists on sub-areas or linguistic families, were consulted.

Having arrived upon a list of languages thought to be relevant to my survey, I then proceeded to examine gender indexicality in each. More or less detailed sources specifically focusing on gender indexicality have been published for only nine languages: Awetí, Chiquitano, also called Bésiro, Iatê, Island Carib, Garifuna, Kadiwéu, Karajá, Kokama, and Mojeño. These sources are given in bold in the Appendix 1. For other languages in my survey, information was gleaned from a variety of sources. A major methodological problem in constructing the list of languages is that gender indexicality is often treated very superficially in grammars, with little discussion on the use of the systems and on possible paths of development. Gender indexicality may be ‘invisible’, with no mention in the table of contents or the index. This makes it very difficult to find data on gender indexicality in grammars. Moreover, some languages have only grammatical sketches; such data were used cautiously, though with less suspicion when they conformed with existing generalizations.

A second methodological problem is distinguishing gender indexicality from lexical or grammatical gender. This will be discussed in detail in the sections on the relevant loci of indexicality: the lexicon (5.1) and the morphology (5.4).

Once the data were collected, the elements indexing gender in each language were analyzed in terms of the loci of indexicality. I have distinguished four loci—lexicon, discourse markers, phonology, and morphology—based on the range and type of data found in the sample. A similar classification of loci was used by Fleming (2012) in his survey of 17 languages of the Americas, and it seems to be cross-linguistically valid, according to the first results of a world-wide survey on categorical genderlects (Rose and Bakker 2014). Each locus groups together a number of languages, and generalizations are possible for each locus, as summarized in the conclusion.

3.2.Distribution of the indigenous South American languages with gender indexicality

The full list of the 41 South American languages showing gender indexicality is given with relevant information in Appendix 1. This number shows that the phenomenon is not as rare as previously thought, though it nevertheless remains quite marginal with respect to the overall number of more than 400 languages in South America.

The 41 languages in the survey belong to 13 different (major and minor) families: Arawak (4 languages), Barbacoan (1), Carib (4), Guaycuru (4), Hibito-Cholonan (1), Lule-Vilela (1), Macro-Jê (8), Nadahup (1), Nambikwara (1), Tacana (1), Tucano (1), Tupi (11), and Uru-Chipaya (1). Also included is the language isolate, Pumé.

A genderlect distinction has never been reconstructed for any of the above-mentioned families. In fact, gender indexicality tends not to function within the same family in comparable ways. For example, within the Jê branch of the Macro-Jê family, Karajá indexes the gender of the speaker in the phonology, and four other Jê languages do so in discourse markers (Costa and Oliveira 2011).¹⁶ However, the four Carib languages under study all show the same gender distinction in some sort of discourse particle, and the Tupi languages mark gender distinctions in their interjections (see 5.2). The issue of whether gender indexicality can be inherited still remains to be investigated.

The 41 languages of the survey are mainly found in Brazil and Bolivia, though this survey also includes languages from Argentina, Paraguay, Peru, Colombia, Ecuador, Venezuela, Surinam, Guyana, French Guiana, Belize, Honduras, and Dominica.

Map 1

As seen in Map 1, there is a concentration of languages with gender indexicality in the Amazon basin. Within this area, two zones are particularly rich in gender-indexicality: lowland Bolivia and the Upper Xingu region. This is probably only due to the fact that they are areas with a high density of languages. Beside Amazonia, the study includes five languages of the Chaco: Abipone, Kadiwéu, Lule, Mocoví, and Toba. In contrast, only one language from the Andes is known to have gender indexicality (Chipaya) and none (to my knowledge) from the Southern Cone.¹⁷ A few languages spoken outside of Amazonia still belong to Amazonian families, for example, the reportedly Macro-Jê language Chiquitano (or Bésiro) spoken in southern Bolivia, Bolivian and Paraguayan Guaraní from the Amazonian Tupi family, and the two mixed languages of Arawak origin—Island Carib, formerly spoken in the Western Indies, and Garifuna, spoken in Belize, Honduras, and in the diaspora in the U.S.A.

Dunn (2014) suggests that “there are hints that high levels of gender variation in language may be an areal feature e.g. in Amazonia.” Many Amazonian languages share some features across linguistic families (Derbyshire and Pullum 1986, Dixon and Aikhenvald 1999), and an Amazonian “type” has been sketched (Aikhenvald 2007). There is, nevertheless, no strong support for Amazonia as a linguistic area, and some support for a larger area comprising adjacent areas (Payne 1990, Constenla Umaña 1991). Also, it has been suggested that the Chaco is a linguistic area (Comrie et al. 2010). Calculating the areal frequency of gender indexicality is uncertain, because it crucially depends on the quality and the availability of the descriptions. Gender indexicality could well be a more frequent feature in South America, and particularly in Amazonia and the Chaco, than elsewhere in the world. While the present survey lists 41 languages with genderlects out of more than 400 languages from South America, a world-wide survey in construction revealed to date only 57 cases in the rest of the world, out of more than 6000 languages (Rose and Bakker 2014). This is clearly statistically

anomalous, and likely not to be due only to the more advanced stage of the South American survey than that of the world-wide survey. Gender indexicality in Amazonia and the Chaco region, nevertheless, remains marginal and is not characteristic of the region as a whole. Additionally, there is no strong case in the literature for gender indexicality having been diffused.¹⁸ And importantly, as mentioned above, some sporadic cases are attested elsewhere in the world.

4. Types of categorical gender indexicality in South America

As seen in Table 2, it is quite clear that the prevalent pattern in South America is that of the ‘male/female speech distinction’; among the 41 languages of the survey, 37 languages index the gender of the speaker (Type 1). Only three languages index the gender of the addressee (Type 2). Only five languages index relational gender (Type 3), though four of these languages (in parentheses in the table) also exhibit Type 1 phenomena, indexing relational gender only in a few items.

Table 2

Type 1 phenomena are well-attested in the survey, and a thorough discussion of gender-of-speaker indexicality will be presented in Section 5. The remainder of this section focuses on Type 2 and Type 3 gender indexicality, for which fewer generalizations are possible due to the scarcity of data.

4.1.Type 2: gender of the addressee

The survey uncovered only three languages that index the gender of the addressee: Southern Nambikwara, Pumé, and Cholón. Southern Nambikwara is an Amazonian language spoken in Brazil.¹⁹ Pumé and Cholón belong to the Andean sphere, away from the core of the Amazonian basin where most genderlects of South America are found. In the three languages, the gender of the addressee is indexed in the morphology.

Bodine (1975:140) has claimed that no language differentiates only the gender of the addressee, and not that of the speaker, except in direct address (2nd person pronouns or vocatives). Cholón supports this claim. It shows a referential gender indexicality system, indexicalizing gender in the second person and vocative markers only. Southern Nambikwara and Pumé contradict this claim. They show a non-referential gender indexicality system: Southern Nambikwara indexicalize gender in the aspectual morphology and Pumé in the person and mode morphology.

The Peruvian language Cholón is the only known example of Type 2 referential gender indexicality in South America.²⁰ Gender indexicality in Cholón is referential, because gender is restricted to second person pronominals (7) and vocative markers (8).

(7) **mi**–kt^sok / **pi**–kt^sok Cholón (Alexander-Bakkerus 2005:131, 309)

2SG.ADD♂–box / 2SG.ADD♀–box

‘your box (male addressee) / your box (female addressee)’

(8) inča–m–**ma** / inča–m–**pa**

what–QM–ADD♂ / what–QM–ADD♀

‘Hey man / woman, what is the matter?’

In Southern Nambikwara (Kroeker 2001, Fabre (in prep. [2004])), four verb-final suffixes express both the gender of the addressee (male or female) and aspect (perfective or imperfective):²¹

- (9) wxã³–na¹–tu¹–**wa**² Southern Nambikwara (Kroeker 2001)
 come–1SG.IO–FUT–**IMPF.ADD**♂
 ‘I will come (male addressee).’

- (10) wxã³–na¹–tu¹–**?a**²
 come–1SG.IO–FUT–**IMPF.ADD**♀
 ‘I will come (female addressee).’

- (11) yxau²–na³–**la**²
 stay–EQUATIONAL–**PF.ADD**♂
 ‘He is here (male addressee).’

- (12) yxau²–na³–**na**²
 stay–EQUATIONAL–**PF.ADD**♀
 ‘He is here (female addressee).’

Pumé, previously called Yaruro, is a language isolate of Venezuela (García 2000). The gender of the addressee is indexed in more than a hundred forms expressing the person of the subject, the object, the possessor and the mode (Mosonyi 1966)²². Examples (13) and (14) illustrate indexical gender of the addressee: the same sentence is expressed differently

(13) iãĩ' jĩĩ' jará nĩ' Pumé (García 2000)
 woman the drink 3F.SG.ADD♀
 'The woman drinks (female addressee).'

Besides these three clear cases of Type 2 gender indexicality, a further case from Island Carib remains disputed. It is generally considered to depend on the gender of the speaker (Taylor 1956), though Hoff (1994) refers to the gender of the addressee. To my knowledge, no one has suggested that it could be a relational gender indexicality system. The language is extinct and recent work depends mainly on the data and explanation offered by Breton (1999[1665]), which are incomplete and sometimes contradictory. This may reflect an early situation where Island Carib was already showing Type 1 gender indexicality, and male speakers were also using a Carib pidgin among them (Taylor and Hoff 1980). Island Carib is classified under Type 1 in the survey.

Relational gender indexicality was found primarily in Chipaya, and as a secondary type after Type 1 (gender of the speaker) in Tembé, Tapirapé, Tupinambá, and Kubeo. In all these languages, it is indexed in discourse markers: in a declarative discourse particle in Chipaya,

and in phatic particles (described as vocatives) in Kubeo, Temb  , and Tupinamb  . Phatic particles, though classified here as discourse markers (see Section 5.2), are strictly speaking lexical elements. This qualifies Dunn (in press)’s suggestion that gender indexicality in Type 3 systems is often expressed morphologically.

Chipaya is the only language in the survey that shows only Type 3 gender indexicality. Based on studies on an earlier stage of the language (Olson 1966; 1967), Cerr  n-Palomino (2006) shows that relational gender is still presently indexed in the declarative discourse particle, cliticized either on the subject or the verb under its scope. This is illustrated in examples (15) to (18).

(15) zhup oqh–u–**tra** Chipaya (Cerr  n-Palomino 2006:166-167)

firewood go–1SG–DECL.  ;   →  

‘I am going for firewood (female or male speaker, male addressee).’

(16) zhup oqh–u–**ma**

firewood go–1SG–DECL.  →  (intimate)

‘I am going for firewood (male speaker, female addressee, intimate).’

