Sociocultural gender in nominal classification: A study of grammatical gender

Marc Allassonnière-Tang, Hiram Ring

SOCIOCULTURAL GENDER IN NOMINAL CLASSIFICATION: A STUDY OF GRAMMATICAL GENDER

MARC ALLASSONNIÈRE-TANG* & HIRAM RING**
** University Lyon 2, France & **University of Zurich, Switzerland*
marc.tang@univ-lyon2.fr & hiram.ring@uzh.ch

ABSTRACT

We analyse how sociocultural gender can be reflected through grammatical gender and select Hindi (Indo–European) and Pnar (Austroasiatic) as case studies. We demonstrate that these grammatical gender systems share universal tendencies based on human cognition, i.e. associating long, thin, and vertical objects with masculine grammatical gender whereas round, flat, horizontal ones are associated with feminine grammatical gender. We also show that these grammatical gender systems distinguish between sociocultural values of the language speakers. Speakers of Hindi maintain a patrilineal kinship system, and in their language objects of large size are generally assigned to the masculine gender. Pnar kinship is matrilineal and in the language large sized objects tend to be associated with feminine gender. Similar asymmetries are observed with regard to generic gender and gender reversal. These results contribute to the impact of universal cognitive principles and culture on grammatical structures by showing that both tendencies are not necessarily complementary and that they can co-exist in the same language.

Keywords: Grammatical gender, Hindi, Pnar, Sociocultural gender, Cognition.

1 Introduction

This paper provides data showing that grammatical gender in language not only can echo universal cognitive principles, but can also mirror cultural idiosyncrasies displayed in sociocultural gender. Grammatical gender is a linguistic system of agreement that can reflect these two types of biases in languages of the world (Corbett, 1991; Seifart, 2010). For example, grammatical gender in French reflects biological gender by categorizing nouns of the lexicon as either masculine or feminine (Corbett, 2013). Languages with masculine/feminine grammatical categories may assign gender to an animate or inanimate object based on the similarity of its salient features with properties generally associated with masculine or feminine characteristics in the society at large, though biological sex and physical properties are not the only types of gender represented in language. Sociocultural gender also plays an important role within languages (Eckert & McConnell–Ginet, 2003; Hall & Donovan, 1996; Hellinger & Motschenbacher, 2015); for example, the use of appropriate speech style and vocabulary may vary according to the social factors associated with women versus men.

Previous studies have proposed that grammatical gender systems share common tendencies

based on human cognition and perception. Long, thin and vertical objects are generally associated with masculine grammatical gender whereas round, flat and horizontal ones tend to be associated with feminine grammatical gender (Aikhenvald, 2012). Size is also a salient feature taken into account by grammatical gender systems. As Kemmerer (2017, p. 412) observes, "Not surprisingly, there is a strong tendency for these languages to treat relatively large things as masculine and relatively small ones as feminine". A few languages such as Hamar (Afro–Asiatic) are attested to have the opposite association, whereby big objects are generally affiliated with feminine gender and vice–versa. Unfortunately, comparative studies targeting gender assignment do not provide a detailed explanation for these different size–gender connections. Furthermore, the distinction between gender assignment based on cognitive principles vs. assignment based on cultural idiosyncrasies has not been investigated from a cross–linguistic perspective.

In this paper, we propose that lexical gender assignment in relation to size (among other features such as generic gender) may be motivated by cultural patterns in particular languages instead of cognitive principles. We demonstrate how sociocultural genders are interpreted, via case studies of grammatical gender in Hindi (Indo–Aryan, Indo–European) and Pnar (Khasian, Austroasiatic). In Hindi, whose speakers have a patrilineal kinship system, large sized objects are assigned the masculine gender (Contini–Morava & Kilarski, 2013). In Pnar, whose speakers have a matrilineal kinship system, large sized objects tend to be associated instead with the feminine gender (see also Khasi: Rabel–Heymann, 1977). The two case studies present a perspective from ethnosyntax (c.f. Enfield, 2002a) highlighting the philosophical alignment of the kinship systems of the speakers with the grammatical realizations of gender marking in each language. This comparison suggests the co-existence of universal cognitive principles and specific sociocultural gender associations for the realization of grammatical gender systems of different languages.

We should be clear here that we do not claim either a local or a universal causal link between the social structures of speakers and the grammatical instantiations of gender marking on the basis of just these two languages, but rather would observe that the plausibility of such a link deserves further investigation. As Enfield (2002b: 19) notes, "Explicit comparison between particular languages and cultures is... a most effective method [for investigating ethnosyntax], especially where they differ in minor, but specific ways, throwing the more subtle linguistic and/or cultural differences into sharper relief." While a quantitative analysis is beyond the scope of the current paper, this is a logical next step, requiring both a more detailed analysis of the sociocultural gender associations for each language/culture (and their possible sources) as well as a larger database of languages with grammatical gender cross-referenced with sociocultural/kinship patterns (see Kirby et al, 2016; Dryer & Haspelmath, 2013 for possible models). In the sections that follow we provide an overview of gender in society and language (§2), describe the qualitative methodology of the current paper (§3), provide an analysis of gender in Hindi and Pnar (§4) with regard to size, generic gender, and gender reversal, and conclude with an outlook toward further research (§5).

2 Gender in society and language

Gender is a salient feature of humans and living beings. Hence, "The multifaceted notion of gender pervades every aspect of life and of living" such that gender may be divided into three main domains (Aikhenvald, 2012, p. 33–35):

- *Biological gender* (natural gender, sex) reflects the physical properties of an individual's reproductive system. In human society, biological gender is generally interpreted as the male versus female distinction.
- *Sociocultural gender* (social gender) refers to a social role based on the biological gender of the person or his/her personal gender identification. As a social category, sociocultural gender is "constituted on the structural, the symbolic, and the individual levels in society" (Scantlebury 2014, p. 1–2). Sociocultural gender involves "gender etiquette, traditional knowledge and social stereotype" (Aikhenvald 2012, p. 32). An example would be the behaviour expected by the society at large for a man and a woman during a specific event (e.g. private conversation, dating, among others).
- *Grammatical gender* (linguistic gender) is the grammatical marking that categorizes entities within the system of a language. It is generally realized through morphological marking and grammatical agreement (Corbett, 1991). This categorization may be based on animacy, biological gender and/or other semantic concepts. For instance in French, all nouns of the lexicon are assigned to either masculine and feminine gender, which is reflected via grammatical agreement on the article, adjective and verb (among other parts of speech).

These three categories are part of a broad concept of gender and show mutual influence. Physical features of biological gender often serve as the basis of sociocultural gender and grammatical gender. As an example, women are commonly attested to use high pitch in conversation. Biologically, this is not only due to a different size of vocal tract compared to men but also an effect of attraction (Fraccaro et al., 2011). Such a tendency may also be reflected in sociocultural genders, e.g. the idea that high pitch is appropriate for women but not for men. This phenomenon is "not a product of physical factors alone... but rather an expressive posture which is socially more appropriate for one sex or another" (Labov, 1972, p. 304). Biological gender is also reflected within nominal classification systems in languages, e.g. grammatical gender. The way objects are classified in a language reflect not only universal tendencies but also "culture–specific ways of categorizing the intra–linguistic realm of nouns and, by extension, the extra–linguistic realm of objects" (Kemmerer 2017, p. 417).

2.1 Sociocultural gender

Biological gender is a universal distinction among human beings and has a strong bearing on societal formation, which gives rise to sociocultural gender through the interactions between people of different sexes. As a concept, sociocultural gender is defined as the "social expectations about behaviour regarded as appropriate for the members of each sex" (Giddens, 2004, p. 689). In this sense, sociocultural gender thus refers to the differences in the socially formed traits of masculinity and femininity rather than differences in physical attributes between sexes.

One way to analyse such interaction across societies regardless of their socio–economic or political orientations is through language. The majority of previous studies on language and gender have focused on the use of language by speakers of different sexes (Holmes, 1991), e.g. different features used by men and women within speech or the behaviour of different sexes in discourse interaction (Lakoff, 1973, 1975; Preisler, 1986). Within this field, the need for cross–

cultural comparison plays an important role (Hymes, 1962; Hymes, 1974; Sherzer, 1987), as the sociology of language is the study of the relationship between language and society (Chen, 1997).

