47

Finnisch-Ugrische Mitteilungen

FINNISCH-UGRISCHE MITTEILUNGEN

Begründet von István Futaky

Band 47

Herausgegeben von

Cornelius Hasselblatt und Beáta Wagner-Nagy
unter redaktioneller Mitarbeit von
Josefina Budzisch

FINNISCH-UGRISCHE MITTEILUNGEN

Herausgeber:
Prof. Dr. Cornelius Hasselblatt
Socrateslaan 13
NL-9404 ER Assen
E-Mail: cornelius@hasselblatt.com

Herausgeberin:
Prof. Dr. Beáta Wagner-Nagy
Universität Hamburg
Fachbereich Sprache, Literatur, Medien II
Institut für Finnougristik/Uralistik
Von-Melle-Park 6, Postfach 29
D-20146 Hamburg
E-Mail: beata.wagner-nagy@uni-hamburg.de
Fax: +49 40 42838-2787

Schriftleitung:
Prof. Dr. Beáta Wagner-Nagy

Wissenschaftlicher Beirat:
Marianne Bakró-Nagy (Budapest)
Andrey Filchenko (Astana)
Riho Grünthal (Helsinki)
Elena Kryukova (Tomsk)
Eva Toulouze (Paris)

Manuskripte und Anfragen, die die Redaktion betreffen, sind an die Herausgeber zu richten. Mit der Annahme eines Manuskripts durch die Redaktion und seiner Veröffentlichung durch den Verlag geht das Verlagsrecht für alle Sprachen und Länder einschließlich des Rechts der fotomechanischen Wiedergabe oder sonstiger Vervielfältigungen an den Verlag über. Die Autor*innen erhalten bei Erscheinen eine PDF-Datei ihres Beitrages.

Für den Inhalt der Beiträge sind allein die Verfasser*innen verantwortlich. Für eingesandte Manuskripte wird keine Haftung übernommen. Unaufgefordert eingesandte Rezensionsexemplare werden nicht zurückgeschickt. Eine Anzeigepflicht besteht nicht, doch sind die Herausgeber um eine zügige Besprechung bemüht. Rezensionsexemplare bitte an die Anschrift der Redaktion (s. o.) senden.

INHALT DES 47. BANDES

Male–female opposition in Mansi

Mátyás Béres (Szeged)

Abstract

The possibility of differentiation between the male and female quality of a denoted object can be found in every language on a certain level. Some languages have grammatical tools as well to denote the biological sex (sexus), meaning that biological sex can be expressed with a separate grammatical category, namely with grammatical gender (genus), whereas in other languages the distinction by sex is restricted to the level of personal pronouns. Most Indo-European languages belong to the first group, excluding English for instance, which can be an obvious example for languages of the second group. However, the role of genus goes beyond semantical functions, because it primarily serves morphosyntactical purposes. The main aim of this paper is to investigate how Mansi, a language without a grammatical gender system, distinguishes between male and female objects. I would like to point out those characteristics that are typical of Mansi, especially its Sosva variant. Even though the concept of genus is unknown to all members of the Uralic language family, different strategies can be found in the individual languages to denote sexus. The examined lexical fields are kinship terms, job titles, and animal names. The sources of data are the Wogulisches Wörterbuch (Mansi Dictionary), the corpora of the Ob-Ugric Database, which contains folklore texts and various newspaper articles from Lūjimā Sēripos.

Keywords: Mansi, grammatical gender, biological sex, word formation, derivation, lexicology, grammaticalization

1 Introduction

There is a way in every language to express biological sex (sexus), namely to distinguish between the male and female quality of a denoted object. In some languages, an explicit way is used, meaning that the differences between the sexes are expressed not only on the level of vocabulary but also via a separate grammatical category (grammatical gender, genus). Therefore, in these languages (e.g. French, Russian, German) grammatical tools are used to make distinctions between the sex of the denoted object. There are also languages, in which the grammatical way of differentiation is present only on the level of personal pronouns (e.g. English), and others that lack the ability to express biological sex on a grammatical level (e.g. Georgian, Chuvash, Hungarian). The latter express gender implicitly. (H. Varga 2016: 307) The aim of this article

is to present a typological overview of grammatical gender in general, which can be considered to be a core element of the linguistic system of languages that have the grammatical category of gender. Furthermore, the main purpose of this paper is to describe how Mansi marks sexus, namely, what kind of strategies can be found in Mansi, especially in its Sosva variant. I would like to point out the characteristics of the language system regarding the different strategies that can be found to denote the biological sex.