(17) zhup oqh–u–**?a**

firewood go–1SG–DECL.  →  

‘I am going for firewood (female speaker, female addressee).’

(18) zhup oqh–u–**qa**

firewood go–1SG–DECL.**children**→  

‘I am going for firewood (child speaker, male addressee).’

The four-form system is organized according to the gender of both the speaker and the addressee, with children under 15 years old forming a separate gender and one of the relations (male speaker to female addressee) being apparently restricted to intimate use between a man and his wife. It seems from the description that children and intimate females are social genders.²⁴

In Kubeo and at least three Tupi languages (Tembé, Tapirapé,²⁵ and Tupinambá), some discourse markers index the gender of the speaker (Type 1), but a few other discourse markers index relational gender (Type 3). For instance, among the Tupinambá discourse particles, some are used exclusively or predominantly by speakers of one gender, illustrating Type 1: the common word for ‘no’ is *aan* or *aan-i*, while *aan-i reĩ* is used exclusively by men and *aan-i reá* by women. Other Tupinambá discourse particles are used only from man to man, illustrating Type 3: *hẽ!* ‘hi!’, used among men (Barbosa 1956:45-46).

A general remark can be made about Type 3 gender indexicality, based on South America languages as well as other well-known cases like Yana and Biloxi. Four forms would be expected to constitute a symmetric system of relational gender indexicality systems, because two binary parameters are combined—the gender of the speaker and that of the addressee. Remarkably, no language is known to conform to this model. Bodine (1975:142) noted that no language with Type 3 gender indexicality differentiates the four possible configurations, and South American data reinforces this claim. Yana shows only two forms (one among males, and the other for all other situations), as exemplified in Section 2. Biloxi (Haas 1944) has three forms (one used among males, a second one used among females, and a third one

used by either a male or a female speaking to a female). Although Chipaya has four forms, they are arranged in an asymmetric system involving parameters other than gender per se.

5. Loci of gender-of-speaker indexicality

Gender indexicality can also be classified according to the locus of indexicality, i.e. the domain where gender is indexed. The survey distinguishes four different loci: lexicon (5.1), discourse markers (5.2), phonology (5.3), and morphology (5.4).²⁶ Discourse markers (including interjections) have been set apart from the lexicon and the morphology, because they are a particularly common locus and many languages of the survey show gender indexicality in this locus only. The relative importance of each type of locus will be discussed in Section (5.5). This section deals with the loci of Type 1 gender indexicality only and is therefore based on a sample of 37 languages indexing the gender of the speaker.

5.1. Lexicon

The survey shows that lexical distinctions encoding the gender of the speaker are found in seven languages (Table 3).

Table 3

Besides the remarkable case of Island Carib discussed below, lexical distinctions generally concern only a few items in a language. For example, only one item shows gender-based lexical distinction in Xavante:

- | | | |
|------|---------------------|--------------------------------|
| (19) | tiha / marĩ | Xavante (Machado Estevam 2011) |
| | ‘thing ♀ / thing ♂’ | |

Lexical distinctions are shown for one item in Xerente and Guarayo (Megan Crowhurst, p.c.), a few items in Aweti (Drude 2002) and Siriono, up to 26 items in Kadiwéu (Souza 2012), and around 50 items in Garifuna (de Pury 2003). Lexical distinctions are used in every day speech and are semantically and formally unpredictable. They can be restricted to a sub-category of words like animal and plant names in Siriono (Schermair 1957, Noé Gasparini, p.c.).

Island Carib is unusual in that the gender-indexing distinctions exist for many lexical items. De Goeje (1939) has analyzed 2809 words from Breton's (1999[1665]) work. Of the 2547 words that are not borrowed from either Spanish or French (these are not specified for gender of the speaker), 1610 words (63 %) are common to both sexes, 554 (22 %) are used only by men, and 383 (15 %) only by women. These numbers show that over a third of this large lexical sample indexes the gender of the speaker. Out of a list of 100 basic lexical items, Taylor (1956) counts that 41 items of vocabulary are common to male and female speakers (30 of Arawak origin and 11 of Carib origin) and 59 items are expressed through different words by females and males, those of females being predominantly of Arawak origin, and those of males of Carib origin.²⁷ Island Carib is a mixed language, with many elements from Carib having entered an Arawak language after some Carib men came to live among an Arawak group (maybe reduced to women), according to the prevailing theory (Taylor and Hoff 1980). The male/female suppletive forms can be explained by different sources of origin.

Dunn (2014) suggests that there are two types of gender-indexing lexical distinctions: "These may be cryptic variants of the same word, where men's and women's forms of the lexeme are clearly related but have some distinctive mutation. [...] There are also cases where men's and women's lexemes have no obvious etymological relationship." The latter case was exemplified above with Xavante (19) and is obviously the case in Island Carib, in which the corresponding lexemes have not only different etymological sources, but also different

language sources. The former case can be exemplified with Kadiwéu and Awetí. In the list of Kadiwéu lexical pairs, I have detected frequent (though not completely regular) correspondences, such as the substitution by female speakers of a V_1CV_2 sequence by a V_2V_2 sequence. Specialists cannot point to the specific factor triggering this rule in a restricted subset of the lexicon (20)-(22), nor the phonological rules explaining other lexical pairs (23)-(24).

(20) *atemati / eemati* Kadiwéu (Souza 2012)

‘tell ♂ / tell ♀’

(21) *nopitena / niitena*

‘arrow ♂ / arrow ♀’

(22) *akami / aami*

‘you ♂ / you ♀’

(23) *eloadi / eemadi*

‘kill ♂ / kill ♀’

(24) *acipe / aaka*

‘drink ♂ / drink ♀’

The results (a small number of languages with a small number of gender-indexing items) may seem to undermine common expectations about the importance of categorically distinct lexicon between men and women. In fact, I have collected many answers from researchers

convinced that some languages index the gender of the speaker in the lexicon, especially in the kinship system. Therefore, two methodological caveats are called for at this point. First of all, some lexical items may include, in their semantics, information on the gender of a participant of the state of affairs. Thus, some verbs can refer to activities typical of men or women. This situation is regularly confused with gender indexicality, although it is logically independent from it. A nice illustrative example can be found in two successive analyses of some Karajá lexical distinctions. An earlier study of gender indexicality in Karajá (Fortune and Fortune 1975) points to a lexical distinction based on the gender of the speaker, illustrated in (25). Confusingly, further explanation given by the authors seems to indicate that these two lexical items express two different activities: “The women cry or chant daily if a child is travelling or hurt. Women also cry for a period of a lunar month: the death chant. In contrast, men cry only during the death chant and even then less than the women” (Fortune and Fortune 1975).

(25) bu / hi Karajá (Fortune and Fortune 1975)
 ‘cry ♀ / cry ♂’

A more recent study (Ribeiro 2012) clarifies the distinction, which actually refers to two different kinds of activities typical of a female or male agent, but which is in no way linked to the gender of the speaker, see (26) and (27). Both terms can be used by speakers of both genders. The gender of the speaker should not be confused with the gender of a participant of the state of affairs expressed by the lexicon.

(26) obu Karajá (Ribeiro 2012:141)
 ‘cry (feminine subject) ♀ / ♂’

(27) hi

‘cry (masculine subject) ♀ / ♂’

A second type of frequent confusion is that between the gender of the speaker and the gender of the ego of a kinship term. *Ego* is the person on which a kinship relationship is based, or in morphosyntactic terms, the ‘possessor’ of the relationship. This can be illustrated with the following examples from Emerillon.

(28) e-r-adʒir

Emerillon (Rose, fieldnotes)

1SG-REL²⁸-daughter

‘my daughter ♂’

(29) e-mēbir

1SG-daughter

‘my daughter ♀’

(30) t-adʒir

3²⁹-daughter

‘his daughter ♀ / ♂’

(31) i-mēbir

3-daughter

‘her daughter ♀ / ♂’

Examples (28) and (29) both refer to a feminine referent and thus seem to be distinguished according to the gender of the speaker. However, when the same roots are used with a 3rd person prefix as in (30) and (31), it appears that the factor for the distribution of the two lexical roots is not the gender of the speaker, but rather the gender of the possessor (the ego of the kinship relation). Consequently, *adʒir* should be glossed ‘daughter of a man’, and *mēbir* ‘daughter of a woman’.

This distinction between ‘daughter of a man’ and ‘daughter of a woman’ parallels that of ‘husband’ and ‘wife’ in English. All speakers can use both these words even though ‘husband’ with a 1st person possessor will generally be uttered by a female speaker, and conversely. As is recommended in the Questionnaire (Appendix 2), it is crucial when researching gender indexicality not to elicit kinship terms only with a 1st person possessor (or only in the vocative form, where a first person possessor is implicitly involved), because it is impossible with such forms to know whether the gender distinction concerns the speaker or the possessor, since they both refer to the same person.

This confusion was already mentioned in Fleming (2012:307). Nevertheless, Fleming considers that Bésiro (a language of Bolivia and Brazil previously known as Chiquitano) is one of the rare languages to encode the gender of the speaker in kinship terms, as was asserted by Adam and Henry (1880:6): “the man says *izai* ‘my father’ [...], and the woman says *ixupu* ‘my father’”. Please note that these terms are given with a 1st person possessor.³⁰ In fact, as example (32) shows, both male and female speakers can use the root *upu* that actually means ‘father of a female’ (as well as both can use the root *ai* ‘father of a male’).

- | | | | |
|------|---|-----------------|--------------------------|
| (32) | n–i–y–upú | n–i–po–sápa | Bésiro (Sans fieldnotes) |
| | N–3–LC–father | N–3–house–SIMIL | |
| | ‘the father of the (female) neighbor ♂ / ♀’ | | |

It can be concluded that so far no South American language is known to distinguish kinship terms on the basis of the gender of the speaker.³¹ Gender indexicality in the lexicon is found in other semantic sub-parts of the lexicon, but shows up as a rather rare and minor phenomenon in South American languages, as summarized in Table 3. This fits with Fleming's classification of languages with gender indexicality (16 American languages and 3 outliers), with only Island Carib indexing gender in the lexicon (Fleming 2012:300). Remarkably, it is the only locus of categorical gender indexicality in just one language of the survey, Siriono.³²

To summarize, lexical distinctions is a locus of gender indexicality in seven languages only. It generally concerns a few items a language, except in Island Carib. The two variants may be completely independent or formally related. Contrarily to a common idea, there is absolutely no attestation of kinship terms indexing gender.