A main contribution of cross–cultural comparisons of sociocultural gender is to distinguish between universal values and cultural idiosyncrasies. As an example, the use of indirect and attenuated forms of speech is commonly attributed to women across languages (Brown, 1980; Koike, 1986; Ide et al., 1986), but comparative sociolinguistic studies demonstrate that the use of polite and indirect speech by the different sexes is culturally dependent (Holmes 1991, p. 212). This is shown by the fact that in Malagasy (Keenan, 1974) and Guyana (Edwards, 1979) men use more indirect and allusive speech while confrontational and direct speech forms are only acceptable for women.

2.2 Grammatical gender

Grammatical gender (or noun class) is a feature of languages that distinguish between (sub–)classes of words (typically nouns) through some morphological means (Seifart, 2010). Typically, the term 'gender' is used for languages in which classes have a clear sex–based (masculine/feminine) component, or for systems with few distinctions. The term 'noun class' is used for languages in which other semantic features (such as animacy) are primary, or for systems with a larger set of class distinctions. In this study, we restrict ourselves to a selected group of smaller noun class systems in which there is a clear sex–based component, in order to explore other sociocultural dimensions of each system. Throughout this study we will use the term 'grammatical gender' to discuss these systems.

Corbett (1991) notes that the primary criteria for distinguishing a grammatical gender or noun class system is the feature of 'agreement'. If a noun triggers morphological agreement on another word class, it can be said that the language has grammatical gender. As demonstrated in (1), the two clauses display similar number, case and syntactic structure yet the different grammatical genders (masculine/feminine) of the nouns are reflected on the numeral, adjective and verb.[1]

(1) Gender agreement in French[2]

a.	*un*	*grand*	*livre*	*est*	*tombé*
	one.MASC	big.MASC	book(MASC)	be	fall.PARTIC.MASC
	'A big book fell.'				
b.	*une*	*grande*	*table*	*est*	*tombée*
	one.FEM	big.FEM	table(FEM)	be	fall.PARTIC.FEM
	'A big table fell.'				

Grammatical gender is not considered a feature of languages that encode the semantic feminine/masculine distinction of nouns without triggering agreement. For instance in Mandarin Chinese (Sinitic), the feminine/masculine contrast exists lexically, e.g. *ge*^{1}*ge*^{1} 'big brother' and *jie*^{3}*jie*^{0} 'sister'. However, there is no grammatical agreement between the noun and other elements of the clause, as shown in (2). Hence, Mandarin Chinese is not considered a grammatical gender language, though it does rely on another nominal classification system, classifiers (Aikhenvald,

2000; Zhang, 2013; Grinevald, 2015).

(2) Absence of gender agreement in Mandarin Chinese (Sino–Tibetan)

a.	*yi*	*ben*	*da*	*shu*	*diaoxialai*	*le*
	one	CLF_{volume}	big	book	fall	PERF
	'A big book fell.'					

b.	*yi*	*zhang*	*da*	*zhouzi*	*diaoxialai*	*le*
	one	CLF_{2D}	big	table	fall	PERF
	'A big table fell.'					

Nouns in standard Hindi are affiliated with either masculine or feminine gender, which is reflected through grammatical agreement on other elements of the clause such as modifiers and verbs (Kachru, 2006, p. 46). As shown in (3), the article, adjective and verb agree with the respective gender of the noun.

(3) Gender agreement in Hindi

a.	*nayā*	*ṭrak*	*āyā*	*hai*
	new.MASC	truck(MASC)	come.PAST.MASC	be.PRES
	'The new truck has come.'			

b.	*naī*	*frij*	*āī*	*hai*
	new.FEM	fridge(FEM)	come.PAST.FEM	be.PRES
	'The new fridge has come.'			

This distinction of grammatical gender is "present in all Hindi dialects (with the exception of lingua franca or pidginized varieties where grammatical gender is often lost altogether)" (Hall, 2002, p. 134). While the gender of animate nouns generally corresponds to the biological gender of the referent (e.g. *larka* 'boy' is masculine and *larki* 'girl' is feminine), gender assignment for inanimate nouns is considered to be arbitrary (e.g. *dvaar* 'table' is masculine and *khirkee* 'window' is feminine). Hence, variation may be observed across dialects, e.g. *dahi* 'yogurt' is affiliated with masculine gender in eastern dialects and feminine in western ones (Nespital, 1990; Hall, 2002). Moreover, due to contact with languages without grammatical gender, some dialects reflect the gender distinction only on animate nouns, while inanimate nouns default to the unmarked masculine form. This is also observed for Indo–European languages located in the North and North–East of India, where there is contact with Sino–Tibetan classifier languages (Barz & Diller, 1985; Aikhenvald, 2000). In Nepali, for example, which also has a masculine/feminine grammatical gender system, all nouns are by default masculine, except for nouns referring to feminine animates (Acharya, 1991, p. 99). In this paper, we focus on standard Hindi (which has a productive gender agreement system), partly due to speaker population and availability of data.

Grammatical gender in Pnar is marked on nouns via proclitics *u*= 'masculine', *ka*= 'feminine', and *i*= 'neuter, diminutive' (Ring, 2015, p. 101). Along with the plural marker *ki*=, these serve as agreement markers within the Pnar noun phrase (4). Gender is also reflected in the pronominal paradigm of Pnar and other Khasian varieties, whereby second person singular forms

show a masculine/feminine distinction and third person singular pronouns have the same 3–way distinction as the proclitics, with similar forms.

(4) Gender agreement in Pnar

a.	*pat*	*u=ni*	*u=syntu*	*u*	*wa*	*heh*
	fall	MASC=PROX	MASC=flower	MASC	REL	be.big
	'This big flower fell.'					

b.	*pat*	*ka=ni*	*ka=miej*	*ka*	*wa*	*heh*
	fall	FEM=PROX	FEM=table	FEM	REL	be.big
	'This big table fell.'					

Pnar gender markers on animate nouns primarily identify the sex of the referent. Inanimate nouns may have assigned gender, but some nouns can occur with all gender markers for different shades of meaning. In many cases the gender marker seems to be associated with other semantically salient features of a noun, such as shape, size, and position; this is similar to Khasi (Rabel–Heymann, 1977) and is discussed further below. Other relevant features of the Pnar gender system are that 1) these markers can derive nouns from verbs, with each marker functioning as a different kind of modifier (*u*= 'nonfinite, gerund', *ka*= 'resultative', *i*= 'abstract'; Ring 2014), and 2) agreement only occurs within the noun phrase and not on the verb.

2.3 Functions of grammatical gender

The benefits of grammatical gender are not immediately obvious, yet previous studies have shown that grammatical gender has various discourse and lexical functions. In terms of discourse functions, the use of grammatical gender marking allows the hearer to constrain the set of nouns that can fit into a particular slot. Knowing that the noun MUST be feminine, for example, reduces the strain on human subjects in processing tasks (Lew–Williams & Fernald, 2007; Lew–Williams & Fernald, 2010). These types of discourse functions of grammatical gender are typically represented by anaphora, deixis and disambiguation (Contini–Morava & Kilarski, 2013). By way of illustration, in (5a), the referents 'son' and 'daughter' are introduced by their proper names in the beginning of the clause. In the second clause (5b), the speaker refers to only one of the referents. Nevertheless, there is no ambiguity since the listener can interpret speaker's intended referent via the masculine gender agreement.[3]

(5) Grammatical gender and disambiguation in Hindi

a.	*merā*	*beṭā*	*aur*	*merī*	*larkī*	*kal*	*aeṁge*
	my.MASC	son	and	my.FEM	daughter	tomorrow	come.FUT.MASC.PL
	'My son and my daughter are coming tomorrow.'						

b.	*lekin*	*vah*	*to*	*der se*	*aega*
	but	3SG	FOC	late	come.FUT.MASC.SG
	'But he, he will come late.'				