2   Genus in the languages of the world

2.1 Definition and the role of genus

In English scientific literature, usually the term *grammatical gender* is used when referring to genus. Nowadays, the word *gender* is frequently associated with the concept of men–women relations, at least in the colloquial use of language. However, it is important to point out that in the linguistic sense, the terms *grammatical gender* and *genus* are not necessarily related to biological sex. In those languages that have a gender system, the term *gender* is only used to refer to word classes of different nominal stems, which are essentially semantic of origin, and these fundamental semantic criteria are not always connected to sexuality. The most substantial factor regarding genus is how a particular nominal stem can belong to a certain gender, as it determines how the stem can agree with other members outside of the nominal phrase (Greenberg 1978: 49).

The generally accepted use of the term *morphological agreement* refers to a function of inflection involving the coordination of sentence constituents, which means that two or more words have to take the same value in relation to a particular morphological category depending on the syntactic environment. Traditionally, eight categories can be identified. Agreement can be based on person, number, case[1], definiteness, class, time, and mood. In the Uralic languages, only four of the aforementioned categories can be found, which are person, number, case, and definiteness. If a language has gender agreement, it works exactly the same way as other categories of agreement. The values of nouns belonging to different genders are taken by certain sentence constituents as well, exactly as they do in other types of agreement.

Besides the division discussed above, two types of morphological agreement can be identified. The first one can tipically occur between a head and a modifier in one syntagma, whereas the other type of agreement takes place between the predicate and the subject (Greenberg 1978: 331–374).

In terms of agreement, grammatical gender functions and works the same as case and number or the previously mentioned categories. The following examples from Russian – which has three genders: masculine, feminine, and

---

[1] Grammatical gender is usually considered a subcategory of class (see Corbett 2013b, Greenberg 1978), but some authors classify it as a category on its own (see H. Varga 2010).

neuter – illustrate agreement in gender between the predicate and the subject. It is important to note that if a Russian sentence has a nominal predicate, gender agreement takes place in every tense, but if the predicate is verbal, the agreement can occur only in the past tense.

(1)   *Žurnal*   *leža-l-Ø*   *na*   *stol-e.*
        newspaper.M   lay-PST-M   LOC   table-LOC
        'The newspaper was on the table.' (Corbett 2013a)

(2)   *Kniga*   *leža-l-a*   *na*   *stol-e.*
        book.F   lay-PST-F   LOC   table-LOC
        'The book was on the table.' (Corbett 2013a)

(3)   *Pis'mo*   *ležal-o*   *na*   *stol-e.*
        letter.N   lay-PST-N   LOC   table-LOC
        'The letter was on the table.' (Corbett 2013a)

Nouns belonging to different genders agree differently with other sentence constituents according to their genus. The verbal stem of the predicate is the same in all three sentences, but the suffix differs depending on the genus of the subject, according to which the masculine gender will be marked with a zero morpheme, the feminine with an *-a,* and the neuter with an *-o*. As in the case of other types of agreements, gender greement can not only occur between the predicate and the subject, but also between the head and the modifier.

(4)   *dlinn-yj*   *žurnal*
        length-ADJZ.M   newspaper
        'long newspaper'[2]

(5)   *dlinn-aja*   *kniga*
        length-ADJZ.F   book
        'long book'

(6)   *dlinn-oe*   *pis'mo*
        length-ADJZ.N   letter
        'long letter'

Morphosyntactically, genus has a key role and its essential feature, according to which genus always implies agreement, is of defining value (Corbett 2013a). A language can be defined as one having a gender system only if a sentence

---

[2] Examples without given source are the author's own.

constituent with a constant genus – namely, the actual gendered noun – agrees with a sentence constituent with a varying genus, which can be a pronoun, an adjective, a numeral, a participle or a verb as well (Fodor 1958: 339). Therefore, a noun belonging to a certain gender has to trigger gender agreement with elements outside of the nominal category as well, otherwise, a language cannot be considered having a gender system. Apart from its role to place the lexemes in the right syntactic environment, genus can also mark sexus. Languages that have a gender system tend to express the biological sex of the denoted object with grammatical tools as well (see in section 3.2).