5.2. Discourse markers

This section is concerned with gender indexicality in various types of discourse markers, including interjections, affirmative and negative words or particles, discourse particles (with illocutionary functions), routines and formulae, and connectives. Discourse markers are defined by their function of encoding the speaker's communicative intentions (Fraser 1996) and by their invariable form. They are treated separately from the lexicon or morphology because they are a very common locus of gender indexicality, often the only locus in the language (see Table 4).³³

Table 4

Interjections are “relatively conventionalized vocal gestures (or more generally, linguistic gestures) that express a speaker’s mental state, action or attitude, or reaction to a situation.” (Ameka 1992:106). In the survey, interjections are the most important sub-class within discourse markers in terms of number of items and number of languages. This is in line with the assumption by several authors that gender indexicality in interjections is cross-linguistically common (Balmori 1962:44, Shibatani 1990:371, Ribeiro 2012:141). Various South American examples of gender-indexing interjections are given in Table 5.

Table 5

Of the different types of interjections noted by Ameka (1992), expressive and phatic interjections are the ones most relevant to gender indexicality in South American languages. Expressive interjections are symptoms of the speaker’s mental state. In the survey, they express pain, surprise, disdain, aversion, admiration, sadness, anger, shock, joy, fright, shame or derision. They are found in Abipone, Bolivian Guarani, Old Guarani, Guarayo, Kali’na, Karajá, Kipea, Lule, Mojeño, Tapirapé, and Tupinambá. Expressive interjections indexing the gender of the speaker in the Tupi-Guarani branch of the Tupi family are also common and seem to be cognate. Phatic interjections are those used in the establishment and maintenance of communicative contact. Such interjections, often called ‘vocatives’, are found in Bolivian Guarani, Karajá, Mocoví and Tupinambá. There are also some rare examples of a conative interjection (i.e. an interjection directed at an addressee) in Kipea, Pará Gavião and Xerente. Interestingly, the survey does not show gender indexicality in descriptive interjections, even though ideophones are a common category in Amazonian languages.

A second sub-set of discourse markers is often referred to as ‘discourse particles’ in studies on Amazonian languages. They do not necessarily serve as discourse connectives but generally convey illocutionary force, with some possible additional epistemic meanings.³⁴ They are frequent in discourse and express emotion, illocutionary, or epistemic meanings that are difficult to define precisely. Typologically, they are distinguished from interjections by their inability to function on their own. Some examples are given in Table 6.

Table 6

Discourse particles with general illocutionary force are found in three Carib languages (Avarico,³⁵ Kali’na, and Tamanaku) and in Kubeo. Gilij (1780-1784:161) says that even if these three Carib languages do not have grammatical gender, they nevertheless have particles distinguishing the gender of the speakers, as shown in (33) and (34).

(33) maje **uè** Tamanaku (Gilij 1780-1784:161)
‘Let’s go! ♂’

(34) maje **me** Tamanaku (Gilij 1780-1784:161)
‘Let’s go! ♀’

Gender-indexing particles with additional meaning (like surprise in (35)) are found in Kali’na, Kama iurá, and Tembé (see Table 6).

(35) h–ajme–ma’e te ’aŋ **pa** Kama iurá (Seki 2000:100)
3–be_sharp–NOM FOC PROX DISC(surprise)♂
‘Wow, how sharp-edged this is! ♂’

Other types of discourse marker indexing the gender of the speaker are affirmative/negative words or particles, as found in Abipone, Bolivian Guarani, Garifuna, Guarayo, Mebengokre, Pará Gavião, Sateré-Mawé, Tembé, Toba, Tupinambá, and Xavante, routines or formulae, as found in Iatê, and connectives, as found in Iatê, Kokama and Omagua. See Table 6 for examples.

To conclude, gender indexicality in discourse markers is quite common, but often restricted to only a few items, usually less than a dozen. Therefore, it is a rather minor phenomenon within these languages. The question of whether these double sets of discourse markers can be reconstructed for Proto-Carib and Proto-Tupi should be further investigated.

5.3. Phonology

Four South American languages have been described as indexing gender in their phonology (see Table 7), and for two of them, it is the only locus of gender indexicality.

Table 7

Two languages, Araona and Guarayo, have been described as indexing the gender of the speaker in the phonetic substitution of a single phoneme. Araona has been said to have a male-female speech distinction: [f] in female speech corresponds to [s] in male speech (Aikhenvald and Dixon 1999:366).³⁶ However, this phonetic substitution is not attested in a recent phonological description, where [s] is the only phonetic realization of /s/ for speakers of both genders (Emkow 2006). In Guarayo, men used to pronounce [ts] or [ds], where

women pronounced [s] (Höller 1932:2). An alternation between [ts] for men and [s] for women has been noted in the 1990's (Megan Crowhurst p.c.).

- (36) otso ~ odso / oso Guarayo (Höller 1932:2)³⁷
- ‘(s)he went away ♂ / (s)he went away ♀’

The best described case of gender indexicality on the phonological level is from Karajá. Karajá does not show phonetic substitution but phonological deletion. In most contexts, where women pronounce /k/, men delete it (Ribeiro 2012:131).

- (37) kɔwɔɾu / ɔwɔɾu Karajá (Ribeiro 2012:131)
 ‘tree, wood ♀ / tree, wood ♂’

- (38) hãlɔkɔɛ / hãlɔɛ
 ‘jaguar ♀ / jaguar ♂’

Deletion of a phoneme can have consequences beyond sound substitution. The male variant in example (38) shows fusion of the two vowels that are separated by /k/ in the female variant. These further changes perhaps make it less likely for the variants to be leveled.³⁸ The fact that this gender distinction was already described more than a century ago (Ehrenreich 1894) shows the stability of the system.

Finally, Kadiwéu shows a prosodic distinction between men and women. According to Sandalo (2011), non-noble men would use binary moraic feet, as in (39); noble men ternary

geographic locations. Eight of these eleven languages have other loci of gender-indexing besides morphology.

Table 8

These 11 languages all index the gender of the speaker in the pronominal/reference system, i.e. in independent or bound pronominals and/or in the gender system. Most of the present section is therefore devoted to describing gender indexicality in the pronominal/reference system. Fleming (2012) had already noticed a robust type of language with gender-indexicality in grammatical categories that refer to humans (3rd person pronominals, determiners, human nouns, gender, etc.). In my survey, three languages additionally index the gender of the speaker in other parts of the morphology. They will be discussed at the end of the section.

Kokama shows indexicality of the gender of the speaker in its pronominal system, in the 1st and 3rd person (Vallejos 2010), as illustrated in Table 9. The gender distinction in the 3rd person pronouns does not concern the gender of the referent but that of the speaker. There is no grammatical gender in this system.

Table 9

Table 10 summarizes which persons in the paradigms of independent or bound pronominals index gender in the 14 systems under study. Three groups of languages can be observed. In the first six languages, gender is indexed in third person pronominals (and in some languages in first person pronominals as well), clearly distinguishing these systems from systems with grammatical gender.⁴¹ The second group of languages, Garifuna, Island Carib, and Kadiwéu, index the gender of the speaker in 1st and 2nd person pronouns only, and not in third person pronouns. Importantly, when the 2nd person pronoun indexes the gender of

the speaker, there is no difficulty in classifying these languages as non-referential Type 1 gender indexicality, i.e. as languages that index the gender of the speaker in forms other than these referring to the speaker (1st person pronominals). The third type of gender indexicality in pronouns is illustrated by Tsafiki (Dickinson 2002:65). It indexes the gender of the speaker only of the 1st person. It can be said to illustrate referential gender indexicality—the element that indexes the gender of a speech act participant also refers to this person as a participant in the event or state. Referential gender indexicality is not well attested in my survey; it occurs in only 2 languages out of 41. Besides the Tsafiki Type 1 case of gender of speaker indexicality just mentioned, referential gender indexicality is also found in Cholón Type 2 gender of addressee indexicality (Cf. Section 4.1).

Table 10

These results are not surprising given that gender distinctions in the 1st and 2nd person are cross-linguistically rare. Siewierska (2011) counted 124 languages with some gender distinction in the independent personal pronouns in a sample of 378 geographically and genetically diverse languages.⁴² Out of these, only 20 languages show a gender distinction in the 1st and/or 2nd person, and two of these do not distinguish gender in the 3rd person. Thus, not only is a gender distinction in the 1st or 2nd person rare, but it is even rarer to find it without a distinction in the 3rd person.

In the process of collecting data on gender indexicals, I have come across seven languages that show grammatical gender in the 1st or 2nd person. Nasa Yuwe shows gender in 1st and 2nd persons only. Itonama and Old Mocoví have gender in the 2nd and also in the 3rd person. Iatê, Kubeo, Ofaié-Xavánte, and Rikbaktsa show gender in the 1st and 2nd persons and also in the 3rd person. I have excluded these seven languages from the survey, since the gender distribution makes an analysis in terms of gender indexicality uselessly complex. For

instance, in Iatê (Table 11), all three persons show a gender distinction that seems to be morphologically analyzable (as a suffix $-sV$ with vowel harmony), and thus it is more economical to consider a single distinction of grammatical gender (i.e. gender of the referent) rather than three different distinctions (indexicality of the gender of the speaker for 1st person, indexicality of the gender of the addressee for 2nd person, and grammatical gender for 3rd person).

Table 11

Unambiguous cases of gender indexicality may interact with grammatical gender, forming reference systems that are extremely complex to describe. Mojeño shows an interesting pronominal paradigm (visible in independent pronouns, verbal and nominal prefixes and suffixes, articles, and demonstratives). For third-person human singular referents, feminine and masculine are distinguished, but the form for masculine differs depending on the gender of the speaker (Rose 2013*b*; 2015).

- (42) **esu** n–kuñar–eni Mojeño (Rose, fieldnotes)
- PRO.F 1SG–sister_in_law–PAST
- ‘She was my sister in law. ♂’
- (43) **esu** s–omuire movimari'i
- PRO.F 3F–too movima–ASS
- ‘She also is a Movima. ♀’
- (44) **eñi** t–yono te to ewire–gne
- PRO.M♀ 3–go PREP.NH ART.NH away–very
- ‘He goes very far away. ♀’

- (45) **ema** ma-no-k-po to senda
 PRO.M♂ 3M♂-make-ACT-PF ART.NH path
 ‘He made the path. ♂’

All speakers use the same feminine independent pronoun *esu*, as in (42) and (43), but female and male speakers use different forms for the masculine independent pronoun, respectively *eñi* (44) and *ema* (45). Thus, in Mojeño, the gender of the speaker is indexed in the pronominal forms referring to a particular grammatical gender.

Also in Bésiro (better known as Chiquitano), gender indexicality is found in the marking of the gender distinction. Female speakers make no gender distinctions, whereas male speakers use additional morphology for masculine only. Previously, this language has been described as showing gender indexicality at the phonological level, with the deletion of the initial phoneme of some nouns referring to humans (or animals), as in (46) and (47) (Adam and Henry 1880, Schuchard 1986:98-99).