The functions of grammatical gender directly related to our study are lexical functions. These functions include expansion of the lexicon, differentiating referents and ascribing properties (Contini–Morava & Kilarski, 2013). The first two functions can be summarized as follows: grammatical gender marking may be used 1) to create a new lexical item via a derivational process (expansion of the lexicon) and 2) to differentiate between referents by using a different gender marker on the same nominal stem (differentiating referents). For instance, in Italian, a different grammatical gender marker helps to distinguish between trees and their fruits, and thus expand the lexicon, e.g., *melo* 'apple tree' is masculine while *mela* 'apple' is feminine (Ferrari, 2005, p. 39–40). As for differentiating referents, in Nepali, different gender suffixes can be used with nominal stems to indicate the sex of an animate referent, for example: *cora* 'son' vs. *cori* 'daughter' and *kaka* 'uncle' vs. *kaki* 'aunt'.

Gender marking can also ascribe different properties to nominals, such as by indexing or expressing the speaker's attitude toward the referent. Such uses are highly language– and culture–specific. In the case of languages with grammatical gender, this function depends not only on the available contrasts but also on the sociocultural connotations that gender has in a given language or dialect. In other words, attitudes toward the referent can be conveyed by the choice of a different gender than the one normally used, e.g. by using a masculine noun with reference to a woman, or by gender shift, i.e. converse use of agreement forms. In Modern Hebrew (Semitic), for example, masculine pronouns and verbal morphology are used to refer to females by both male and female friends and relatives to convey affection and intimacy (Tobin, 2001). Such possibilities are also considered below.

3 Methodology

In our study of two languages with grammatical genders (Pnar and Hindi), we adopt a functional approach (Contini–Morava & Kilarski, 2013) and demonstrate that even though these grammatical gender systems share common tendencies dependent on human cognition and perception, they also reflect the cultural idiosyncrasies of the two different languages.

3.1 Data

We selected nominal classification and more specifically grammatical gender as our scope of analysis for sociocultural gender for several reasons. As stated by Lakoff (2003, p. 162–63), "In order [for us] to understand the world and [to] function in it, we have to categorize, in ways that make sense to us, the things and experiences that we encounter". This need is fulfilled in language via various means of categorization, one of the most prominent being grammatical gender (Aikhenvald, 2000; Corbett, 1991). Moreover, grammatical gender assignment is underpinned by cognitive and sociocultural principles (Aikhenvald, 2012; Corbett, 1991; Kemmerer, 2017), so it is expected to reflect certain aspects of sociocultural gender (Boroditsky, 2009; Deutscher, 2010). Sociocultural gender is even more transparent through grammatical gender if such a system is based on biological gender (i.e., masculine/feminine), since the biological and sociocultural distinction is "so fundamental to social organization and social structure that linguistic means to refer to this category are indispensable for speech communities" (Stahlberg et al., 2007, p. 163).

We chose Pnar and Hindi according to the following criteria. First, our analysis requires languages with relatively stable grammaticalized gender. Within gender languages, we may distinguish between languages in which gender is 'fixed' (i.e. one lexical root can be associated with one and only one gender; as in French and German) versus those in which it is 'flexible' (i.e. a lexical root can be associated with multiple genders to distinguish different shades of meaning). Pnar and Hindi are examples of the latter – the grammatical gender system allows for variation such that the kind of sociocultural dimensions we are interested in can be reflected via gender markers. In the case of Pnar, multiple genders can be associated with a single lexical root (particularly in the case of animate referents). In the case of Hindi, variation in terms of gender assignment is also present within the language, as mentioned in §2.2.

Second, although the two languages each have a flexible grammatical gender system, they reflect different sociocultural features with regard to how gender is practised and understood. Hindi society is patrilineal (whereby inheritance passes to male heirs) and has been attested to reflect patriarchy and its related gender hierarchy (Shabadi, 2005, p. 258). Pnar, on the other hand, is a matrilineal society (whereby inheritance passes to female heirs) in which woman have a larger amount of social and economic responsibility, and is suggested to reflect matriarchal structural properties (Pakyntein, 2000, p. 28). Finally, as noted briefly in the introduction, we acknowledge that this study could be criticized for its lack of comparative scope, since we limit ourselves to two languages from separate language groups (Austroasiatic and Indo–European) with relatively different population sizes. Although this means that our data cannot be used for any statistically-motivated analysis to infer correlation, we find it beneficial to focus on these two languages in order to provide a qualitative analysis that may serve as a beachhead for future large–scale cross–linguistic comparison.

3.2 Method

Our analysis follows the methodology of Hall (2002) and targets three aspects of interaction between grammatical gender and sociocultural gender: *gender assignment*, *generic gender* and *gender reversal*. First, *gender assignment* relates to how grammatical gender is assigned to nouns of the lexicon. By way of illustration, a certain physical property of a referent may result in the noun being affiliated with either masculine or feminine gender in Hindi and Pnar. Second, *generic gender* identifies which gender represents the unmarked gender in a language. For instance, a particular form of gender marking may refer to the generic sense of the noun. Finally, *gender reversal* is also a lexical function of grammatical gender. A speaker may deliberately switch the grammatical gender of a noun to convey his or her attitude toward the referent of the noun, as in the example of Modern Hebrew in §2.3.

With regard to gender assignment, our analysis focuses on language–specific properties of the two languages in question rather than gender assignment strategies which are shared cross–linguistically. Languages of the world do share common features and "all human groups use language in a similar way" (Calude & Pagel, 2011, p. 1106). One of the shared behaviours across languages is how categories of object concepts are realized in the human brain (Clarke & Tyler, 2015; Clarke, 2015; Gainotti, 2006; Mahon & Caramazza, 2009; Martin, 2007; L. Chen & Rogers, 2014). Grammatical gender serves to identify particular nouns as part of a class, allowing speakers to organize these items conceptually. Contini–Morava & Kilarski (2013) describe how this organizational strategy follows patterns that seem to reflect a human cognitive and perceptive bias.

Specifically, when we compare the form of objects that are typically viewed cross–linguistically as belonging to the 'masculine' gender, they tend to be long, thin, or vertical. Similarly, those that are typically 'feminine' include round, flat, and horizontal objects (Kemmerer, 2017, p. 408).

Such a statement is also valid in Pnar and Hindi. Pnar's grammatical gender subsumes a mixed bag of semantic features, including sex, shape, and size. Even though the gender system in Pnar is not as fully fixed as in a language like French or German, the following generalizations can still be made. First, masculine gender *u=* and feminine gender *ka=* are associated with the biological gender of animate nouns. Second, for inanimate nouns, masculine gender is associated with long, thin, or upright objects (trees, sticks, narrow stones), while feminine gender is associated with flat, covering, or circular objects, objects with openings or objects used to hold things (a floor, cloth, a cave, bowls, pots). Third, exceptions to the second generalization often have to do with whether the object is conceptualized as a tool associated with activities typically performed by one or another of the sexes (in which case it takes the gender/sex corresponding to the prototypical doer of the activity), or whether the object is anthropomorphized in traditional stories (in which case it takes the gender/sex of the animate referent in the story).

Such tendencies are also observed in Hindi. For instance, names of trees are generally masculine, e.g., *piipal* 'fig–tree', *niim* 'neem tree'. Moreover, minerals and jewels such as *lohaa* 'iron' and *hiiraa* 'diamond' are also masculine (Agnihotri, 2007, p. 48). As summarized in Table 1, basic principles of cognition are reflected through grammatical gender both in Hindi and Pnar, such that masculine is associated with long, thin and upright objects, while feminine is associated with flat, circular objects.

Table 1: Cognitive patterns in gender assignment for Hindi and Pnar

Feature	**Hindi**	**Examples**	**Pnar**	**Examples**
Long	masc	*kalam* 'pen'	masc	*pseiñ* 'snake', *tylle* 'rope'
Thin	masc	*taar* 'wire'	masc	*tkut* 'sword'
Vertical	masc	*dayaar* 'door', *ghar* 'house'	masc	*jhep* 'plant', *klong* 'bottle'
Round	fem	*plate* 'plate'	fem	*leiñ* 'boat'
Flat	fem	*mez* 'table', *kitaab* 'book'	fem	*meij* 'table', *kot* 'book'
Horizontal	fem	*chat* 'roof'	fem	*nan* 'lake'

Another interesting feature is how variation in the association of gender marking on nouns corresponds to the conceptualization of the properties of the noun. This reflects the lexical function of differentiating referents. For example, in Pnar the noun *deiñ* 'tree' can be marked by both masculine and feminine proclitics.[4] The masculine form *u=deiñ* refers to a standing, upright tree, while *ka=deiñ* refers to a fallen tree (log) or to wood more generally. Similarly, *madan* 'ground' (being flat) is usually marked by the feminine proclitic, except in the case of compound nouns like *u=madan futbol* 'football pitch', where the associated game is typically a male sport.

