Determining Gender Markedness in Wari'

Joshua Birchall
University of Montana

1. Introduction

Grammatical gender is rare across Amazonian languages. The language families that possess grammatical gender properties are Arawak, Chapakuran and Arawá. Wari', a Chapakuran (Txapakura) language of Brazil, is unusual among these languages in that it possesses three gender categories: what we can refer to as masculine, feminine and neuter. While efforts have been made to posit the functionally unmarked gender in Arawak and Arawá languages, no such analysis has been given for any Chapakuran language. In this paper I propose that neuter is the functionally unmarked grammatical gender of Wari'.

Aikhenvald (1999:84) states that masculine is the functionally unmarked gender for non-Caribbean Arawak languages, while Dixon (1999:298) claims that feminine is the unmarked gender in Arawá languages. Functionally unmarked gender refers to the gender that is obligatorily assigned when its designation is otherwise opaque. With attempts, as in Greenberg (1987), to claim a genetic relationship between Chapakuran languages and the Arawak and Arawá families, the presence of gender is a major grammatical similarity and possible motivating factor for such a hypothesis and thus should be examined.

In order to investigate this claim, I first present an outline of the Wari' gender distinctions and then proceed to describe how gender is manifested within the clause. A case is then presented for neuter as the functionally unmarked gender. This claim is further examined for its historical implications, especially regarding genetic affiliation. All of the data[1] used in this paper come from Everett and Kern (1997).

2. Gender Distinctions

The three gender distinctions we find in Wari' are largely determined by the semantic domain of the noun. Even though gender assignment is somewhat arbitrary in the language, Everett and Kern (1997) use two features as general domains for distinguishing the genders: human and female.

The feminine gender is by far the most semantically restricted of the gender classes, containing only human females, children, and human collective nouns that are of mixed sexes. Some examples feminine nouns are *narima* "woman", *panxi* "child", and *'oro wari'* "people". Semantically, feminine nouns can be described as [+feminine], [+human].

The masculine gender has a less restricted semantic domain than feminine, containing nouns that are human males, certain animals, plants, and insects, along with "culturally significant objects" (Everett and Kern 1997:296). Cultural significance is described as "familiar to the Wari' before contact" and "played an important role in the pre-contact worldview". Some example masculine nouns are *tarama'* "man", *hwam*

Santa Barbara Papers in Linguistics, Volume 19, 2008
Proceedings from the eleventh Workshop on American Indigenous Languages
Joye Kiester and Verónica Muñoz-Ledo, eds.

"fish" and *panawo'* "moon". Given the semantic features mentioned above, masculine nouns can be described as [-feminine].

The neuter gender has the least restricted semantic domain for assignment. Non-human entities can be assigned to the neuter class, including inanimate objects, most plants, some animals, fish, and insects, as well as loanwords, body parts and subordinate clauses. Some example neuter nouns are *xe* "fire", *tononoin* "grass", *came* "capivara", and *tocoxi'* "eye". Neuter nouns can be described as having the semantic feature of [-human].

3. Gender Agreement

Gender is not overtly marked on the head noun, such as in the –o/-a distinction in certain Romance languages. Instead, gender is realized through agreement within the clause. Gender agreement features are manifested through verbal inflectional clitics[2] (VICs), possessive constructions, demonstratives and prepositions. Gender is only marked in the third person and only human nouns trigger number agreement.

Verbal inflectional clitics appear at the right edge of the verb and indicate subject and object agreement information along with tense and mood. Agreement between the verb and its arguments is realized on the VIC. A verbal inflectional clitic can be composed of one or two morphemes, depending on whether the clause is transitive. Intransitive VICs, as seen in (1), contain tense and mood information and agree with the subject in number and person. As example (2) illustrates, transitive VICs also contain a morpheme that agrees with the object in person, number and gender.

(1)   mao **na.**
        go   **3s:rp/p**
        "He went" (Everett and Kern 1997:121)

(2)   cao' **'ina-on**   hwam.
        eat   **1s:rp/p-3sm**   fish
        "I ate the fish" (Everett and Kern 1997:152)

Depending on the construction, some VICs can be tenseless, carrying only argument agreement properties. Intransitive and transitive tenseless VICs can be seen in (3) and (4) respectively.