- (46) ixaaras / xaras Bésiro (Adam and Henry 1880)
 ‘Spaniard ♂ / Spaniard ♀’

- (47) ñoñeís / oñeís
 ‘man ♂ / man ♀’

Pierric Sans (p.c.) argues that these pairs should in fact be interpreted in terms of morphological alternation. The female forms are the plain roots. The male forms are made up of the root and a masculine prefix *i-*.⁴³ Indeed, comparison with first-hand data shows that all

the examples given by Adam and Henry are in fact masculine nouns. And indeed, Bésiro encodes masculine gender only in male speech, as illustrated below with more recent data. Male speakers, but not female speakers, encode masculine gender by adding a masculine clitic (= *ti* in the singular and = *ma* in the plural) to refer to a third person male subject or possessor (48)-(50) (Sans 2013).⁴⁴

(48)	ba-páche-ro= ti	n-i-kisé-s	Bésiro (Sans fieldnotes)
	3-look_for-TAM=3SG.M♂	N-3-knife-DET	
	‘He looks for her knife. ♂’		

(49)	ba-páche-ro	n-i-kise-s= ti
	3-look_for-TAM	N-3-knife-DET=3SG.M♂
	‘She looks for his knife. ♂’	

(50)	ba-páche-ro= ti	n-i-kise-s= ti
	3-look_for-TAM=3SG.M♂	N-3-knife-DET=3SG.M♂
	‘He looks for his knife. ♂’	

In contrast, female speakers never encode grammatical gender (51).

(51)	ba-páche-ro	n-i-kisé-s
	3-look_for-TAM	N-3-knife-DET
	‘She looks for her knife. / She looks for his knife. / He looks for her knife. / He looks for his knife. ♀’	
	‘She looks for her knife. ♂’	

Male speech has no marker for feminine gender, so male speakers use the same unmarked forms as female speakers (51), but these refer only to a female subject or possessor.

Island Carib and notably Garifuna also show a complex system of gender indexicality in the expression of grammatical gender. The attribution of grammatical gender differs depending on the gender of the speaker. The languages show a masculine/feminine distinction. Yet in Garifuna, males use feminine agreement (52) and females use masculine agreement (53) with abstract nouns, as well as in some specific syntactic constructions (de Pury 2003, Munro 2013).

- (52) Gúndan–tina t–au idemual t–ó. Garifuna (de Pury 2003:159)
 happy–1SG 3F–with help 3F–DEICT
 ‘I am happy with this help. ♂’

- (53) Gúndan–tina l–au idemual l–é.
 happy–1SG 3M–with help 3M–DEICT
 ‘I am happy with this help. ♀’

Thus, the gender of the speaker is indexed by the use of a specific grammatical gender for some naturally non-gendered items. This extremely complex situation is clearly exceptional in the survey.

Although all languages indexing gender at the morphological level show some distinction within the reference system, it was noted earlier (see Table 8) that four languages (Iatê, Island Carib, Kokama, Omagua) additionally index gender in other parts of the morphology.

(54) =kana / =nu Kokama (Vallejos 2010:42)

=PL♂ / =PL♀

(56) yaepe / raepe
there, then ♂ / there, then ♀

(57) ta samake-hlě-ne Iatê (Costa and Silva 2005:25)
3SG marry-PF-IND ♀

‘(s)he got married ♀’

Finally, Hoff (1994) mentions some Island Carib morphemes that are used only in male speech, and only on lexemes of Carib origin. Most were very likely borrowed from Carib. Taylor (1954: 29) also lists five morphemes that are mentioned in Breton (1999[1665]) as belonging only to men’s speech: a negator, a causative and three pronominal prefixes. For Island Carib, Taylor shows that men used the negator *–pa* of Carib origin, while women used the Arawak negative prefix *m–* (58).

(58) arámêtoupátina / marámêtontina Island Carib (Taylor 1954:29)
 ‘I am not hiding (s.th) ♂ / I am not hiding (s.th) ♀’

It seems these were remains of Carib morphology in the mixed language. What is clear for Island Carib is that morphology distinctions in male and female speech were quite marginal given the amount of distinct lexicon.

To summarize, gender-indexing morphology was found in 11 languages, in the pronominal or reference system of each, and in other morphology for four languages. Indexical gender and grammatical gender interact in complex ways in some of the languages of the survey.

5.5. The loci and the degree of gender indexicality

The previous four sections discussed the various loci where gender-of-speaker indexicality occurs in the languages of my sample. Table 12 shows the number of languages with gender indexing for each locus.

Table 12

What stands out from these results is the importance of the discourse markers as a locus of gender indexing. Also, languages with gender-indexing lexicon are not as prevalent as suggested in the literature.

Another perspective on the issue is to consider the degree of gender-of-speaker indexicality that one finds in any given language. The degree of indexicality can be evaluated on the basis of a combination of three parameters: number of loci, number of items in each locus, and the frequency of the items.

As far as the number of loci is concerned, 26 of the 37 languages in the sample index the gender of the speaker in one locus only, most often in discourse markers. Six languages out of 37 index gender in two loci and 5 languages in three loci. Appendix 1 lists the loci for each language.

Generally, the number of items indexing gender per locus is low (under a dozen). The number of gender-indexing items is minor in both the grammar and the lexicon. The one exception to these generalizations is Island Carib, where a great part of the lexicon indexes the gender of the speaker. Even though it is the best known case of the sample, it is clearly not representative of the sample in this respect.

The frequency of gender-indexing items is difficult to discuss without text-based research. One observation is that the frequency of the gender-indexing items seems to vary depending on the locus. The literature does not comment on the frequency of the gender-indexing lexical items, and the examples provided do not suggest them to be neither particularly rare nor frequent. I assume that the degree of indexicality is therefore somewhat medium in languages indexing gender only in this locus. Gender-indexing discourse markers vary in terms of frequency—for example, an assertive marker is likely to be much more frequent than an

interjection expressing pain. No counts of either lexical or discourse-marking elements are available.

In contrast, we can surmise that gender-indexicality in the phonology results in a robust genderlect distinction. For example, Karajá discourse is said to index gender phonologically every three words on average (Fortune and Fortune 1975). Gender-indexicality in the morphology is also highly visible in the languages of the survey because it is found almost exclusively in pronouns or other referential items, and pronominal cross-reference on the predicate is frequent in Amazonian languages (Dixon and Aikhenvald 1999). For example, counts on a Mojeño Trinitario text (Rose 2013*b*) shows that gender-indexing items are very common (in a majority of sentences on average), though they are irregularly distributed, with an average of 2.5 per sentence when present.⁴⁵ This is illustrated in the following extract, with three elements (a pronoun, an article and a prefix) expressing masculine singular, and at the same time, indexing the gender of the male speaker.

- (59) **ema**–rich'o **ma** papa Piyo **mu**–ejare⁴⁶ Mojeño (Rose, fieldnotes)
PRO3M♂–still **ART.M♂** pope Pio **3M♂**–name
‘He was still the pope, Pius was his name. ♂’

In sum, phonological and morphological loci are particularly salient distinctions in that they tend to pervasively gender indexicality.

The languages in the survey are quite diverse in terms of the degree of indexicality. Most languages of the survey have a low degree of indexicality—this probably explains why gender-indexicality in particular languages is sometimes presented as anecdotal or is described only sketchily. Among the languages with a high degree of gender indexicality,

there are many more languages with a few frequent gender-indexing items (such as Karajá in the phonology and Mojeño or Bésiro in the morphology) than languages with a high number of gender-indexing items (such as Island Carib and its lexicon).

6. Conclusion

This paper has surveyed the phenomenon of gender indexicality in South America, limiting the discussion to cases of the categorical indexation of the gender of one or both speech act participants in clauses that do not necessarily refer to them as participants in the state of affairs expressed. A questionnaire was used in conducting the survey (see Appendix 2), and it may prove useful for further research on individual languages. Based on the survey of 41 languages (see list in Appendix 1), the paper offers a typology of gender indexicality. To my knowledge, this is the largest published survey of gender indexicality systems to date. Some possible reasons for categorical gender indexicality to be under-described are suggested below.

First of all, gender indexicality can be very unobtrusive in a language. It affects only two words in Xavante, and often only interjections in Tupi-Guarani languages. Interjections are often treated only briefly or not at all in grammars (Schachter and Shopen 2007[1985]:57). The minor scope of gender indexicality within grammars, especially in contrast with the exceptional pervasiveness of gender indexicality in the lexicon of Island Carib, the best known case of gender indexicality in the region, may have led scholars to underestimate the interest of their data.

Second, a careful methodology is required for researching gender indexicality. Data obviously have to be collected with speakers and addressees of both genders. The lexicon should ideally be systematically checked for its use by both genders of speakers. Data should

also consist of spontaneous discourse to order to include interjections and discourse particles and to illustrate the discourse use of gender-indexing items.

A third difficulty could be that speakers may refuse or not be able to discuss gender indexicality. This possibility seems not to hold strongly. Only one of the investigators consulted has noted that there could be a kind of taboo with the issue.⁴⁷ All detailed descriptions mention that not only can speakers of both genders evaluate a word as belonging to one lect or another, but they can also produce both lects with ease and correct children. My own experience with Mojeño speakers indicate a total consciousness of the gender indexicality system, and knowledge of both variants by all speakers (Rose 2013b). The survey shows that the norm is indeed that all speakers master the two variants. This is obvious from the discourse use of speaker gender indexicality: a speaker uses the variant of the other gender when quoting the speech of a speaker of the other gender.⁴⁸ In the following textual example, the female Mojeño speaker Florencia Carire Tamo uses the female speech singular masculine *ñi* when talking about a man with a dog-face. She uses the masculine speech singular masculine *ma* when quoting this man's speech about another man (for Kokama examples, cf. Vallejos (2010:43)).

(60) **ñi**-cho'o-uko-pu-iji eno **ñi**-janemuri-ono Mojeño (Rose 2013b:122-123)

3M_♀-call-PL-PF-RPT 3PL **3M**_♀-fellow-PL

“j-**ma**-kñ-ojoo'i kristianu ”

DEM-**3M**_♂-INVIS-COP human_being

‘He called to his fellows _♀: “There’s a human being around _♂.”’

Given these difficulties, the survey of 41 languages reported on here serves as a preliminary treatment. Hopefully this paper will inspire research on gender indexicality in

other languages. The remainder of this conclusion first summarizes the results of the paper at all levels (geographically, genetically, pragmatically, typologically, degree of indexicality, use), and then underlines the two major findings of the paper.