The consistent usage of grammatical gender agreement itself may also represent different sociocultural registers in discourse. In Hindi, grammatical gender agreement is considered as standard by both speakers of Eastern and Western varieties, but it can also be given different social evaluations (Hall, 2002, p. 138–39). While speakers of Western varieties tend to associate the lack of standard grammatical gender agreement with illiteracy and disrespect, speakers of certain Eastern varieties consider the use of standard grammatical gender agreement as marked and distant (Simon, 1993, 1996).

Such observations support the statement of McConnell–Ginet (2014, p. 36–37): "To what extent the kind of gender system in a language constrains or promotes gender equity is not clear. What is clear however... [is that] Neither 'linguistics' nor sociocultural gender is natural if by that we mean impervious to change, isolated from human ideas and actions." Following this claim, even though the analysis of gender assignment as a whole also relates indirectly to the topic at hand, due to the limitations of space and resources (e.g. lack of large corpora for both languages) we focus on gender assignment in terms of size, along with generic gender and gender reversal (Hall, 2002).

4 Analysis

In this section, we first provide a brief overview of sociocultural gender in Hindi and Pnar societies. Then we analyse the interaction between sociocultural gender and grammatical gender in terms of size, generic gender and gender reversal. It is important to point out that we approach the sociocultural gender hierarchy within both societies from a descriptive view, describing the interaction between men and women in the societies based on the existing literature. The hierarchy mentioned is by no means an absolute rule of the respective societies. It is rather a tendency observed by previous researchers.

With regard to Hindi society, official laws in India stipulate that men and women share the same rights, and rethinking the gender perspective in Indian politics is also a recent trend in Hindi society (Chakrabarty & Pandey, 2008, p. 316). Nevertheless, sociocultural gender in Hindi society tends to position men higher than women in a hierarchy (Shabadi, 2005). This division based on biological gender is observed among different class groups and often attested within the context of family. By way of illustration, the family ideology suggests that women should eat less than men during a meal or that their incomes should be placed within the hands of the male decision–makers in the family. Women are mostly in charge of home–related issues, but can also play an important role in religious rituals such as wedding ceremonies and funeral rites, while men are typically the priests and overseers of such rites and concerns outside of the home.

Sociocultural gender has been a topic of interest for researchers in the Indian State of Meghalaya (where Pnar is located) for some time, namely due to the prevalence in this area of social organization predicated on matrilineal systems (Gurdon, 1914; Ehrenfels, 1955; Nongbri, 1988, 2000; Gneezy et al., 2009). In these systems, inheritance passes from the youngest daughter of a family to her youngest daughter (Pnar: *ka=kha–dooh* lit. 'F=born–last' or 'last–born female'), which comes with a responsibility to maintain the family's wealth for the benefit of all relatives. Children are considered part of the mother's clan rather than the father's, and both female and male children take their mother's surname (or clan name). At the same time, family decisions are shared, and the eldest brother has major decision–making and religious responsibilities. This matrilineal kinship pattern affords much respect to women in the society. Within the Khasian context, however, studies of linguistic facets of sociocultural gender are lacking. Thus, we also aim at beginning to fill this gap.

4.1 Sociocultural gender with regard to size

Besides the connection with shapes of objects, grammatical genders also return sociocultural

values of language speakers in different forms. As an example, the patrilineal kinship system of Hindi society and patriarchal tendencies seems to be reflected in the Hindi language by the fact that objects of larger size are affiliated with masculine gender (Contini–Morava & Kilarski, 2013, p. 271). As an example, in Hindi, a small needle is feminine while a big needle is masculine. In Pnar, large objects tend to be associated instead with feminine gender, which may also reflect the sociocultural values of the Pnar people, whose kinship system is matrilineal. For instance, in Pnar, a small mountain is masculine while a big mountain is feminine. Additional details and examples are provided in the following paragraphs.[5]

In Hindi, the gender of nouns may be used to expand the lexicon (Contini–Morava & Kilarski, 2013, p. 271–72) along some semantically consistent lines depending on sex and size (Shukla, 2009, p. 496). Among objects from the same class, those with such features as big, rough and coarse are assigned masculine gender while smaller, finer and more refined objects are feminine. As an example among automobiles, cars and buses are feminine whereas trucks and dumpers are masculine. This distinction is commonly realized by assigning different gender suffixes on the root of a word, e.g., if you look at needles as a class of object, *sui* 'small needle' is feminine while *sua* 'big needle to stitch sacks' is masculine. In ropes, *rassi* 'small rope' is feminine whereas *rassa* 'thick rope' is masculine. Regarding growing plants, *ped/vraksh* 'tree' is masculine whereas *jhadi* 'shrub' and *lata/bel* 'creeper' are finer, smaller and thus feminine. More examples are shown in Table 2.

Table 2: Sociocultural gender and grammatical gender affiliation in Hindi

Masculine	**Feminine**
dabbaa 'box'	*dabbii* 'little box'
ghantaa 'bell'	*ghantii* 'small bell'
jholaa 'bag'	*jholii* 'bag or wallet'
laathaa 'large club'	*laathii* 'club, nightstick'
lakaraa 'log'	*lakarii* 'stick'
pattaa 'strap, belt'	*pattii* 'ribbon, cloth strip'

Similarly to Hindi, the Pnar gender markers are also associated with size, but in an inverse manner: masculine gender usually (but not always) denotes smaller relative size than feminine gender (i.e. *u=lom* 'small hill/mountain' vs. *ka=lom* 'big hill/mountain' and *u=pung* 'small pond' vs. *ka=pung* 'large pond'). This parallels the differential status of women in Hindi (patrilineal) and Pnar (matrilineal) society, and suggests that gender markers in both languages are indexing sociocultural status. For Pnar, we should clarify that the availability of the neuter/diminutive gender marker *i=* can obscure this size feature of the gender system, in part because diminutive size may be indexed by a separate grammatical marker that fits within the gender system. However, it may be better to view the three gender markers as part of a triad (mother, father, child) with reference to sex of the referents. Under this conception mother=female, father=male, and child=neuter. If we strictly confine ourselves to size, we see that (all other things being equal) mother=female=large, father=male=medium, child=neuter=diminutive (this applies to non-human animates as well). But for inanimate objects that take variable gender assignment, the more salient contrast is between female and male, whereby the former is typically associated with a larger version of the object and the latter is associated with a smaller version of the object. When neuter is used to mark the object, this often means it is under–specified for size. These generalizations

have apparent exceptions in the language, as noted for Khasi by Rabel–Heymann (1977). Many such exceptions can be accounted for as borrowings, whereby borrowed words tend to take the feminine gender marker *ka=* as 'default', but the complexities of this system still require more careful study. Some masculine/feminine word pairs in Pnar are given in Table 3.

Table 3: Sociocultural gender and grammatical gender affiliation in Pnar

Masculine	**Feminine**
u=pung 'small pond'	*ka=pung* 'big pond'
u=loom 'small hill'	*ka=loom* 'big hill or mountain'
u=chang 'small basket'	*ka=chang* 'large basket'
u=kchu 'small pot/pitcher'	*ka=kchu* 'large pot/pitcher'
u=tadong 'small cup'	*ka=tadong* 'large cup'

The 'default' usage of the feminine marker and the association of feminine gender with large size seems to reflect a larger sociocultural value of the Khasian family of lects, whose speakers share a matrilineal kinship system. As noted above, Khasian clans trace their origin to a founding mother and the inheritance of a nuclear family passes to the youngest daughter. The centrality of the female sex thus seems to be reflected in the grammatical gender system of these varieties. To summarize in Table 4, we observe an apparent synchronization between sociocultural gender and grammatical gender in both languages.