2.2 The relation between genus and sexus

Besides the fact that genus is not a universal linguistic category, languages that have a gender system do not provide a uniform picture, and genus is not even connected to sexus in some cases.

To illustrate how varied the realization of gender systems can be, let us examine the following examples from German and the previous ones from Russian (cf. examples 1–6). Both languages have three genders (masculine, feminine, and neuter), but while there is gender agreement in Russian between the predicate and the subject, German does not have this feature.

(7)   *Die*   *Zeitung*   *lag*   *auf*
       ART.DEF.F   newspaper   lay.PST   LOC
       dem   *Tisch.*
       ART.DEF.DAT.M   table
       'The newspaper was on the table.'

(8)   *Das*   *Buch*   *lag*   *auf*
       ART.DEF.N   book   lay.PST   LOC
       dem   *Tisch.*
       ART.DEF.DAT.M   table
       'The book was on the table.'

(9)   *Der*   *Brief*   *lag*   *auf*
       ART.DEF.M   letter   lay-PST   LOC
       dem   *Tisch.*
       ART.DEF.DAT.M   table
       'The letter was on the table.'

Although there is no gender agreement between the predicate and the subject in German, there is between the head and the modifier, such as in Russian.

(10) *lang-e*   *Zeitung*
      long-F   newspaper
      'long newspaper'

(11) *lang-es*   *Buch*
      long-N   book
      'long book'

(12) *lang-er*   *Brief*
      long-M   letter
      'long letter'

Genus may appear as an arbitrarily working system, but gender distinctions always have a semantic core, meaning that a noun belonging to a certain genus has an overlap with a certain concept. Although, the degree of this overlap can be different in the languages (Corbett 2013b). The grammatical gender of Russian nouns can be determined based on their stem types. According to the rule, every noun ending with a consonant is masculine because masculine gender is marked by a zero morpheme (e.g. *gorod* 'city', *luk* 'onion', *čas* 'hour'). Every noun ending with -*a* or -*ja* is feminine, and so are most of the nouns that end with a soft consonant (e.g. *mašina* 'car', *sem'ja* 'family', *noč'* 'night'). Neuter nouns usually end with -o, -e or -*mja* (e.g. *pis'mo* 'letter', *more,* 'sea' *imja* 'name'). It is important to point out that these endings apply only for the nominative case of the nouns (Timberlake 2004: 130). However, this rule can be overridden by the biological sex of the denoted object. Words such as *mužčina* 'man' or *deduška* 'grandfather' end with -*a*, which is supposed to mean that they belong to the feminine gender, but instead they belong to the masculine gender grammatically speaking as well, despite their nominative case ending cf. **staraja deduška*, but *staryj deduška*). (H. Varga 2016: 309–310).

All the previously mentioned gender systems are sexus-based, as, besides Russian, German has a masculine-feminine-neuter gender system as well. Division into three – or in the case of some languages, such as Danish and Swedish: two (common and neuter – genders is considered a common feature in most of the Indo-European languages, although some of them completely lack a gender system (e.g. Persian, Greek).

In the Fula language, which is spoken in West Africa, the number of genders is around twenty depending on the dialect. However, male and female referents belong to the same gender, because the semantic core of this gender category covers males and females as well. In Fula, the role of sexus is of much less importance considering genus, as the grammatical classification of nouns has

to do with concepts that do not involve biological sex (Corbett 2013b). Alex Ortiz distinguished 19 different genders in the dialect investigated by him.[3]