The discussion of the geographical and genetic distribution of the languages of the survey is still preliminary, due to insufficient evidence. Geographically, gender indexicality is found in many parts of South America (see Map 1), with most occurrences in Amazonia and the Chaco region. At present, there is no robust evidence of gender-indexicality having spread by contact. Genetically, the phenomenon is found in 13 linguistic stocks, with very little evidence that it is the result of inheritance. Some regularities were observed, such as the recurring pattern of gender-indexing discourse markers in Tupi-Guarani and Carib languages, and gender indexicality or grammatical gender in all persons in several Macro-Jê languages (Iatê, Ofayê, Rikbaktsa). Whether these shared patterns should be seen as a common inheritance still remains to be investigated.

Pragmatically, all three types of gender indexicality (of the speaker, of the addressee, and relational) have been observed, with a notable predominance of indexicality of the gender of the speaker (37/41).

The systems were further classified according to the loci of indexicality, with four major types (lexicon, discourse markers, phonology, morphology). Discourse markers (including interjections) have been set apart from the lexicon and the morphology, because they are a particularly common locus, often the only locus for individual languages. Lexical distinctions are far less common than expected on the basis of the existing literature, and are not found in kinship terms, contrary to common expectations. Some regularities were observed within the loci, such as the predominance of pronominal forms within morphology, and of expressive interjections and discourse particles with illocutionary force within discourse markers. Only

two languages were considered to show referential gender indexicality (in the 1st or 2nd person only); all others index gender in elements that do not specifically refer to the speech act participants.

A great diversity was observed in the degree of indexicality, from languages like Hup with just one gender-indexing interjection to rare languages like Island Carib where the greater part of the lexicon indexes the gender of the speaker. On the whole, gender indexicality remains a generally marginal phenomenon, far from diglossia (except for Island Carib), contrary to what the expression ‘male vs. female speech’ suggests.

While descriptions of the phenomenon vary in depth, from a brief mention in a grammar to specific papers on the topic (on 9 languages only, sources in bold in Appendix 1), almost all descriptions note a similar use of the gender indexicality system: all speakers know the two variants, are conscious of the system, and use the variant of the opposite gender when citing a speaker of the opposite gender.

Finally, on the social level, the present survey uncovered minor uses of a genderlect by someone of the opposite biological sex. For instance, in Bolivian Guarani, children are socially ungendered.⁴⁹ Boys change their speech at the initiation ritual when they are fitted with the labret (Giannecchini 1996:189-306). Prestigious elder Karajá men who withdraw from most public activities and stay within the women’s sphere can use female speech without being criticized (Ribeiro 2012:154-155).

The first major finding of the survey is that categorical gender-indexicality is much more common than previously thought, at least in South America (compare the 41 languages of my survey with the 6 languages in Fleming’s 2012 study). Besides the famous reports of how gender is statistically indexed in discourse, many languages also obligatorily index gender in

various loci of their grammar. This survey thus provides a large and rather robust dataset disproving the position that gender indexicality is always statistical, i.e. that there is only a tendency of association of one form to a gender. It supports the idea that social distinctions are encoded not only in discourse, an area which has already received a lot of attention, but also in grammar (for another example, see Evans 2003). It also suggests that this phenomenon may be more common in South America than in the rest of the Americas, where it has previously been more frequently discussed. Hopefully, this paper will be an incentive for the discovery of new systems of gender indexicality and a better description of the systems already uncovered.

The second major result of the survey is the line drawn between occasional gender indexicality, for example, as found in the lexicon or discourse markers, and regular gender indexicality, for example, occurring in the phonology or the morphology. Although gender indexicality in the phonology or the morphology is rare cross-linguistically, it is found in almost half of the languages in my sample (18/41). Furthermore, it is well-attested within the languages that exhibit it because of the high frequency of its use. So not only is gender-indexicality not so rare as previously thought, it is also much more pervasive than previously thought.

Appendix 1. Languages and data sources

Type of gender indexicality is symbolized as I : gender of speaker, II : gender of addressee, III: relational gender. Locus of marking is abbreviated as DISC: discourse markers, LEX: lexicon, M: pronominal morphology, M(+) : pronominal morphology and other type of morphology, PH: phonology. Sources of data in bold are studies that specifically deal with gender indexicality.

Language	Language family	Type	Locus	Sources of data
Abipone	Guaycuru	I	DISC	Dobrizhoffer (1822: vol 2, p. 197), Najlis (1966:73) ⁵⁰
Araona	Tacana	I	PH	Aikhenvald and Dixon (1999:366)
Avarico	Carib	I	DISC	Gilij (1780-1784:161)
Awetí	Tupi	I	M, LEX	Drude (2002)
Bésiro	Macro-Jê	I	M	Adam and Henry (1880), Falkinger (2002) , Sans (2013, fieldnotes)
Bolivian Guarani	Tupi	I	DISC	Giannechini (1996[1898]), Dietrich (1986:168), Ortiz and Caurey (2012)
Chipaya	Uru-Chipaya	III	DISC	Olson (1967:300-304), Cerrón-Palomino (2006:150-151, 165)
Cholón	Hibito- Cholonan	II	M, DISC	Alexander-Bakkerus (2005)
Garifuna	Arawak	I	M, LEX, DISC	De Pury (2003; 2004), Munro (2013)
Guarayo	Tupi	I	PH, LEX, DISC	Höller (1932), Crowhurst (2000)
Hup	Nadahup	I	DISC	Epps (2008)
Iatê	Macro-Jê	I	M(+), LEX, DISC	Costa and Silva (2005)
Island Carib	Arawak	I (II?)	M (+), LEX	Taylor (1954; 1956), Taylor and

				Hoff (1980), De Pury (2003; 2004), Hoff (1994)
Kadiwéu/Mbaya	Guaycuru	I	PH, M, LEX	Sandalo (2011), Souza (2012), Sanchez Labrador (Sánchez Labrador 1971a[1760], 1971b[1760])
Kali'na	Carib	I	DISC	Gilij (1780-1784:161), DeGoeje (1946[1910]:43-44), Odile Renault-Lescure (p.c.)
Kamaiurá	Tupi	I	DISC	Seki (2000:100-101)
Karajá	Macro-Jê	I	PH, DISC	Fortune and Fortune (1975), Ribeiro (2001; 2012: Chapter 3)
Kayabí	Tupi	I	M	Dobson (1997:13-14)
Kipeá	Macro-Jê	I	DISC	Mamiani (1699:116-117)
Kokama	Tupi	I	M(+), DISC	Vallejos (2010), Pottier (1972), Faust (1959)
Kubeo	Tucano	I	DISC (+III)	Chacon (2012)
Lule	Lule-Vilela	I	DISC	Machoni (1732:87-88), Zamponi and de Reuse (in preparation)
Maipure	Arawak	I	DISC	Gilij (1780-1784:161), Zamponi (2003:11)
Mebengokre	Macro-Jê	I	DISC	Salanova (2001), Bernat Bardagil p.c.
Mocoví	Guaycuru	I	DISC	Paucke (2010:286), Bucca

				(1981:233) ⁵¹
Mojeño	Arawak	I	M, DISC	Marbán (1702), Olza Zubiri et al. (2002), Rose (2013b)
Old Guarani	Tupi	I	DISC	Ruiz de Montoya (1640)
Omagua	Tupi	I	M(+), DISC	Zachary O'Hagan (2011:55, p.c.)
Pumé	isolate	II	M	García (2000:567)
Pará Gavião	Macro-Jê	I	DISC	Costa and Oliveira (2011)
Sateré-Mawé	Tupi	I	DISC	DaSilva (2010:206)
Siriono	Tupi	I	LEX	Schermair (1949; 1957)
Southern Nambikwara	Nambikwara	II	M(+)	Kroeker (2001:65-66)
Tamanaku	Carib	I	DISC	Gilij (1780-1784:161)
Tapirapé	Tupi	I	DISC	Walkiria Praça (p.c.)
		(+III)		
Tembé	Tupi	I	DISC	Bendor-Samuel (1972)
		(+III)		
Toba	Guaycuru	I	DISC	Susnik (1971)
Tsafiki	Barbacoan	I	M	Dickinson (2002:65)
Tupinambá	Tupi	I	DISC	Barbosa (1956)
		(+III)		
Xavante	Macro-Jê	I	LEX, DISC	Machado (2011)
Xerente	Macro-Jê	I	LEX, DISC	Sousa Filho (2007:97-98)

Appendix 2. Questionnaire on gender indexicality

This questionnaire is made for linguists seeking to detect and describe gender indexicality. It was designed for a survey of gender indexicality in South American languages. Please address questions, suggestions, or results to the author of the questionnaire.

‘Gender indexicality’ refers to a formal distinction depending on the gender of the speech act participants, and is distinct from grammatical gender, which indicates the gender of a referent. Gender indexicality can be found in sentences where neither the speaker nor the addressee is involved as a participant in the event.

1. name of language
2. linguistic family
3. geographical location
4. sources (published or fieldnotes)
5. researcher contact information

Type of indexicality

6. Does language X index the gender of the speaker (a.k.a. male/female speech), the addressee or both? If both, please describe the number of categories and their context of use.
7. Is the distinction categorical (forms exclusively used by a gender) or statistical (forms statistically more frequently associated with a gender)?

Locus of marking

8. Is gender indexed in the lexicon? If so, in how many items? In which semantic subpart of the lexicon (i.e. in animal names)?

NB: Please make sure that the lexical distinction is not based on gender of the referent for nouns (as in *son/daughter*), on gender of an argument of the verb (for verbs expressing, say, typical activities of women *vs.* men), on gender of the possessor for kinship terms (DO NOT use 1st person possessor or vocative forms when eliciting kinship terms). In all these cases, there is no gender indexicality.

9. Is gender indexed in interjections (expressive, phatic, descriptive –ideophones-, conative)? Examples of expressive interjections include expression of pain, surprise, disdain, aversion, admiration, sadness, anger, shock, joy, fright, shame and derision.
10. Is gender indexed in discourse particles?
11. Is gender indexed in affirmative/negative words or particles?
12. Is gender indexed in routines and formulae?
13. Is gender indexed in connectives?
14. Is gender indexed in the phonology? If so, how (phonetic substitution, deletion, syllabic structure, prosodic rules)?
15. Is gender indexed in the morphology? If so, is it found in the pronominal system? Is it found elsewhere in the morphology as in the pronominal system?

NB: If found only in the 1st or 2nd person, consider analyzing it as grammatical gender especially if grammatical gender is found in the 3rd person.

Synchronic comparison

16. Are the two forms distinguishing indexical gender (lexemes, morphemes, phonemes, etc.) equivalent in terms of complexity?
17. Can one of the forms be considered to be derived from the other? By what type of operation (addition, deletion, etc.)?