Table 4: Sociocultural gender and grammatical gender affiliation in Hindi and Pnar

Feature	**Hindi**	**Examples**	**Pnar**	**Examples**
Large	masc	*dabbaa* 'box'	fem	*ka=chang* 'large basket'
Small	fem	*dabbii* 'small box'	masc	*u=chang* 'small basket'

In the patrilineal Hindi society, within the same class of objects, large size tends to be affiliated with the masculine grammatical gender while small size is generally associated with the feminine grammatical gender. The reversed tendency is attested in Pnar, a matrilineal society, whereby objects described by the same lexeme tend to be assigned masculine and feminine gender depending on their small and large size respectively. Once again, we should clarify that the current analysis is qualitative and additional data should be considered to investigate this correlation quantitatively. Further details on this limitation are provided in §5.

4.2 Sociocultural gender and generic gender

Generic gender refers to which grammatical gender is the unmarked or default value in the language. This asymmetry between marked and unmarked gender can thus reflect the sociocultural gender of the society. By way of illustration, if the generic gender of a language is androcentric, the feminine status is less visible in discourse since it is not lexically marked when it does not need to be (Cameron, 1998; Hellinger & Bußmann, 2001). Such grammaticalization of gender thus "strongly contributes to the salience, or accessibility, of the social category gender" and is viewed as a "reflection of intergroup hierarchies" (Gabriel & Gygax, 2016, p. 1). For instance, the animate referent of a noun may be unspecified in terms of biological gender and be given the

generic interpretation in reference. In Hindi, for example, *sathi* 'companion' is assigned masculine grammatical gender even though *sathi* may refer to either a male or a female companion. Additional examples in Table 5 show that the masculine grammatical gender in Hindi is used as a generic inclusive of both male and female referents (Hall, 2002, p. 143–44). Feminine versions of the nouns listed in Table 5 are attested, but such terms refer specifically to a female referent and cannot be used in the generic sense. By way of illustration, the feminine equivalent of *yatri* 'traveler' would be *yātriṇī* (or *yātrā karnevālī*). However, the feminine forms may only refer to female travelers, as opposed to *yatri* which can designate both men and women.

Table 5: Generic gender in Hindi

Noun	**Meaning**
dost	'friend (male or female)'
sathi	'companion (male or female)'
mitra	'ally (male or female)'
yatri	'traveler (male or female)'
ghursavar	'rider (male or female)'

This tendency is also observed within verbal agreement of grammatical gender (Kachru, 2006, p. 191) and is displayed in (6) via an example from Hall (2002, p. 144). The term *koi* 'someone' may refer to both female and male referents semantically. However, only the masculine agreement on the verb carries such a double interpretation. Feminine agreement on the verb solely identifies a female referent.

(6) Generic gender and verbal agreement in Hindi

a. *koi* *aya* *hai*
someone come.PERF.MASC.SG be.PRES
'Someone (male or female) has come.'

b. *koi* *ayi* *hai*
someone come.PERF.FEM.SG be.PRES
'Someone (female) has come.'

Conjoined nouns also reflect sociocultural gender through grammatical agreement. In Hindi, when a conjoined noun includes both masculine and feminine animate conjuncts, the verb is consistently in the masculine agreement form. This is shown in (7a), where the conjoined noun includes both female and male animates, and the verb agrees in masculine gender, whereas feminine agreement (7b) occurs only when all the animate referents are feminine.

(7) Verbal agreement with masculine and feminine subjects in Hindi

a. *Marc* *aur* *Anna* *kal* *viśvavidyālay* *gaye*
Marc and Anna yesterday university go.PAST.MASC.PL
'Marc and Anna went to the university yesterday.'

b. *Anna* *aur* *Pori* *kal* *viśvavidyālay* *gaīṁ*
Anna and Pori yesterday university go.PAST.FEM.PL

‘Anna and Pori went to the university yesterday.’

It is also interesting to observe that when a complex subject is formed by coordinating inanimate nouns, the verb agrees in gender with the nearest noun. As shown in (8), if the masculine noun ‘truck’ is closer to the verb, the verb carries the masculine agreement. However, if the word order is changed and the feminine subject ‘car’ immediately precedes the verb, gender agreement on the verb changes to feminine.

(8) Verbal agreement with masculine and feminine subjects in Hindi

a.	*kār*	*aur*	*ṭrak*	*kal*	*yahāṁ*	*thā*
	car.FEM.SG	and	truck.MASC.SG	yesterday	here	be.PAST.MASC.SG
	‘The car and the truck were here yesterday.’					
b.	*trak*	*aur*	*kār*	*kal*	*yahāṁ*	*thī*
	truck.MASC.SG	and	car.FEM.SG	yesterday	here	be.PAST.FEM.SG
	‘The truck and the car were here yesterday.’					

Generic gender is likewise observed in Pnar, though it operates differently and makes use of the feminine marker.[6] There are several pieces of evidence that the feminine gender is the generic marker, and they are given here without a claim regarding the prominence of a particular feature. First, as mentioned in §4.1, foreign words borrowed into Pnar generally receive the feminine gender marker (i.e. *palong* ‘bed’, *wait* ‘sword, dao’), and as such it seems to serve as a ‘default’. Second, the feminine marker serves as the primary means by which verbs receive a resultative or generic nominalization that can refer to a regularized activity (i.e. *ræp* ‘to farm’ > *ka=ræp* ‘farming, farmer(s)’) – this is also the most common kind of nominalization in Pnar. Third, collective nouns referring to cultural activities (i.e. *balang* ‘(a) gathering, association’ [from the verb *lang* ‘be together’], *jaitbru* ‘tribe, group’; example 9) all receive the feminine gender marker to refer to a particularized group, or the plural marker to refer to multiple members of the group.[7]

(9) Pnar collective nouns with generic reference

a.	*tæ*	*mih*	*ki=ni*	*ki=tnat*	*balang*	*na*
	NVIS	bring.out	PL=PROX	PL=branch	association	ABL
	ka= tæ		*ka=balang*			
	FEM=NVIS		FEM=association			
	‘so these branch associations came out from that association’					
b.	*ka=jaitbru*	*Pnar*	*ka*	*æm*	*ko*	
	FEM=tribe	Pnar	3SG.FEM.ACC	have.exist	3SG.FEM.NOM	
	‘(how) the Pnar tribe existed..’					

Another collective noun *durbar* ‘council’ is borrowed from Indo–Aryan languages, and whereas in the source language (Sanskrit) it refers to a masculine event conceptualized as a location (a ruler’s court), in Pnar the same event and location is marked as feminine. Finally, animals that are unspecified for gender often receive either the neuter marker *i=* or the feminine marker *ka=*, which serves as a generic marker. These features of gender marking suggest that *ka=*

is the generic marker in Pnar.[8] Hence, Pnar displays an opposite pattern to Hindi in terms of which gender marker serves as 'generic'.

4.3 Sociocultural gender in relation to gender reversal

Sociocultural gender may also be echoed by gender reversal. Gender reversal refers to a change of grammatical gender, i.e., a masculine referent may be referred to via a noun with feminine grammatical gender. One of the main functions of such strategies is to ascribe subjective properties to the referent (Contini–Morava & Kilarski, 2013). By way of illustration, in Nepali a speaker may refer to a boy using the feminine grammatical gender to convey affection and vice–versa. Such use of gender agreement commonly occurs in the speech of parents addressing their children. For instance, the parents of a young girl may address her as a young boy to show appreciation and affection towards her. As shown in (10), masculine nouns with masculine agreements can be used with reference to a daughter, while feminine nouns and feminine agreements may make allusion to a son. In both cases, the agreement is marked on the possessive and verbal forms (Pokharel, 2010, p. 46–52). Similar usage is attested in other grammatical gender languages of the world, such as baby–talk of Arabic and Marathi (Ferguson, 1964, p. 106–109).

(10) Gender reversal to show affection in Nepali

a. *mero* *keto* *basyo*
my.MASC boy Sit.PAST.3SG.MASC
'My boy sat down (with reference to a girl).'

b. *meri* *keti* *basi*
my.FEM girl sit.PAST.3SG.FEM
'My girl sat down (with reference to a boy).'