(3)   'om   ca   mao **ca.**
        not:exist infl:nrp/p go   **3sm**
        "He did not go" (Everett and Kern 1997:121)

(4)   ma-in   ca   to' 'ac   **ca-em**?
        that:prox:hearer-n infl:nrp/p hit travel **3sm-2s**
        "Where (on your body) did he hit you?" (Everett and Kern 1997:51)

A second way that gender agreement is realized within a clause is through demonstrative constructions. Demonstratives show a gender agreement relationship with the noun that they modify. In addition to marking for proximity to the speaker or hearer, they agree with the gender of the head noun. The forms only distinguish between neuter and non-neuter nouns, as illustrated in the following examples:

(5) hwam **cwa'**   (6) xirim **ca'**
  fish **this:m/f**   house **this:n**
  "this fish"   "this house"

(7) pana **cain**   (8) carawa **ma'**
  tree **that:n:distal**   animal **that:prox:hearer**
  "that tree (over there)"   "that animal (close to you)"

A third way that gender agreement is realized within a clause is through possessive constructions. Possession relationships can be marked in two ways[3]: with a nominal inflectional clitic (NIC) following the head noun (9) or a possessive suffix (10). Both NICs and the possessive suffixes occur at the right edge of possessed nouns and contain information on the person, number and gender of the possessor.

(9) xirim **nucun** Mirin (10) xiri-**con**
  house poss:**3sm** name:m   house-**3sm**
  "Mirin's house"   "his house"

Prepositional constructions are the final major way that agreement is realized within a clause. There is a single inflected prepositional paradigm used to mark oblique objects on the verb. Prepositions agree with their complements in person, number and gender. Selection of a verbal complement is dependant on the semantic hierarchy in (11) below.

(11) Goal > Circumstance > Theme > Benefactive > Comitative > Location > Time
              (Everett and Kern 1997:4)

All other arguments present in the clause are marked as oblique objects through the use of a preposition. In example (12) below, the Patient *hwam* (m) "fish" is treated as an oblique object through agreement with the preposition. *Humaxicam* (f) "her children" is acting as the Goal and is therefore marked on the VIC.

(12) hoc mi' **nanam** **con** hwam humaxicam
  cook give **3s:rp/p-3pf prep:3sm** fish children-3sf
  "She cooked fish for her children" (Everett and Kern 1997:125)

With an understanding of how gender is realized within a clause, it is now possible to examine the case for neuter as the functionally unmarked gender.

4. Case for Unmarked Neuter
In this section I argue that neuter is the functionally unmarked gender in Wari' based on the following evidence: the grammatical incorporation of loanwords, gender assignment to mixed groups, interrogatives and subordinate clause agreement. This claim is of significance because Everett and Kern (1997) do not analyze for gender markedness.

       The most convincing piece of evidence for an unmarked neuter is that when new words enter the lexicon through borrowing they are "assigned to the neuter gender class" (Everett and Kern 1997:301). The gender of borrowed lexical items in the donor language does not influence gender assignment, e.g., *sal* "salt" is masculine in Portuguese yet is assigned neuter in Wari'.

(13) Example Loanwords from Portuguese

sal "salt"	*canoa* "canoe"
semana "week"	*quilometro* "kilometer"
motor "motor"	*dinheiro* "money, price"
mesa "table"	*segunda-feira* "Monday"

Although noting that multiple factors like phonological shape and semantic domain can influence the gender assignment of loanwords, Poplack et al. (1982:21) state "that loanwords tend to adopt the unmarked gender of the language into which they are borrowed." While Wari' does not condition gender assignment on phonological factors, semantic domain plays a central role. It must be noted that these instances of loanwords from Portuguese, which are all incorporated as neuter nouns, may not fall into the semantic domain of assignment into any gender besides neuter. However, examples in (13) are only the instances in the available materials. I expect that more data will demonstrate that all loanwords in Wari' belong to the neuter gender class.

       A second piece of evidence for an unmarked neuter is that in instances where gender is assigned to groups composed of entities that are both masculine and neuter, the group is always treated as neuter[4]. This process is known as gender resolution (see Corbett 1991: Ch.9). In example (14) below, we see that a group composed of two masculine entities *me* "bird" and *jowin* "monkey" and a single neuter entity *carawa jimao* "different animals" triggers neuter object agreement on the VIC.