18. Is the genderlect distinction based on the presence vs. absence of a feature/element/category?

Diachrony

19. Is the same gender indexicality pattern found in other related languages?
20. Are both forms inherited, i.e. reflexes of attested or reconstructed proto-forms?
21. Can one form (or the marked form) be considered an internal development of the other form?
22. Is the same distinction found in a genetically-unrelated neighboring language?
23. Is one of the form a result of borrowing?

Use

24. Are the lects given a special name?
25. Is gender indexicality obligatory or optional?
26. In case of indexicality of gender of the speaker, check in texts whether the opposite gender form is used in citing a speaker of the opposite gender?
27. How is social gender defined? (classification of children, elders and homosexuals)
28. How pervasive is the phenomenon (minor, visible in all sentences...)? Please specify whether this results from the number of elements concerned (types) and/or their frequency (tokens)?
29. Approximate age of acquisition.
30. Attitudes towards (exclusive or statistical) distinctions related to gender of the speech act participants.
31. Attitudes towards errors and correcting.

References

- ADAM, LUCIEN. 1879. Du parler des hommes et du parler des femmes dans la langue caraïbe. *Mémoires de l'Académie de Stanislas (Nancy)* 4ème série, no.Tome XI: 145-176.
- ADAM, LUCIEN AND HENRY, VICTOR. 1880. *Arte y vocabulario de la lengua Chiquita*. Paris: Maisonneuve et Cie.
- ADELAAR, WILLEM F.K. 2004. *The languages of the Andes*. Cambridge Language Surveys. Cambridge: Cambridge University Press.
- AIKHENVALD, ALEXANDRA Y. 2007. Languages of the Pacific Coast of South America. *The Vanishing Languages of the Pacific Rim*, ed. O. Miyaoka, O. Sakiyama and M. Krauss, 183-205. Oxford: Oxford University Press.
- AIKHENVALD, ALEXANDRA Y. 2012. *The Languages of the Amazon*. New York, NY: Oxford University Press.
- AIKHENVALD, ALEXANDRA Y. AND DIXON, R.M.W. 1999. Other small families and isolates. *The Amazonian languages*, ed. R.M.W. Dixon and Alexandra Y. Aikhenvald, 341-383. Cambridge: Cambridge University Press.
- ALBERDI, JABIER. 1995. The Development of the Basque System of Terms of Address and the Allocutive Conjugation. *Towards a History of the Basque Language*, ed. J.I. Hualde, J.A. Lakarra and R.L. Trask, 279-293. Amsterdam: John Benjamins.
- ALEXANDER-BAKKERUS, ASTRID. 2005. *Eighteenth-Century Cholón*. Utrecht: LOT.
- AMEKA, FELIX 1992. Interjections: The universal yet neglected part of speech. *Journal of Pragmatics* 18, no.2-3: 101-118.
- ANTONOV, ANTON. to appear. Verbal allocutivity in a cross-linguistic perspective. *Linguistic Typology*.
- BAKKER, PETER. 2013. "Genderlects", Course given at the Historical Sociolinguistics Summer School, Lesbos.

- BALMORI, CLEMENTE HERNANDO. 1962. Habla mujeril. Filología VIII.
- BARBOSA, A. LEMOS. 1956. Curso de Tupi Antigo: Gramática, Exercícios, Textos. Rio de Janeiro: Livraria São José.
- BENDOR-SAMUEL, DAVID. 1972. Hierarchical Structures in Guajajara. Norman, Oklahoma: SIL.
- BODINE, ANN. 1975. Sex differentiation in language. Language and sex: differentiation and difference, ed. Barrie Thorne and Nancy Henley, 130-151. Rowley, Massachusetts: Newbury House.
- BORGES, MÔNICA VELOSO. 2004. Diferenças entre as falas feminina e masculina no Karajá e em outras línguas brasileiras : aspectos tipológicos. Llamas 4: 103-113.
- BRETON, PÈRE RAYMOND. 1999 [1665]. Dictionnaire caraïbe-français du Révérend Père Raymond Breton. Paris: Editions IRD/Karthala.
- BUCCA, SALVADOR. 1981. Palabras y frases mocovíes de Colonia Dolores. Cuadernos del Sur 14: 231-238.
- CERRÓN-PALOMINO, RODOLFO. 2006. El chipaya o la lengua de los hombres del agua. Lima Fondo Editorial PUCP.
- CHACON, THIAGO. 2012. The Phonology and Morphology of Kubeo: The documentation Theory and Description of an Amazonian Language, PhD Dissertation, University of Hawai'i, Manoa.
- COMRIE, BERNARD; GOLLUSCIO, LUCÍA; GONZÁLEZ, HEBE AND VIDAL, ALEJANDRA. 2010. El Chaco como área lingüística. Estudios de lenguas amerindias 2: contribuciones al estudio de las lenguas originarias de América, ed. Zarina Estrada Fernández and Ramón Arzápalo Marín, 85–131. Hermosillo: Editorial Unison.
- CONSTENLA UMAÑA, ADOLFO. 1991. Las lenguas del área intermedia : introducción a su estudio areal. San José, Costa Rica: Editorial de la Universidad de Costa Rica.

- CORBETT, GREVILLE. 1999. *Gender*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.
- COSTA, JANUACELE DA AND SILVA, FÁBIA PEREIRA DA. 2005. Dêixis de Gênero em Yaathe, língua indígena brasileira (Macro-jê). *Leitura* 35: 15-29.
- COSTA, RODRIANA DIAS COELHO AND OLIVEIRA, CHRISTIANE CUNHA DE. 2011. A distinção entre as falas masculina e feminina em algumas línguas da família Jê, *Proceedings of the Registros da 63a. Reunião Anual da SBPC*: 1-5.
- CROWHURST, MEGAN. 2000. Informe sobre algunas diferencias lingüísticas que existen entre el guarayu y el guaraní, ms.
- DA SILVA, RAYNICE GERALDINE PEREIRA 2010. Estudo morfossintático da língua sateré-mawé, PhD, Programa de Pós-Graduação em Linguística, Unicamp, Campinas.
- DAHL, ÖSTEN. to appear. Siriono. *Lenguas de Bolivia*, ed. Mily Crevels and Peter Muysken. La Paz: Plural Editores.
- DE GOEJE, CHARLIUS HENRICUS. 1939. Nouvel examen des langues des Antilles. *Journal de la Société des Américanistes* 31: 1-120.
- DE GOEJE, CHARLIUS HENRICUS. 1946 [1910]. *Études linguistiques caribes*. Amsterdam: KNAW.
- DE PURY, SYBILLE. 2003. "Vice-versa". *Le genre en garifuna*. *Faits de Langues* 21, no.2: 155-162.
- DE PURY, SYBILLE. 2004. El género en garífuna. Un análisis dinámico, *Proceedings of the Symposium Dinámica lingüística de las lenguas en contacto*, 51º Congrès International des Américanistes.
- DERBYSHIRE, DESMOND AND PULLUM, GEOFFREY. 1986. Introduction. *Handbook of Amazonian languages*, ed. Desmond Derbyshire and Geoffrey Pullum, Vol. I, 1-28. Berlin: Mouton de Gruyter.

- DICKINSON, CONNIE. 2002. Complex Predicates in Tsafiki, PhD dissertation, University of Oregon, Eugene.
- DIETRICH, WOLF. 1986. El idioma chiriguano: gramática, textos, vocabulario. Madrid: ICI.
- DIXON, R.M.W. AND AIKHENVALD, ALEXANDRA Y. 1999. Introduction. The Amazonian languages, ed. R.M.W. Dixon and Alexandra Y. Aikhenvald, 1-21. Cambridge: Cambridge University Press.
- DOBRIZHOFFER, MARTIN. 1822. An account of the Abipones, an equestrian people of Paraguay. London: John Murray.
- DOBSON, ROSE. 1997. Gramática Prática Com Exercícios da Língua Kayabi. Sociedade Internacional de Linguística. Cuiabá.
- DRUDE, SEBASTIAN. 2002. Fala masculina e feminina em Awetí. Línguas Indígenas Brasileiras. Fonologia, Gramática e História. Atas do I Encontro Internacional do GTLI da ANPOLL, ed. Ana Suelly Cabral and Aryon Rodrigues, 177-190. Belem: Editoria Universitária U.F.P.A.
- DUNN, MICHAEL. 2000. Chukchi Women's Language: A Historical-Comparative Perspective. *Anthropological Linguistics* 42, no.3: 305-328.
- DUNN, MICHAEL. 2014. Gender determined dialect variation. The expression of gender, ed. Greville Corbett, 39-68. Berlin: De Gruyter.
- EHRENREICH, PAUL. 1894. Materialien zur Sprachenkunde Brasiliens I. Die Sprache der Caraya (Goyaz). *Zeitschrift für Ethnologie* 26: 20-37, 49-60.
- EMKOW, CAROLA. 2006. A Grammar of Araona, an Amazonian language of Northwestern Bolivia, PhD Dissertation, Reserach Center for Linguistic Typology, LaTrobe University, Melbourne.
- EPPS, PATIENCE. 2008. A Grammar of Hup. Mouton Grammar library 43. Berlin / New York: Mouton de Gruyter.

- EVANS, NICHOLAS. 2003. Context, culture, and structuration in the languages of Australia. *Annual Review of Anthropology* 32: 13-40.
- EVERETT, DANIEL. 1979. Aspectos da fonologia do Pirahã, Dissertação de Mestrado em Lingüística, Unicamp, Campinas.
- FABRE, ALAIN. 2004. Gender in language families of the Gran Chaco (Paraguay and Argentina) and surrounding areas: A comparative and typological study, ms.
- FALKINGER, SIEGLINDE. 2002. Diferencias entre el lenguaje de hombres u mujeres en Chiquitano (Besiro). *Current studies on South American Languages*, ed. M. Crevels, S. van de Kerke, S. Meira and H. van der Voort, 43-56. Leiden: CNWS Publications.
- FAUST, NORMA WILLE FAUST. 1959. El lengua de los hombres y mujeres en Cocama. Peru: Instituto Linguistico de Verano.
- FLEMING, LUKE. 2012. Gender indexicality in the Native Americas: Contributions to the typology of social indexicality. *Language in Society* 41, no.3: 295–320.
- FORTUNE, DAVID AND FORTUNE, GRETCHEN. 1975. Karajá men's-women's speech differences with social correlates. *Arquivos de Anatomia e Antropologia* 1: 111-124.
- FRASER, BRUCE. 1996. Pragmatic markers. *Pragmatics* 6, no.2: 167 -190.
- FRAZER, JAMES G. 1900. A suggestion as to the origin of gender in language. *The Fortnightly Review* 73: 79–90.
- GARCÍA, JORGE RAMÓN. 2000. Yaruro (Pumé). *Manual de lenguas indígenas de Venezuela*, ed. Esteban Emilio Mosonyi and Jorge Carlos Mosonyi, Vol. I, 544-593: Fundación Bigott.
- GIANNECCHINI, DOROTEO. 1996 [1898]. *Historia natural, etnografía, geografía, lingüística del Chaco boliviano*. Tarija: Centro Eclesial de Documentación.