However, such bi–directional use of gender reversal for positive connotations, as in Nepali, is not consistently observed in all languages. Rather, a negative connotation of the feminine gender is more commonly attested in languages of the world (Schulz, 1975; Sautermeister, 1985; Cherry, 1987). In Amharic (Afro–Asiatic), referring to a woman with a masculine pronoun represents admiration, while addressing a man with a feminine pronoun conveys an insult (Aikhenvald, 2012, p. 70). In Modern Hebrew, only women are attested to be addressable via gender reversal for the purpose of attributing subjective properties onto the referent. Women can be referred to by close relatives via the masculine as a sign of affection and intimacy, while gender reversal whereby males are addressed with a feminine marker convey an insult (Tobin, 2001, p. 187).

Such asymmetric distribution emulates sociocultural gender and is analogously found in the Hindi society and language. In Hindi, parents may convey affection by referring to their female child using the opposite grammatical gender of the referent's biological gender, i.e. refer to a girl using the term *beta* 'boy' instead of *beti* 'girl'. Further, the use of a masculine term when referring to women generally elevates the status of the referent (Hall, 2002, p. 145). By way of illustration, if a woman is invoked as *bhai* 'brother', it conveys equality and intimacy. Likewise, if a woman is introduced via the honorific term *sahab*, it represents respect. However, the opposite effect is attested when men are referred to via feminine, such that referring to a man via feminine grammatical gender or associating a man with feminine traits conveys an insult instead.

Interestingly, Pnar does not clearly show this contrast. When referring to children, the neuter/ diminutive marker may be used, and there is generally no negative or positive effect of using the incorrect gender marker for an animate referent except perhaps for humorous purposes. For adults, however, it is possible to use the feminine marker for male referents in a derogatory manner, but not vice–versa. This seems to mirror the effect of gender mismatch in languages like French, for which incorrect categorization results in processing difficulty for the listener rather than extra–linguistic effects such as social derogation of the referent. For Pnar, this may also be due to the use of the feminine *ka=* as a generic or 'default' marker and the availability of the neuter/diminutive *i=*, which allows speakers to make no claims about the gender/sex of a referent.

5 Discussion and conclusion

As a summary, we suggest that grammatical gender paradigms are not only affected by universal cognitive principles but also influenced by cultural idiosyncrasies, perhaps especially with regard to languages that maintain a masculine/feminine distinction affecting the lexicon. Three facets of the two languages Hindi and Pnar were involved in our study: semantics, grammar and pragmatics. Regarding semantics, the association is relatively transparent between semantic features such as size and the masculine/feminine status in the respective sociocultural gender patterns of the speakers' society. Further, the grammatical system itself furnishes additional evidence, as the generic or 'default' gender varies according to the sociocultural gender of the speakers. Likewise, pragmatics–related strategies such as gender reversal seem to mirror sociocultural gender to some degree, since the effect of gender reversal is asymmetric between men and women, though this may itself be a language–specific feature.

In this paper, we have provided a qualitative analysis in terms of grammatical gender and sociocultural gender for two languages. These analyses are expected to serve as a beachhead for further research in terms of human cognition, language typology, and ethnosyntax. Further, we combine observations of linguistic data with a sociocultural dimension. Our approach enhances the application of cross–disciplinary studies in ethnosyntax by pointing out the fascinating patterns that we find at the intersection of sociocultural and linguistic gender. Our findings also provide additional support to functional typology regarding the lexical functions of gender marking in expanding the lexicon, differentiating referents, and ascribing properties. This demonstrates that the functions of grammatical gender are not restricted to reference tracking, but have a variety of uses within language and society. Finally, the extensive set of functions of grammatical gender we have discussed represent multiple reasons for the existence and maintenance of a particular grammatical paradigm.

With regard to limitations and future studies, we acknowledge that in order to fully support a finding that sociocultural gender is reflected in the grammatical gender system of particular languages, information on grammatical gender and sociocultural gender in individual languages needs to be fine–grained enough to fully capture the size/shape categories of the language alongside information on cultural/kinship patterns. A database of such information would enable more comprehensive analysis of lexical semantics to be conducted so as to observe whether such a claim is supported cross-linguistically and cross-culturally. Since the two languages and societies we compared in this paper differ in multiple respects, not the least of which is population size, a future study of the impact of sociocultural gender on grammatical gender could control for such

potentially related factors. A more comprehensive survey should also include a phylogenetically weighted sample of languages with different types of grammatical gender systems, so as to check the statistical correlation between grammatical gender and sociocultural gender.

Last but not least, the interaction of universal cognitive principles and cultural idiosyncracies should also be investigated in terms of diachrony. For instance, Hindi resolved the reduction of three genders to two from older Indo-Aryan, which gives rise to several questions, including: to what degree are the semantic and pragmatic considerations under discussion reflected in gender reassignment? What was the weight of influence from cognitive principles and cultural idiosyncracies during the development of gender systems in Indo-Aryan and Khasian languages in general? It is possible that the distribution of nouns within the gender system of a language may be the historical residue of its evolution, in which case cognitive principles and sociocultural factors only played a minor role. Additional analyses should be conducted to investigate the correlation of grammatical gender and sociocultural gender from a diachronic point of view.

In conclusion, we have shown that sociocultural gender may be reflected by grammatical gender. Two languages of two different societies, Hindi and Pnar, both mark nouns of their lexicon according to masculine and feminine grammatical gender. Yet Hindi and Pnar have different patterns in terms of how sociocultural gender is realized, and could be considered to be near opposite poles of a cultural continuum (in terms of kinship/inheritance) as the two societies are patrilineal and matrilineal respectively. We support the claim of previous studies that certain tendencies are shared across languages and can be ascribed to common cognition, e.g., long shaped objects tend to be masculine and round shaped objects tend to be feminine. At the same time, we show that sociocultural gender is broadly conjointly reflected in the two languages through grammatical gender, i.e. that big objects tend to be masculine and small objects tend to be feminine in Hindi, while the opposite is observed in Pnar. Similar observations are made with regard to generic gender and gender reversal, lending support to the analysis that grammatical gender in each language is closely connected to sociocultural gender in each society, as part of the larger philosophical orientations of the respective speakers.

ABBREVIATIONS

ABL = ablative; ACC = accusative; CLF = classifier; FEM, F = feminine; FUT = future; MASC, M = masculine; NOM = nominative; NVIS = non-visible; PARTIC = participle; PAST = past; PERF = perfective; PL = plural; PRES = present; PROX = proximal; REL = relative; SG = singular; 3SG = third person singular.

COLOPHON

Our work on this paper was fully collaborative; the order of the authors' names is alphabetical and does not reflect any asymmetry in contribution. We wish to acknowledge support from an IDEXLYON (16-IDEX-0005) Fellowship grant (Allassonnière-Tang) as well as an MOE grant (MOE2012- 1-100) at Nanyang Technological University in Singapore (Ring), and an SNSF grant (SNSF100015_176264) at the University of Zürich in Switzerland (Ring) while this research was being carried out. We are also grateful for the fruitful discussion with the audience of the *Workshop on Gender and classifiers: Diachronic and synchronic variation* held in 2016 at the 46th Poznań Linguistic Meeting in Poland. We thank our colleagues and two anonymous referees for

their constructive comments. Special thanks to D. R. Michael Buam for answering specific queries on Pnar. We are fully responsible for any remaining errors.

REFERENCES

Acharya, J. 1991. *A Descriptive Grammar of Nepali and an Analyzed Corpus*. Washington, D.C.: Georgetown University Press.

Agnihotri, R.K. 2007. *Hindi: An Essential Grammar*. Routledge Essential Grammars. London ; New York: Routledge.

Alpher, B. 1987. Feminine as the unmarked grammatical gender: Buffalo girls are no fools. *Australian Journal of Linguistics*, 7, 169–187

Aikhenvald, A.Y. 2012. Round Women and Long Men: Shape, Size, and the Meanings of Gender in New Guinea and Beyond. *Anthropological Linguistics*, 54(1), 33–86.

Aikhenvald, A.Y. 2000. *Classifiers: A Typology of Noun Categorization Devices*. Oxford: Oxford University Press.