(14)   Om   ca   taraju xo'   camain' **cain**   'oro me
          not:exist infl:nrp/p hear   correctly at:all   3sm-**3n** coll. bird(m)

          'oro jowin   cwa' 'oro carawa   jimao.
          coll monkey:spec.(m) this coll. animal(n) different

          "He doesn't hear correctly at all, the birds, the *jowin* monkeys, all different animals." (Everett and Kern 1997:494)

Corbett (1991) notes that markedness is an inadequate explanation for all gender resolution rules since trinary gender systems must also account for groups that fall outside of the domain of the unmarked gender. For example, in Wari' groups of mixed gender humans are treated as feminine because the entities fall outside of the [-human] semantic domain of the neuter gender class. Since semantic factors are primary in the assignment of gender in Wari', this should be of no surprise. However, the fact that neuter overrides masculine even in instances where there is no difference in semantic domain within the group, as in (14), strengthens the claim that neuter acts as a default when gender assignment is unclear.

Opacity in gender assignment is most evident in interrogative structures where it is uncertain which gender the object being questioned belongs to. Wari' primarily employs two interrogative markers, *cain'* "that:n:distal" and derivatives of *ma'* "that:prox:hearer". Both forms are morphologically derived from the demonstrative paradigm. Derivatives of *ma'*, which can be inflected for gender, are used in questions denoting "who?", "whom?" or "which?". Except for the neuter form *main*, which can also mean "how?" or "where?", the derivatives of *ma'* are used in instances where the object being questioned is semantically [+human].

The other interrogative marker *cain'* is used to mark questions denoting "what?", "when?", "how?", "how many?". The objects being questioned typically fall outside of the semantic domain of [+human]. We can see through agreement with the neuter inflectional particle *ca* "infl:nrp/p" that *cain'* still carries its neuter properties.

(15) **Cain'**   **ca**   tomi' cama?
     that:n:distal infl:nrp/p speak 3sf
     'What did she say?" (Everett and Kern 1997:51)

Since it is unclear what gender class the object being questioned belongs to, a neuter form is employed. The use of the unmarked gender in interrogative structures where the gender of the referent is unavailable is not uncommon[5]. Dixon (1995) uses similar criteria from interrogatives in Jarawara (Arawá) to posit feminine as its functionally unmarked gender.

The final piece of evidence that points towards neuter as the functionally unmarked gender in Wari' is agreement with a subordinate clause. All subordinate clauses are introduced by a preposition. Treated as a prepositional object, subordinate clauses trigger agreement as if they were a neuter entity.

(16) Querec te   'inem   **pain**   ca   hwet   mapa'
     See   father:1s 1s:rp/p-2s prep:**3n** infl:nrp/p appear 2s-1s
     "I am happy that you came to me" (Everett and Kern 1997:297)

Since a clause cannot typically possess a grammatical gender, Wari' defaults to neuter because agreement is necessary in a prepositional construction.

The four pieces of evidence presented above all demonstrate that when gender agreement is necessary but opaque, the language uses neuter as the functionally

unmarked gender class for assignment. While semantic properties are the main driving force behind the assignment of gender, when these properties alone are insufficient to determine proper agreement, markedness helps to motivate the designation of grammatical gender. This claim not only contributes to the typological data available for Amazonian languages, but also has some historical implications.

5. Historical Implications

Greenberg (1987) posits a genetic relationship between the Chapakuran language family and the Arawak and Arawá families. He includes Chapakuran languages in an "Arawakan" subgrouping with these two other families. Given the broad scope of his research, Greenberg presents little evidence to motivate this claim. The only evidence he cites is that there are traces of *n-* 1st person singular and *m-* 2nd person singular in the pronouns (p.51). However, this evidence is unconvincing considering the high occurrence of these phonemes in pronominal sets of unrelated languages all across the Americas, e.g. the Algonquian and Salishan families. Acknowledging the paucity of data, he states that his analysis is based off of earlier works by Paul Rivet and others.