- GILIJ, FILIPPO SALVATORE. 1780-1784. Saggio di Storia Americana, o sia, Storia Naturale, Civile a Sacra de'Regni e delle Provinzia Spagnuole di Terra-Ferma nell'America Meridionale/4. Rome: L. perego erede Salvioni.
- GÜNTNER, SUSANNE 1996 Male-Female Speaking Practices across Cultures. *Contrastive Socio-linguistics*, ed. Marlis Hellinger and Ulrich Amm, 447-473. Berlin: Mouton de Gruyter.
- HAAS, MARY. 1944. Men's and Women's Speech in Koasati. *Language* 20, no.3: 142-149.
- HOFF, BEREND. 1994. Island Carib, an Arawakan language which incorporated a Lexical register of Cariban origin, used to address men. *Mixed Languages: Fifteen case studies in language intertwining*, ed. Peter Bakker and Maarten Mous, 161–168. Amsterdam: Institute for Functional Research into Language and Language Use (IFOTT).
- HÖLLER, ALFREDO. 1932. Grammatik der Guarayo-Sprache. Guarayos, Bolivia.
- HUNT, RICHARD. 1937. Mataco-English and English-Mataco Dictionary, Vol. 5. *Etnologiska studier*. Göteborg.
- KROEKER, MENNO. 2001. A Descriptive Grammar of Nambikuara. *International Journal of American Linguistics* 67, no.1: 1-87.
- LAKOFF, ROBIN. 1973. Language and woman's place. *Language in Society* 2, no.1: 45-79.
- LEVINSON, STEPHEN C. 1983. *Pragmatics*. Cambridge textbooks in linguistics. Cambridge Cambridgeshire ; New York: Cambridge University Press.
- LI, PAUL JEN-KUEI. 1983. Types of lexical derivation of men's speech in Mayrinax. *Bulletin of the Institute of History and Philology, Academia Sinica* 54, no.3: 1-18.
- MACHADO ESTEVAM, ADRIANA. 2011. *Morphosyntaxe du Xavante*, PhD Dissertation, Sciences Du Language, Université Paris Diderot (Paris 7).
- MACHONI DE CERDEÑA, ANTONIO 1732. *Arte, y vocabulario de la lengua lule, y tonocote*. Madrid.

- MAMIANI. 1699. *Arte de grammatica da lingua brasilica da naçam Kiriri*. Lisboa: Miguel Deslandes.
- MARBÁN, PEDRO. 1702. *Arte de la lengua Moxa, con su Vocabulario, y Cathecismo*. Lima: Imprenta Real de Joseph de Contreras.
- MCCONNELL-GINET, SALLY. 1988. Language and gender. *Linguistics: The Cambridge survey*, ed. Frederick J. Newmeyer, Vol. 4: Language: The socio-cultural context, 75–99. New York: Cambridge University Press.
- MOSONYI, ESTEBAN E. 1966. *Morfología del verbo yaruro. Estudio de los sufijos personales*. Caracas: Universidad Central de Venezuela.
- MUNRO, PAMELA. 2013. “Garifuna Gender Revisited”, SSILA, Boston.
- NAJLIS, ELENA L. 1966. *Lengua Abipona*, Vol. 2. Buenos Aires: Centro de Estudios Lingüísticos, Facultad de Filosofía y Letras, Universidad de Buenos Aires.
- O'HAGAN, ZACHARY J. 2011. *Proto-Omagua-Kokama: Grammatical Sketch and Prehistory*, Honor's Thesis, University of California, Berkeley.
- OAKLEY, ANN. 1972. *Sex, Gender and Society*. London: Temple Smith.
- OCHS, ELINOR. 1992. Indexing gender. *Rethinking context: language as an interactive phenomenon*, ed. Alessandro Duranti and Charles Goodwin, 335-358. Cambridge: Cambridge University Press.
- OLSON, RONALD D. 1966. Morphological and syntactical structures of Chipaya, ms.
- OLSON, RONALD D. 1967. The Syllable in Chipaya. *International Journal of American Linguistics* 33, no.4: 300-304.
- OLZA ZUBIRI, JESÚS; NUNIDE CHAPI, C. AND TUBE, J. 2002. *Gramática Moja Ignaciana*. Caracas: Universidad Católica Andres Bello.
- ORTIZ, ELIO AND CAUREY, ELÍAS. 2012. *Diccionario etimológico y etnográfico de la lengua guaraní hablada en Bolivia*. La Paz: Plural Editores.

- PAUCKE, FLORIAN. 2010. *Hacia allá y para acá (memorias)*. Santa Fe: Ministerio de Innovacion y Cultura de Santa Fe.
- PAYNE, DORIS. 1990. Morphological Characteristics of Lowland South American languages. *Amazonian Linguistics, Studies in Lowland South American Languages*, ed. D. Payne, 213-241. Austin: University of Texas Press.
- POTTIER, BERNARD. 1972. Langage des hommes et langage des femmes en cocama (tupi). *Langues et Techniques, Nature et Société*, ed. Jacqueline Thomas and Lucien Bernot, Vol. I, 385-387. Paris: Editions Klincksieck.
- RIBEIRO, EDUARDO. 2012. *A grammar of Karajá*, PhD Dissertation, Chicago University.
- RIBEIRO, EDUARDO RIVAIL. 2001. "Female versus male speech in Karajá", SSILA / WAIL, Santa Barbara, July 6 2001.
- ROMAINE, SUZANNE. 2003. Variation in language and gender. *Handbook of Language and Gender*, ed. Janet Holmes and Miriam Meyerhoff, 98-118. Oxford: Basil Blackwell.
- ROSE, FRANÇOISE. 2013*a*. Le genre du locuteur et de l'allocutaire dans les systèmes pronominaux: genre grammatical et indexicalité du genre. *Bulletin de la Société de Linguistique de Paris* 108, no.1: 381-417.
- ROSE, FRANÇOISE. 2013*b*. Los generolectos del mojeño. *Liames* 13: 115-134.
- ROSE, FRANÇOISE. 2015. Innovative complexity in the pronominal paradigm of Mojeño. A result of contact? *Borrowed morphology*, ed. Francesco Gardani, Peter Arkadiev and Nino Amiridze. Berlin: De Gruyter Mouton.
- ROSE, FRANÇOISE AND BAKKER, PETER. 2014. "Les génerolectes", Séminaire TypoUlm, Paris, 3 mars 2014.
- RUIZ DE MONTTOYA, ANTONIO S.J. 1640. *Arte, y vocabulario de la lengua guarani*. Madrid: Iuan Sanchez.

- SALANOVA, ANDRÉS PABLO. 2001. A nasalidade em Mebengokre e Apinayé: o limite do vozeamento soante, Mestrado, Departamento de Lingüística, Campinas.
- SÁNCHEZ LABRADOR, JOSÉ. 1971*a* [1760]. Gramática eyiguayegi - mbayá. Familia Guaycuru, ed. Branislava Susnik, Vol. I, 1-170. Asuncion del Paraguay: Museo Ethnografico Andres Barbero.
- SÁNCHEZ LABRADOR, JOSÉ. 1971*b* [1760]. Vocabulario Eyiguayegi. Familia Guaycuru, ed. Branislava Susnik, Vol. II, III. Asuncion del Paraguay: Museo Ethnografico Andres Barbero.
- SANDALO, FILOMENA. 2011. Estratificação social e dialetos prosódicos no kadiwéu. Kadiwéu: senhoras da arte, senhores da guerra, ed. Giovani J. Silva. Cuiabá: Editora CVR
- SANS, PIERRIC. 2013. Elementos de la gramática del Bésiro. Sociolingüística – Fonología – Morfología – Textos.
- SAPIR, EDWARD. 1963 [1929]. Male and female forms of speech in Yana. Selected writings of Edward Sapir in language, culture and personality, ed. David Mandelbaum. Los Angeles: University of California Press.
- SCHACHTER, PAUL AND SHOPEN, TIMOTHY. 2007 [1985]. Parts-of-speech systems. Language typology and syntactic description, ed. Timothy Shopen, Vol. I. Cambridge: Cambridge University Press.
- SCHERMAIR, ANSELMO EBNER. 1949. Gramática de la lengua sirionó. La Paz.
- SCHERMAIR, ANSELMO EBNER. 1957. Vocabulario sirionó-castellano, Vol. Sonderheft 5. Innsbruck/Áustria.
- SCHUCHARD, BARBARA. 1986. Apuntes sobre la gramática. Zúbaka. La Chiquitanía: visión antropológica de una región en desarrollo, ed. Jürgen Riester, Vol. I: Vocabulario Español-Chiquito y Chiquito-Español de Max Fuss. Cochabamba/ La Paz: Editorial Los Amigos del Libro.

- SEKI, LUCY. 2000. Gramática do Kamaíurá. Campinas: Editora da Unicamp.
- SHIBATANI, MASAYOSHI. 1990. The languages of Japan. Cambridge Language Surveys. Cambridge: Cambridge University Press.
- SIEWIERSKA, ANNA. 2011. Gender Distinctions in Independent Personal Pronouns. The World Atlas of Language Structures Online, ed. Matthew S. Dryer and Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology.
<<http://wals.info/chapter/44>>.
- SILVERSTEIN, MICHAEL. 1985. Language and the culture of gender: At the intersection of structure, usage, and ideology. Semiotic mediation: Sociocultural and psychological perspectives, ed. Elizabeth Mertz and Richard Parmentier, 219–59. Orlando: Academic Press.
- SOUSA FILHO, SINVAL MARTINS DE. 2007. Aspectos morfossintáticos da língua Akwe-Xerente, PhD Dissertation, UFG.
- SOUZA, LILIAN MOREIRA AYRES DE. 2012. Descrição da fala masculina e da fala feminina na língua Kadiwéu, Universidade Federal do Mato Grosso del Sul.
- SUSNIK, BRANISLAVA. 1971. Los patrones estructurales de la lengua Toba/ Guaycurú. Familia Guaycuru, ed. Branislava Susnik, Vol. I, 1-102. Asuncion del Paraguay: Museo Ethnografico Andres Barbero.
- TAYLOR, DOUGLAS. 1954. Diachronic note on the Carib contribution to Island Carib. International Journal of American Linguistics 20, no.1: 28-33.
- TAYLOR, DOUGLAS. 1956. On dialectal divergence in Island Carib. International Journal of American Linguistics 25, no.1: 62-68.
- TAYLOR, DOUGLAS AND HOFF, BEREND. 1980. The linguistic repertory of the Island-Carib in the seventeenth century: the men's language- A Carib pidgin ? International Journal of American Linguistics 46, no.4: 301-312.