Aikhenvald, A.Y. 2016. *How gender shapes the world*. Oxford: Oxford University Press.

Audring, J. 2008. Gender Assignment and Gender Agreement: Evidence from Pronominal Gender Languages. *Morphology,* 18(2), 93–116.

Barlow, M. 1992. *A Situated Theory of Agreement*. New York: Garland.

Barz, R.K., & Diller, A. 1985. Classifiers and Standardization: Some South and South–East Asian Comparisons. In Bradley, D. (Ed.), *Papers in Southeast Asian Linguistics 9* (pp.155–84). Canberra: Pacific Linguistics, the Australian National University.

Boroditsky, L. 2009. How Does Our Language Shape the Way We Think? In Brockman, M. (Ed.), *What's Next? Dispatches on the Future of Science* (pp.116–129). New York: Vintage.

Brown, P. 1980. Why and How Are Women More Polite: Some Evidence from a Mayan Community. In McConnell-Ginet, S., Borker, R., & Furman, N. (Eds.), *Women and Language in Literature and Society* (pp.111–136). New York: Praeger.

Calude, A.S., & Pagel, M. 2011. How Do We Use Language? Shared Patterns in the Frequency of Word Use Across 17 World Languages. *Philos Trans R Soc Lond B Biol Sci,* 366, 1101–1107.

Cameron, D. 1998. *The Feminist Critique of Language: A Reader*. New York: Routledge.

Chakrabarty, B., & Pandey R.K. 2008. *Indian Government and Politics*. New Delhi: Sage.

Chen, Lang., & Rogers, T.T. 2014. Revisiting Domain–General Accounts of Category Specificity in Mind and Brain. *WIREs Cognitive Science,* 5, 327–344.

Chen, S-C. 1997. Sociology of Language. In Hornberger, N.H. & Corson, D. (Eds.), *Encyclopedia of Language and Education: Research Methods in Language and Education* (pp.1–13). Dordrecht: Springer Netherlands.

Cherry, K. 1987. *Womansword: What Japanese Words Say About Women*. Tokyo: Kodansha International.

Clarke, A. 2015. Dynamic Information Processing States Revealed Through Neurocognitive Models of Object Semantics. *Language, Cognition, and Neuroscience*, 30, 409–419.

Clarke, A., & Tyler, L.K. 2015. Understanding What We See: How We Derive Meaning from Vision. *Trends in Cognitive Sciences*, 19, 677–687.

Contini–Morava, E., & Kilarski, M. 2013. Functions of Nominal Classification. *Language Sciences,* 40, 263–299.

Corbett, G.G. 1991. *Gender*. Cambridge: Cambridge University Press.
Corbett, G.G. 2013. Number of Genders. In Dryer, M.S. & Haspelmath, M. (Eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
Deutscher, G. 2010. *Through the Language Glass: Why the World Looks Different in Other Languages*. New York: Metropolitan Books.
Dryer, M.S. & Haspelmath, M. (eds.) 2013. *The World Atlas of Language Structures Online*. Jena: Max Planck Institute for the Science of Human History. (Available online at https://wals.info)
Eckert, P., & McConnell–Ginet, S. 2003. *Language and Gender*. Cambridge: Cambridge University Press.
Edwards, W.F. 1979. Speech Acts in Guyana: Communicating Rituals and Personal Insults. *Journal of Black Studies*, 10(1), 20–39.
Ehrenfels, U.R. 1955. Three Matrilineal Groups of Assam: A Study in Similarities and Differences. *American Anthropologist*, 57(2), 306–21.
Enfield, N.J. (Ed.) 2002a. *Ethnosyntax: Explorations in Grammar and Culture*. Oxford: Oxford University Press.
Enfield, N.J. 2002b. Introduction. In Enfield, N.J. (Ed.). *Ethnosyntax: Explorations in Grammar and Culture*. Oxford: Oxford University Press.
Ferrari, F. 2005. A Syntactic Analysis of the Nominal Systems of Italian and Luganda: How Nouns Can Be Formed in the Syntax. PhD dissertation, New York: New York University.
Fraccaro, P.J., Benedict C.J., Vukovic, J., Smith, F.G., Watkins, C.D., Feinberg, D.R., Little, A.C., & Debruine, L.M. 2011. Experimental Evidence That Women Speak in a Higher Voice Pitch to Men They Find Attractive. *Journal of Evolutionary Psychology,* 9(1), 57–67.
Gabriel, U., & Gygax, P. 2016. Gender and Linguistic Sexism. In Giles, H. & Maass, A. (Eds.), *Advances in Intergroup Communication* (pp.1–26). New York: Peter Lang.
Gainotti, G. 2006. Anatomical, Functional, and Cognitive Determinants of Semantic Memory Disorders. *Neuroscience and Biobehavioral Reviews,* 30, 577–94.
Giddens, A. 2004. *Sociology*. Oxford: Polity Press.
Gneezy, U., Leonard, K.L., & List, J.A. 2009. Gender Differences in Competition: Evidence from a Matrilineal and a Patriarchal Society. *Econometrica,* 77(5), 1637–1664.
Grinevald, C. 2015. Linguistics of Classifiers. In Wright, J.D. (Ed.), *International Encyclopedia of the Social & Behavioral Sciences* (pp.811–818). Oxford: Elsevier.
Gurdon, P.R.T. 1914. *The Khasis*. London: Macmillan; Co.
Hall, K. 2002. 'Unnatural' gender in Hindi. In Hellinger, M. & Bußmann, H. (Eds.), *Gender Across Languages: The Linguistic Representation of Women and Men* (pp.133–162). Amsterdam: John Benjamins.
Hall, K., & Veronica, O.D. 1996. Shifting Gender Positions Among Hindi–Speaking Hijras. In Bergvall, V., Bing, J., & Freed, A. (Eds.), *Rethinking Language and Gender Research: Theory and Practice* (pp.228–266). London: Longman.
Hellinger, M., & Bußmann, H. 2001. The Linguistic Representation of Women and Men. In Hellinger, M. & Bußmann, H. (Eds.), *Gender Across Languages: The Linguistic Representation of Women and Men* (pp.1–25). Amsterdam: John Benjamins.
Hellinger, M., & Motschenbacher, H. 2015. *Gender Across Languages: The Linguistic Representation of Women and Men*, Vol. 4. Amsterdam: John Benjamins.
Holmes, J. 1991. Language and Gender. *Language Teaching,* 24(4), 207–20.