Table 1: Grammatical gender system in a dialect of Fula

	Class name	Meaning / Grouping	Example
1	on	Persons and French loanwords	gorko on 'the man'
2	ⁿgen	Various	nage ⁿgen 'the cow'
3	ⁿgun	Various	teu ⁿgun 'the beef'
4	ⁿgon	Various	felo ⁿgon 'the mountain'
5	ⁿgol	Various	cangol ⁿgol 'the river'
6	ⁿgal	Augmentative	lega ⁿgal 'the tree'
7	ⁿgil	Augmentative	nagi ⁿgil 'the (oversized) cow'
8	ⁿdun	Various and animals	sondu ⁿdun 'the bird'
9	ⁿden	Various and round objects	wofo ⁿden 'the egg'
10	ⁿdan	Fluids	ⁿdiaⁿdan 'the water'
11	ⁿdin	Various	ⁿgari ⁿdin 'the (male) cow'
12	kin	Various	leki kin 'the medicine'
13	kun	Diminutive	paɪkun kun 'the (little) boy'
14	kon	Various non-singular nouns	maro kon 'the rice'
15	kan	Various	kafa kan 'the machete'
16	ɓen	Plural persons / loanwords	worɓe ɓen 'the men'
17	ɗin	Plural various	bariɲi ɗin 'the dogs'
18	ɗen	Plural various	bole ɗen 'the snakes'
19	koɪ̂	Plural diminutive	paɪkoɪ koɪ 'the boys'

Thus, the gender system of Fula differs considerably from the concept of genus that can be found in most Indo-European languages. While in the case of Fula words gender markers are clearly visible, in languages where genus marking does not appear on nouns, it is realized only in agreement (Greenberg 1978: 49). However, the most significant difference between gender systems lies within the relation between sexus and genus.

---

[3] Source: https://www.academia.edu/19668989/Nominal_Class_System_of_Fula 04.30 (latest access: 20.10.2023.)
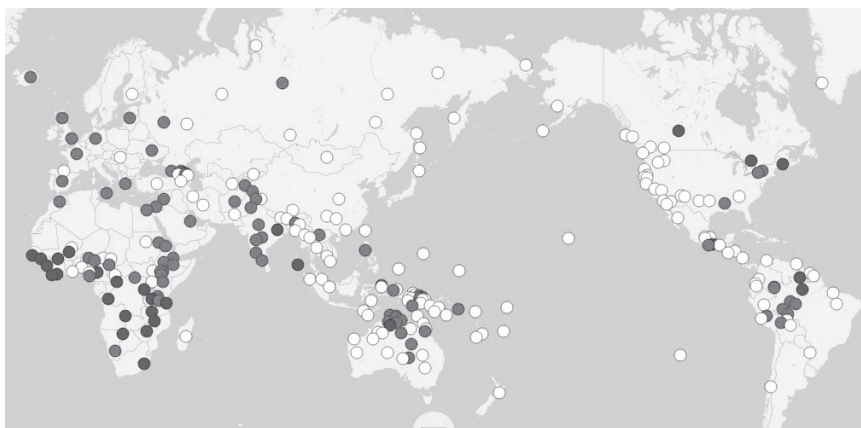
Figure 1: Sex-based and non-sex-based gender systems in the languages of the world (Corbett 2013b)

Despite the fact that languages having a grammatical gender system are quite diverse, gender systems are based on biological sex in most of the languages in general. The relevant chapter of the *World Atlas of Language Structures* (Corbett 2013b) offers concrete statistical data. 112 of the 257 languages examined have a gender system, and exactly three quarters of them, 84 languages, have a system based on sex. In a non-sex-based system, animate-inanimate distinction always has a role (Corbett 2013b), which is of high importance from the point of view of the Indo-European languages as well as their gender systems developed from a similar type of distinction. Historically, the male and neuter gender of the proto-language is a result of an earlier animate-inanimate distinction. The former developed from the animate, the latter from the inanimate category, whereas the feminine gender is a subsequent development that took place after the split of the Anatolian branch of the language family (Lurarghi 2011).

3   Strategies of marking biological sex in the languages of the world

It is important to discuss the strategies from a typological point of view that are used to mark the biological sex of the denoted object. There are two different ways of marking, lexical and grammatical. Both of them can be found in languages with or without a gender system, but most of the grammatical tools can be found only in such languages that express gender explicitly.

3.1 Lexical tools

To mark the biological sex of an object, three types of lexical tools can be found. A language can use lexeme pairs, adjective phrases, and compound words. Amongst these types, lexeme pairs can be considered fundamental be-

cause such lexemes already have the semantic features that denote the sexus of the object (H. Varga 2016: 300), therefore there is no need to add any other elements to specify it.