- TRECHTER, SARA. 1995. Categorical gender myths in Native America: Gender deictics in Lakshota. *Issues in Applied Linguistics* 6, no.1: 5-22.
- VALLEJOS, ROSA. 2010. A grammar of kokama-kokamilla, PdD Dissertation, Department of Linguistics, University of Oregon.
- ZAMPONI, RAOUL. 2003. Maipure. *Languages of the world Materials*. München: Lincom Europa.
- ZAMPONI, RAOUL AND DE REUSE, WILLEM. in preparation. *Lule Corpus*, ms.

¹ Special thanks to Mily Crevels, Fernando Zúñiga, Marie-Claude Mattei-Müller, Spike Gildea, Eduardo Ribeiro, Wolf Dietrich, Antoine Guillaume, Diego Vilar, Alain Fabre, Pedro Viegas Barros, and Willem de Reuse for sharing their knowledge on the topic. Thanks also to Patience Epps, Willem de Reuse, Pierric Sans, Diana Lewis, Peter Bakker, Anton Antonov, and anonymous reviewers and editors, especially Donna Gerdt, for commenting on preliminary versions of the paper.

² Gender is a socially constructed classification of people built on the basis of the biological difference between males and females, but not completely identical with it (Oakley 1972). “The word gender [...] refers to the complex of social, cultural and psychological phenomena attached to sex, a usage common in the behavioral and social sciences” (McConnell-Ginet 1988).

³ Basque is unique in having markers that index the gender of the addressee without carrying a referential or grammatical meaning.

⁴ These symbols refer to the gender of the speaker (♀ = female, ♂ = male) or that of the addressee (ADD♀, ADD♂). Symbols referencing speaker or addressee are usefully distinct from abbreviations for grammatical gender (M = masculine, F = feminine).

⁵ Abbreviations used in this paper are: ♀ – female speaker; ♂ – male speaker; ACT – active; ADD – addressee; ART – article; ASS – assertive; CLF – classifier; COP – copula; DEICT – deictic;

DEM – demonstrative; DET – determiner; DISC – discourse marker; EXCL – exclusive; F – feminine; FOC – focus; FUT – future; IMPF – imperfective; INCL – inclusive; IND – indicative; INDEF – indefinite; INVIS – invisible; IO – individual verification, observation orientation; LC – linking consonant; M – masculine; N – *n-* prefix; NH – non-human; NOM – nominalizer; PAST – past; PF – perfective; PL – plural; PREP – preposition; PRO – pronoun; PROX – proximal; QM – question marker; REL – relational; REP – repetitive; RPT – reportative; S – intransitive subject; SG – singular; SIMIL – similitive; TAM – tense aspect mood.

⁶ The three outliers are Kūrux (Dravidian), Chuchckee (Chukotko-Kamtchakan), and Yanyuwa (Pama-Nyungan).

⁷ Additional case studies are presented in Dunn (2014) and Bakker (2013).

⁸ A fifth parameter could be markedness: can one of the variants be analyzed as equivalent to the other in terms of complexity? Or as derived from it via addition, elision, etc? Or as being restricted to a genderlect and nonexistent in the other? A sixth parameter could be the historical relation of the two co-existing linguistic forms: is one historically derived from the other? Or innovated? These questions lead to investigating the development of gender indexicality systems, and are beyond the scope of this paper (see the Questionnaire in the Appendices).

⁹ The term “allocutivity” is traditionally used for this phenomenon in Basque. A recent cross-linguistic study of allocutivity (Antonov to appear) surveys situations in which “an addressee who is not an argument of the verb is systematically encoded in all declarative main clauses conjugated verb forms.” It includes cases other than Type 2 gender indexicality, because the verbal indexing of arguments does not necessarily encode gender cross-linguistically.

¹⁰ Fr. “parler des hommes / des femmes”, Sp. “habla varonil / mujeril; habla masculino / femenino”, Port. “fala feminina / masculine”.

¹¹ In Type 3 gender-indexicality, the gender of both the speaker and addressee is indexed. The distinction is therefore not binary but still exclusive.

¹² It is in fact now generally asserted within sociolinguistic studies that statistical gender deixis cannot be separated from other social factors and must be dealt with in terms of social deixis.

¹³ A related caveat is that languages with grammatical gender may sometimes look as if they were indexing the gender of the speaker or the addressee, in sentences including 1st or 2nd person. The use of grammatical gender in agreement with 1st or 2nd person (as in French *je suis français / française* ‘I am French’) is therefore regarded as insufficient to speak of gender indexicality. The distinction *français/française* accounts for the gender of the subject in the examples given, and not for the gender of the speaker, as shown by the agreement with a 3rd person subject (*il est français / elle est française* ‘he/she is French’), whatever the gender of the speaker.

¹⁴ Pirahã has been excluded from the survey, because gender indexicality in that language has been described as being statistical. Everett (1979) states that Pirahã's allophones [ɽ] and [ɽ̃] of the phonemes /g/ and /b/ are used more frequently by female speakers and are restricted to familiar situations for male speakers.

¹⁵ <http://lista.etnolinguistica.org/3304>

¹⁶ Within these four languages, only one pair of forms seems to be shared by two languages.

¹⁷ It is also possible that Amazonia and the Chaco may have a similar proportion of languages with gender indexicality as other regions in South America, but that the overall

numbers are higher because there are simply more languages with decent descriptions. This proportion is quite difficult to estimate.

¹⁸ Rose (2015) investigates the hypothesis that the Mojeño genderlect distinction has been borrowed, or more precisely, that the forms of one of the genderlects have been borrowed, but admits the lack of strong evidence for this hypothesis.

¹⁹ Southern Nambikwara is a dialect complex. According to Kroeker (2001:1), the various bands of the Southern Nambikwara nation “all speak a dialectic variant of the same language, which is mutually intelligible with all the other variants.”

²⁰ A cross-linguistic study of gender in 1st and 2nd person pronouns notes that Iraqw (Cushitic) and Minangkabau (Austronesian) also show Type 2 referential gender indexicality (Rose 2013*a*).

²¹ Numbers indicate tones.

²² Mosonyi’s data was kindly made accessible to me by José Alvarez as a Toolbox database.

²³ Pumé also has grammatical gender: the clitics in (13) and (14) refer to the feminine subject. The interaction of grammatical gender with indexical gender is treated in Section 5.4.

²⁴ This four-form system of relational gender indexicality was simplified in the expression of the imperative, now reduced to a two-form distinction depending on the gender of the addressee (Type 2), with *-a* ‘imperative male addressee’ and *-um(a)* ‘imperative female addressee’ (Cerrón-Palomino 2006:151).

²⁵ The specific meaning of the sentence-final discourse particle of Tapirapé is not explicated.

²⁶ One difficulty encountered in classifying phenomena into these four loci are the fuzzy boundaries of the ‘discourse marker’ category. For instance, ‘connectives’ could have been

placed under the lexicon, and ‘illocutionary’ particles are functionally very similar to the mood suffixes discussed in the morphological section.

²⁷ It is actually hypothesized that the source for the Carib items in Island Carib is a pidgin used as a vehicular language along the coast of South America, and on the islands (Taylor and Hoff 1980).

²⁸ This relational prefix is found between a 1st or 2nd person prefix and a vowel-initial lexical root.

²⁹ The *t-* ~ *i-* allomorphs are distributed on a lexical basis.

³⁰ The prefix *i-* is used for both 1st and 3rd person, but it is followed by a different linking consonant in these two contexts before vowel-initial roots.

³¹ As an associate editor noted, in some groups, kinship terms are used overwhelmingly by ego (they mark endearment in address or reference), so that their use in discourse regularly index the gender of the speaker.

³² Siriono also showed a male/female alternation in the phonology (realization [d] ~ [r]) in the past, though it is not attested nowadays (Dahl to appear, Noé Gasparini, p.c.). However, since the distinction is presented as statistical (Schermaier 1957:17) rather than categorical gender indexicality, Siriono is excluded from this survey.

³³ A locus termed ‘illocutionary force operators’ was used in Fleming 2012.

³⁴ This use is contrary to the definitions given by Ameka (1992:107) for which “particles are typically used to express speaker attitudes or perspectives towards a proposition and to modify the illocutionary force of utterances”, while discourse markers or particles “indicate the way a speaker intends a subsequent discourse fragment to be related to the previous unit.”

³⁵ This is a long extinct language attested only in Gilij’s work.

³⁶ Aikhenvald and Dixon (1999:366) do not provide examples or reference source.

³⁷ This phonetic transcription is based on Höller's description, with the help of Swintha Danielsen.

³⁸ However, the differences have been almost completely leveled in the Javaé dialect of Karajá (Ribeiro 2012:155-167).

³⁹ Sandalo did not collect data for non-noble women's speech.

⁴⁰ This may be the case in Kadiwéu. This is also probably the case in the Atayal dialect of Mayrinax spoken in Taiwan. This dialect shows more than a hundred lexical pairs indexing the gender of the speaker (Li 1983). These forms may be explained by numerous unpredictable phonological or morphological derivations classified under five different types, plus suppletion.

⁴¹ Hunt (1937:14-15) suggests that the two sets of independent pronouns of Mataco (now called Wichí in Argentina), one used in polite or ordinary speech and the other in the intimacy of family life, are vestiges of an ancient system of male/female speech in all persons. In the absence of any data from this former stage, this language was excluded from the database.

⁴² See Rose 2013a for a typology of systems with gender in 1st and 2nd person pronouns in the languages of the world.

⁴³ In (47), the prefix *i-* turns into a glide before a V, and then undergoes a regular process of nasalization.

⁴⁴ In Bésiro, the third person *i-* prefix is ambiguous as to whether or not the possessor is coreferential with the subject. A clitic can be added to specifically mark coreferentiality (Pierric Sans, p.c.).

⁴⁵ Additional counts can be found in Rose (2013b).

⁴⁶ *ma-* with initial /i/ or /e/ is realized *mue* [mwɛ].

⁴⁷ "Yaathe speakers don't like to talk about it" (Januacele da Costa, p.c.).

⁴⁸ The only description that contradicts this is found in Souza (2012), but the methodology is biased. It seems men were asked whether they could speak like women, and their testimony has not been verified in texts.

⁴⁹ Children use female speech until they reach adulthood.

⁵⁰ Abipone data was kindly sent to me by Willem de Reuse.

⁵¹ Mocoví data was kindly sent to me by Pedro Viegas Barros.