Hymes, D. 1962. The Ethnography of Speaking. In *Anthropology and Human Behaviour* (pp.13–53). Washington, D.C.: Anthropological Society of Washington.
Hymes, D. 1974. *Foundations in Sociolinguistics*. Philadelphia: University of Pennsylvania Press.
Ide, S., Hori, M., Kawasaki, A., Ikuta, S., & Haga, H. 1986. Sex Difference and Politeness in Japanese. *International Journal of the Sociology of Language,* 58, 25–36.
Kachru, Y. 2006. *Hindi*. Amsterdam: John Benjamins.
Keenan, E. 1974. Norm–Makers, Norm–Breakers: Uses of Speech by Men and Women in a Malagasy Community. In Bauman, R. & Sherzer, J. (Eds.), *Explorations in the Ethnography of Speaking* (pp.125–43). Cambridge: Cambridge University Press.
Kemmerer, D. 2017. Categories of Object Concepts Across Languages and Brains: The Relevance of Nominal Classification Systems to Cognitive Neuroscience. *Language, Cognition and Neuroscience,* 32(4), 401–24.
Kirby, K.R., Gray, R.D., Greenhill, S.J., Jordan, F.M., Gomes-Ng, S., Bibiko, H-J., Blasi, D.E., Botero, C.A., Bowern, C., Ember, C.R., Leehr, D., Low, B.S., McCarter, J., Divale, W., & Gavin, M.C. 2016. D-PLACE: A Global Database of Cultural, Linguistic and Environmental Diversity. *PLoS ONE*, 11(7): e0158391. doi:10.1371/journal.pone.0158391.
Koike, D.A. 1986. *Differences and Similarities in Men's and Women's Directives in Carioca Brazilian Portuguese*. Austin: University of Texas Press.
Labov, W. 1972. *Sociolinguistic Patterns*. Oxford: Blackwell.
Lakoff, G., & Johnson, M. 2003. *Metaphors We Live by*. London: The University of Chicago Press.
Lakoff, R. 1973. Language and Woman's Place. *Language in Society,* 2(1), 45–80.
Lakoff, R. 1975. *Language and Woman's Place*. New York: Harper Colophon.
Lew–Williams, C., & Fernald, A. 2007. Young Children Learning Spanish Make Rapid Use of Grammatical Gender in Spoken Word Recognition. *Psychological Science,* 18(3), 193–98.
Lew–Williams, C., & Fernald, A. 2010. Real–Time Processing of Gender–Marked Articles by Native and Non–Native Spanish Speakers. *Journal of Memory and Language,* 63(4), 447–64.
Mahon, B.Z., & Caramazza, A. 2009. Concepts and Categories: A Cognitive Neuropsychological Perspective. *Annual Review of Psychology,* 60, 27–51.
Martin, A. 2007. The Representation of Object Concepts in the Brain. *Annual Review of Psychology,* 58, 25–45.
McConnell–Ginet, S. 2014. Gender and Its Relation to Sex: The Myth of 'Natural' Gender. In Corbett, G.G. (Ed.), *The Expression of Gender* (pp.3–38). Berlin: Mouton de Gruyter.
Nespital, H. 1990. On the Relation of Hindi to Its Regional Dialects. In Offredi, M. (Ed.), *Language Versus Dialect: Linguistic and Literary Essays on Hindi, Tamil and Sarnami* (pp.3–23). New Delhi: Manohar Publications.
Nongbri, T. 1988. Gender and the Khasi Family Structure: Some Implications of the Meghalaya Succession to Self–Acquired Property Act, 1984. *Sociological Bulletin,* 37(1–2), 71–82.
Nongbri, T. 2000. Khasi Women and Matriliny: Transformations in Gender Relations. *Gender, Technology and Development,* 4(3), 359–95.
Pakyntein, V. 2000. Gender Preference in Khasi Society: An Evaluation of Tradition, Change and Continuity. *Indian Anthropologist,* 30(1/2), 27–35.
Pokharel, M.P. 2010. Noun Class Agreement in Nepali. *Kobe Papers in Linguistics,* 7, 40–59.
Preisler, B. 1986. *Linguistic Sex Roles in Conversation*. Berlin: Mouton de Gruyter.
Prewitt–Freilino, J.L., Caswell, T.A., & Laakso, E.K. 2012. The Gendering of Language: A

Comparison of Gender Equality in Countries with Gendered, Natural Gender, and Genderless Languages. *Sex Roles,* 66(3–4), 268–281.
Rabel–Heymann, L. 1977. Gender in Khasi Nouns. *Mon–Khmer Studies*, 247–272.
Ring, H. 2014. Nominalization in Pnar. *The Mon–Khmer Studies Journal,* 43, 16–23.
Ring, H. 2015. A Grammar of Pnar. PhD Disseration, Singapore: Nanyang Technological University.
Sautermeister, C. 1985. The Woman Before Language. *Women's Topics in Foreign Language Teaching: Working Papers*, 63–97.
Scantlebury, K. 2014. Sociocultural Perspectives and Gender. In Gunstone, R. (Ed.), *Encyclopedia of Science Education* (pp.1–2). Dordrecht: Springer.
Schulz, M. 1975. The Semantic Derogation of Women. In Thorne, B. & Henley, N. (Eds.), *Language and Sex: Dominance and Difference* (pp.64–73). Rowley: Newbury House.
Seifart, F. 2010. Nominal Classification. *Language and Linguistics Compass,* 4(8), 719–36.
Shabadi, K. 2005. Speaking Our Gendered Selves: Hinduism and the Indian Woman. In Jule, A. (Ed.), *Gender and the Language of Religion* (pp.257–269). London: Palgrave Macmillan UK.
Sherzer, J. 1987. A Diversity of Voices: Men's and Women's Speech in Ethnographic Perspective. In Philips, S.U., Steele, S. & Tanz, C. (Eds.), *Language, Gender, and Sex in Comparative Perspective* (pp.95–120). Cambridge: Cambridge University Press.
Shukla, S. 2009. "Hindi." In Brown, K. & Ogilvie, S. (Eds.), *Concise Encyclopedia of Languages of the World* (pp.494–497). Amsterdam: Elsevier.
Siewierska, A. 2004. *Person*. Cambridge: Cambridge University Press.
Simon, B.L. 1993. Language Choice, Religion, and Identity in the Banarsi Community. In Hertel, B.R. & Humes, C.A. (Eds.), *Living Banaras: Hindu Religion in Cultural Context* (pp.245–268). Albany: State University of New York Press.
Simon, B.L. 1996. Gender, Language, and Literacy in Banaras, India. In Warner, N., Ahlers, J., Bilmes, L., Oliver, M., Wertheim, S. & Chen, M. (Eds.), *Gender and Belief Systems: Proceedings of the Fourth Berkeley Women and Language Conference* (pp.679–687). Berkeley: Women; Language Group.
Stahlberg, D., Braun, F., Irmen, L., & Sczesny, S. 2007. Representation of the Sexes in Language. In Fiedler, K. (Ed.), *Frontiers of Social Psychology. Social Communication* (pp.163–187). New York: Psychology Press.
Tobin, Y. 2001. Gender Switch in Modern Hebrew. In Hellinger, M. & Bußmann, H. (Eds.), *Gender Across Languages: The Linguistic Representation of Women and Men* (pp.177–198). Amsterdam: John Benjamins.
Zhang, N.N. 2013. *Classifier Structures in Mandarin Chinese*. Berlin: Mouton de Gruyter.

[1] Languages such as English display gender differences for pronouns but not for verbs, e.g. in *he is tall* and *she is tall* the pronouns do change according to masculine/feminine subjects but the verb keeps the same form. This type of grammatical gender system is referred to as pronominal gender (Audring, 2008, p. 96). It is still counted as grammatical gender since the connection between the anaphoric pronoun and its antecedent is analysed as agreement rather than co–reference (Barlow, 1992, p. 134–152; Corbett, 1991; Siewierska, 2004, p. 221–227).

[2] In this and subsequent examples, we use the romanised orthography of the language being exemplified according to conventions found in the literature.

[3] We are aware that in the case of two nouns being affiliated with the same gender, the disambiguation function would not be able to operate via grammatical gender and other cues such as deixis or context would be necessary. However, this does not contradict the statement that disambiguation is one of the functions of grammatical gender, i.e. grammatical gender may fulfill the task of disambiguation but it does not imply that all disambiguation in discourse is realized via grammatical gender in languages which have grammatical gender systems.

[4] Also the neuter proclitic, which will be discussed further below.

[5] It is important to point out that our analysis is limited in its scope of how the big/small contrast is interpreted. Different languages may interpret this contrast in different ways. For instance, some languages may consider 'big' within a vertical dimension while some other languages may consider 'big' from a horizontal perspective. In the current study, we restrict ourselves to the analysis of the big/small contrast from a perspective of volume.

[6] This observation is an additional motivation for choosing Pnar in our analysis, as "the only known languages in which the generic is female, are in some Iroquois languages (Seneca and Oneida), as well as some Australian aboriginal languages" (Prewitt–Freilino, et al., 2012, p. 270; Alpher, 1987). For additional examples, see also Aikhenvald (2016, p. 114)'s discussion on 'Markedness, status, and power in linguistic gender choice'.

[7] Another criterion could be the use of *ka=* for occupations where the gender is unspecified, but there are too few examples of such cases in Pnar corpora to verify this possibility. Yet another criterion could be the use of a gender marker with indefinite pronouns, but in Pnar the indefinite *won* is often used with *bru* 'person' marked either with the feminine, neuter, or plural, and occasionally with the masculine. In these cases, the indefinite seems best translated as 'any' (i.e. *u=won u=won u=bru* 'any man', *ki=won ki=won ki=bru* 'any person'). Future research will help to clarify this point.

[8] Interestingly, similar *ka* forms in other Austroasiatic languages are simply the third person marker, with no gender component.