Distinction by sex in kinship terms is typically carried out by lexeme pairs, as the following examples show: English *uncle ~ aunt*, Italian *il padre* 'father' ~ *la madre* 'mother'. This kind of distinction usually occurs in forms of address as well: German *der Herr* 'sir' ~ *die Frau* 'madame' but lexeme pairs can also be used to distinguish between animals: Russian *петух* 'cock' ~ *курица* 'chicken', English *bull ~ cow*. The use of lexeme pairs can be explained by functional reason as their occurrence is restricted to beings of high importance in the life of a community. In the case of humans, this means that if a role in the society is traditionally associated with certain gender values in the colloquial sense of the word, then two separate lexemes are used to denote men and women: Hungarian *férfi* 'man' ~ *nő* 'woman', *apa* 'father' ~ *anya* 'mother'. In the case of animals, lexeme pairs primarily appear referring to livestock because different agricultural functions are the result of biological sex (H. Varga 2016: 300–301). The high importance of functionality and role is proved by the fact that castrated animals and those that are unable to reproduce have a different role than male or female companions, and thereby different lexemes are used when referring to them: German *der Wallach* 'gelding', English *mule*, *wether*.

The second way of marking is to use adjective phrases. They are usually used with words that have a so-called common gender. Words with common gender can be used to refer to both sexes but they belong to one certain grammatical gender, if the language has a gender system. Both Latin *parēnt* 'parent' and Russian *родитель* 'parent' are grammatically masculine but they can be used to refer to a father or a mother as well, just as in English for instance. Genus, thus, does not determine the denotation. If one wants to specify the meaning by sex, a modifier can be used, for example *männlicher Wolf* 'male wolf' ~ *weiblicher Wolf* 'female wolf' (H. Varga 2016: 301).

The third lexical way of marking is the use of compound words. Both the first and the last component of a compound word can denote the sexus, e.g. German *der Küchenjunge* 'scullion' ~ *das Küchenmädchen* 'kitchenmaid', Russian *ženŝina-vrač* 'female doctor'. The component that defines sexus can derive from a personal pronoun – which is usually in the third person singular –, e.g. *she-wolf* or from a lexeme meaning male or female, e.g. *cock-pigeon* or from a proper name, e.g. *tomcat* (H. Varga 2016: 301).

If there is no sexual dimorphism in certain species, not even in the slightest degree, meaning that the male and female individuals look similar, the same lexeme is used for both sexes: English *wolf*, German *der Fuchs* 'fox', just as in the case of animals at a taxonomically lower position: Romanian *cârtiţă* 'mole', English *butterfly* (H. Varga 2016: 304).

3.2 Grammatical tools

There are also gramamtical tools available to mark the biological sex of the denoted object. Derivational suffixes are commonly used to express the biological sex of an object in languages that are explicit considering gender expressions as well as in languages that are implicit. English *-ess* belongs to these kinds of suffixes, e.g. *god ~ goddess*, such as the German *-in* of the same function, e.g. *Sänger* 'singer' ~ *Sängerin* 'female singer', *Hund* 'dog' ~ *Hündin* 'bitch'.

Some nouns can switch their gender in some languages, meaning that the article and hereby the genus changes as well, marking this way the male or female quality of the denoted object, e.g. Italian *il maestro* 'teacher' ~ *la maestra* 'female teacher', *il gatto* 'tomcat' ~ *la gatta* 'queen' (H. Varga 2016: 305–306).

4 Concepts of marking

4.1 Morphological marking

Morphological markedness is based on the presence or absence of a particular element of form. The opposition between the male and the female form is expressed by a certain morpheme. The previously mentioned *-ess* derivational suffix and all the other ways of marking discussed above could be good examples of that. By using *-ess* as an example, the word *hostess* can be considered morphologically marked because of the *-ess* suffix, which its male equivalent *host* lacks and thus it can be considered unmarked (Lyons 1979: 305–307).