In Crequi-Montfort and Rivet (1913), the possibility of a genetic relationship between Chapakuran and Arawak languages was entertained. However, after a comparative analysis between the families, they conclude regardless of notable resemblances that:

> "Despite the considerable number of these matches, we cannot believe that we can, for the moment at least, conclude a parent, even distant. The presence of these foreign words in our vocabulary may indeed be explained by borrowing, since all the peoples of our group are in contact on all sides with Arawak tribes. It is indeed remarkable that these borrowed words never coincide with those that we have established the parents of from different languages of the Chapakuran group" (Crequi-Montfort and Rivet 1913:146, translated by author)

It seems that even Rivet was in doubt of a genetic relationship between Arawak and Chapakuran languages. It is thus unclear what led Greenberg to propose a closer genetic relationship between Arawak and Chapakuran languages than other languages in his Equatorial grouping.

Everett and Kern (1997:3) suggest that another factor may have contributed to Greenberg's analysis: the presence of grammatical gender. However, they dismiss the claim of a genetic relationship between Arawak, Arawá and Chapakuran languages by stating that "they share no obvious cognates and the presence of gender is hardly a sufficiently reliable criterion alone to posit any genetic affiliation." Upon close inspection of these families' gender systems, we find that they are considerably different.

Arawak languages typically have a binary gender class distinction. Aikhenvald (1999:84) claims that masculine is the functionally unmarked gender in the non-Caribbean Arawak languages. Only Ignaciano (Southern) and Palikur (Eastern) show a

neuter distinction in their pronominal systems. At least in the case of Ignaciano, the presence of a third gender category may be a result of contact.[6]

Arawá languages also only possess a masculine/feminine gender distinction. Dixon (1999:298) analyzes feminine as the functionally unmarked gender category. Although its genetic relationship to Arawak languages has been proposed since Ehrenreich (1897), recent scholarship has cast doubt on such a claim. Aikhenvald and Dixon (1999:14) state that "it must be concluded that there is no evidence whatsoever that… the Arawá and Arawak language families are genetically related".

Without a complete reconstruction of Proto-Chapakuran phonology and grammatical forms[7], a comparative approach to resolving the question of its relationship to Arawak languages is untenable. However, we can rely on grammatical properties such as gender systems to provide insight into such questions.

The available data on Chapakuran languages all point towards an unmarked neuter across the family. The Moré lexicon in Angenot de Lima (2000) and a Torá wordlist from Nimuendajú (1925) show that these languages use similar semantic criteria for the assignment of gender as Wari'. The only available research on 'Oro Win in Popky (1999) shows a very similar gender system to Wari'. Without an extensive analysis we can provisionally assume that neuter is the functionally unmarked gender across Chapakuran languages.

If one were to posit a genetic relationship between Arawak, Arawá and Chapakuran languages, an account of the differences in their gender systems is in order, especially since we see a high degree of regularity in these systems within their respective families. An account for the innovation of a third grammatical gender class in Chapakuran languages, as well as a shift in (un)markedness to this novel category, would be critical in demonstrating a genetic relationship among these groups.

Since Greenberg's claims in *Language in the Americas* were published, they have been widely adopted by other branches of social sciences and the public at large. While grammatical properties can help us investigate such claims, more work involving comparative reconstruction is necessary before any definitive claim on genetic affiliation can be reasonably accepted. Therefore, due to a lack of evidence as well as considerable methodological concerns (see Campbell 1991), the claim that Chapakuran, Arawak and Arawá languages share a genetic relationship must be rejected.

6. Conclusions

In this paper, I have demonstrated that neuter is the functionally unmarked grammatical gender in Wari'. This claim is supported by four main facts. First, loanwords are always assigned to the neuter gender class. Second, gender resolution shows that groups of mixed genders trigger neuter agreement when semantically permitted. Third, the neuter gender is used in interrogative structures where the gender of the object being questioned is unclear. Finally, subordinate clauses trigger neuter agreement with the preposition that introduces the clause. This analysis is of significance because no work on gender markedness in Wari' has been attempted.

This claim of gender markedness can be provisionally extended to all Chapakuran languages. In regards to proposals of a genetic relationship between Chapakuran,

Arawak and Arawá languages, there is not significant evidence to support this conclusion and the claim must be rejected. In general, the terms "Arawakan" or "Macro-Arawakan" should be avoided since they acknowledge the existence of this relationship without any basis in an empirically driven approach of comparative reconstruction. Gender markedness in Wari' helps to shed light on the dubiousness of such a proposal by highlighting the stark differences between the gender systems of these languages. As more data are available and the analyses of these languages further develop, we will be able to get a clearer perspective on the historical relationships in Western Amazonia.