4.2 Distributional marking

Usually, morphologically marked forms have a limited distribution. This correlation is the second important factor in marking concepts because the restricted distribution of morphologically marked forms is independent of the morphological marking itself, as the following example presents. The first member of the *lion ~ lioness* lexeme pair can refer to the whole panthera leo species and a male individual of that species as well, whereas the other member can be used only to refer to female individuals. For this reason, *male lion* and *female lion* are both acceptable collocations, whilst *male lioness* is a contradiction and *female lioness* is a tautological expression. That is not the case with *count* and *countess* for instance, although the morphological structure is the same as in the previous example. *Female count* and *male countess* are contradictory, whereas *male count* and *female countess* are tautological expressions. Of course, from a morphological point of view the word *count* is unmarked such as the word *lion*, and the word *countess* is marked such as the word *lioness* but in the case of the *lion ~ lioness* lexeme pair the opposition is marked not only by morphology but also by the different distribution of the two words. Therefore distribution has to be treated as a distinct concept because there are morphologically marked words that are not marked by distribution, e.g.

countess opposed to *count*, and there are such lexemes that are not marked by morphology, only by distribution, e.g. the negative members of the following word pairs: *good ~ bad*, *high ~ low*. In this case the low level of distribution shows markedness (Lyons 1979: 305–307).

4.3 Semantic markedness

Semantic marking frequently correlates with distributional marking, sometimes it is even labeled as its defining feature. A lexeme is qualified as semantically marked if it has a more specific connotation and therefore more semantic features than the corresponding, semantically unmarked form. The second member of the previously mentioned *lion ~ lioness* word pair has a more specific, more concrete meaning than the lexeme *lion*, thus it can be considered as a marked form because of the earlier described semantic reasons. The difference between morphological and semantic marking is well illustrated by the *dog ~ bitch* lexeme pair, where *bitch* relates to *lioness* from a semantic point of view exactly the same, as *dog* relates to *lion*, whereas from a morphological point of view *bitch* and *dog* are not related to each other, the lexeme *bitch* is not derived from the lexeme *dog*, but *lioness* is derived from the word *lion* (Lyons 1979: 307). In fact, one might ask if this phenomenon was a case of polysemy, meaning that, for example, the word *lion* has two meanings, the first one being 'panthera leo' and the second one being 'a male individual of the panthera leo species'. Likewise in the case of the word *dog*: 'canis lupus familiaris' and 'a male individual of the canis lupus familiaris species'. This would be an appropriate explanation of why these words are superordinates and antonyms of the semantically marked female equivalents at the same time. However, since such dual use is not a unique feature in English nor in many other languages, but it is a widespread phenomenon in the whole lexicon, it should be considered a corollary of semantic marking, not polysemy (Lyons 1979: 308).

4.4 The unmarked base form

Most of the previous examples show that in most cases from two words that denote sex the one denoting the male conterpart is the unmarked base form and the words denoting the female counterpart is formed from that (Huszár 2009: 37).

In general, it can be stated that the unmarkedness of a certain form denoting a male object and, therefore, the markedness of a form denoting the female one is more common in the languages of the world. The phenomenon can be even observed outside of the nominal category, where biological sex has no role at all. The following example shows how the marked feminine predicate is formed from the masculine form.

(13) *Žurnal*   *leža-l-Ø*   *na*   *stol-e.*
      newspaper.M   lay-PST-M   LOC   table-LOC
      'The newspaper was on the table.' (Corbett 2013a)

(14) *Kniga*   *leža-l-a*   *na*   *stol-e.*
      book.F   lay-PST-F   LOC   table-LOC
      'The book was on the table.' (Corbett 2013a)

The concept of generic masculine is mostly used in the context of languages with grammatical gender. It can be described as a grammatically masculine[4] lexeme with a general meaning that can cover both sexes, e.g. German *der Arzt* 'doctor', English *teacher*. However, words that can be labeled as generic masculine are used mostly to refer to male objects (Huszár 2009: 34, H. Varga 2016: 301) and the phenomenon can also be found in languages that do not have a gender system. For instance, most of the job titles in Hungarian are gender-neutral, but by hearing a high-prestige title, e.g. *igazgató* 'director, headmaster'; ügyvéd 'lawyer' or *képviselő* 'representative' one would associate these titles with a male person, and by hearing a low-prestige job title, e.g. *ápoló* 'nurse' or *takarító* 'cleaner' one would associate the titles with women. In general, the fact that the form denoting the male counterpart is the base form and the one denoting the female counterpart is formed from these, means that the latter is a subsequent development caused by changes in society.[5] It is important to point out that the use of forms denoting female counterparts can be typologically heterogeneous as well. For instance, in Hungarian, these forms are used only if the female quality of the denoted object is of higher importance. In lexemes, such as *katonanő* 'female soldier', *hegesztőnő* 'female welder', or *vadásznő* 'female hunter', the intention of expressing the female quality of the person who occupies the certain job is strong. If this piece of information is not relevant, gender-neutral forms are used. On the other hand, in most Indo-European languages, the female form does not have key information regarding the significance of sex (H. Varga 2016: 302). In German, female-denoting words are used to refer to women, even if the sex of the denoted object is not important, e.g. *die Lehrerin* 'female teacher'.