Notes
* Thanks to Leora Bar-el, Ryan Denzer-King, Mizuki Miyashita, Anthony Mattina and the audience at WAIL 2008 for their help and comments. Additional thanks goes to Barbara Kern for help with some of the data. All errors are my own.

[1] Because the traditional Wari' orthography employs an apostrophe ['] to represent a glottal stop, a double quotation mark will be used to contain the gloss accompanying the data.
[2] These bundles of agreement information have been analyzed as clitics rather than affixes principally on a phonological basis. For a more detailed discussion of this analysis see Everett and Kern (1997:332-3).
[3] Inalienably possessed nouns (e.g. kinship terms and body parts) may only take the possessive suffix.
[4] Groups consisting of entities of both masculine and feminine gender are treated as feminine. No data is available for groups that contain feminine and neuter entities.
[5] Thanks to Brook Lillehaugen for pointing out that the use of the gender with the broadest semantic domain in questions may result from the need for sufficient vagueness in the proposition.
[6] d'Orbigny (1839) notes a high degree of contact between the Chapakura and the Moxo (Ignaciano) people upriver.
[7] An attempt at a phonological reconstruction of Proto-Chapakura is presented in Angenot de Lima and Angenot (2000).

Abbreviations
rp/p = realis past/present
m = masculine
f = feminine
n = neuter
s = singular
p = plural
coll = collectivizer
infl = inflectional particle
prep = preposition
prox = proximal
poss = possessor

References

Aikhenvald, A. and R.M.W. Dixon (eds.). 1999. *The Amazonian Languages*. Cambridge: University Press.

Aikhenvald, Alexandra. 1999. 'Arawak'. In A. Aikhenvald and R.M.W. Dixon (eds.), *The Amazonian Languages*. 65-106. Cambridge: University Press.

Angenot de Lima, G. 2000. *Description Phonologique, Grammaticale et Lexicale du Moré, Langue Amazonienne de Bolivie et du Brésil*. Porto Velho: EDUFRO

Angenot de Lima, G. and J.P. Angenot. 2000. 'A reconstrução do Protochapakura.' In Hein van der Voort and Simon van de Kerke (eds.), *Indigenous Languages of Lowland South America.* 149-164. Leiden: ILLA, CNWS Publications.

Campbell, Lyle. 1991. 'On so-called Pan-Americanisms (in Notes and Reviews).' *International Journal of American Linguistics*, Vol. 57, No. 3: 394-399.

Corbett, Greville. 1991. *Gender*. Cambridge: University Press.

Crequi-Montfort, G. and P. Rivet. 1913. 'Linguistique bolivienne: La famille linguistique čapacura.' *Journal de la Société des Américanistes de Paris,* 10:119-171.

Dixon, R.M.W. 1995. 'Fusional Development of Gender Marking in Jarawara Possessed Nouns.' *International Journal of American Linguistics*, Vol. 61, No. 3:263-294.

Dixon, R.M.W. 1999. 'Arawá'. In A. Aikhenvald and R.M.W. Dixon (eds.), *The Amazonian Languages*. 292-306. Cambridge: University Press.

d'Orbigny, A.D. 1839. 'Nación Chapacura.' *L'homme américain de l'Amérique méridionale, considéré sous ses rapports physiologiques et moraux*. 340-3. Paris.

Ehrenreich, P. 1897. 'Materialen zur Sprachenkunde Brasiliens: Vokabuläre von Purus-Stämmen.' *Zeitschrift für Ethnologie*, 24:59-71.

Everett, Daniel and Barbara Kern. 1997. *Wari': The Pakaas Novos Language of Western Brazil.* London: Routledge.

Greenberg, Joseph. 1987. *Language in the Americas*. Stanford: Stanford University Press.

Nimuendajú, Curt. 1925. 'As tribus do Alto Madeira.' *Journal de la Société des Américanistes de Paris*, 17:137-172.

Popky, Donna. 1999. *Oro Win: A Descriptive and Comparative Look at an Endangered Language*. University of Pittsburg. M.A. Thesis.

Poplack, S., A. Pousada and D. Sankoff. 1982. 'Competing Influences on Gender Assignment.' *Lingua,* 57:1-28.