Not only the concept of generic masculine is known but also the concept of generic feminine. This phenomenon is less common, for instance, in German, it is only used in animal names. *Die Spinne* 'spider' and *die Katze* 'cat' are

---

[4] Or in languages that are implicit by gender: formally masculine.

[5] According to H. Varga, this can be meant especially for languages that have a gender system, but, in my opinion, it can be considered universal.

both grammatically feminine but they can be used to refer to male individuals of the species as well.[6]

It is considered rare, but in some languages, the forms denoting the male are morphologically marked and formed from the one denoting the female, which is considered the base form. In Yanyuva, which is spoken in Australia, and in the language of the Bolivian natives, the Chiquitas, the base form denotes the female (Huszár 2009: 37). It is also a typologically rare phenomenon, when both the male and the female quality are expressed with a component of a compound word, which is otherwise not the base word. This is a frequent strategy in Finnish, in which *mies-* and *nais-* are used as first components (*mieslääkäri* 'male doctor' ~ *naislääkäri* 'female doctor') and *-mies* and *-nainen* as last components (*tiedemies* 'male scientist' ~ *tiedenainen* 'female scientist'). Notwithstanding, not only this symmetric use is present in Finnish as some job titles are formed from a word that originally denotes men, e.g. *lehtimies* ('journal' + 'man') 'male journalist' ~ *lehtimiesnainen* (*lehti* 'journal' + *mies* 'man + *nainen* 'woman') 'female journalist' (H. Varga 2016: 303).

5   Marking of male-female opposition in Mansi

5.1 The Mansi language and society

The Mansi people – or the formerly used exonym: Voguls – live in North-West Siberia in the territory of the Khanty-Mansi Autonomous Okrug, traditionally in the area surrounded by the Ob river and the Ural mountains. Their language belongs to the Ugric group of the Finno-Ugric branch of the Uralic language family, and together with Khanty, the two languages form the Ob-Ugric languages. The first Mansi language records appear in the 16th century in Russian Chronicles, but dedicated research in linguistics and ethnography only began in the 19th century.

At the time when the Mansi language became a topic of research, it had four living dialects: northern, western, southern, and eastern. Nowadays, only the northern dialect is spoken, the Sosva variant being its most widely spoken subdialect. The Mansi literary language is based on this subdialect as well (Kálmán 1963: 7–12). The representation of the language is low. According to the latest census,[7] only 12,228 persons identified as ethnic Mansis, and only 1,346 persons declared themselves native speakers of the language. The language of public education is Russian, and there are only ten rural elementary schools that offer Mansi language classes (Horváth 2020: 70). On the other hand, there is a regular radio broadcast in Mansi, and Lūjimā Sēripos ('northern

---

[6] Source: https://www.goethe.de/ins/hu/hu/kul/sup/klt/21458969.html (latest access: 30.08.2023.)

[7] Source: https://rosstat.gov.ru/vpn_popul (latest access: 01.03.2022.)

light'), the only newspaper in Mansi, is also regularly published, the physical copies being only in Mansi.

The marking of male-female opposition in a language is based on socio-cultural factors as well. These social and cultural factors must be defined in the case of Mansi too, because they strongly affect the use of the language. Traditionally, exogamy was a well-known concept amongst the Ob-Ugric people, which can be defined as a custom of marrying outside of their own phratry (clan). Levirate marriage was a known tradition as well, meaning that a man was able to marry the widow of his deceased brother or the sister of his own wife. The third marriage custom in the Ob-Ugric culture is the sororate marriage, which is actually the opposite of levirate marriage, this is which allows women to marry the brother of their deceased husband or the husband of their sister. These customs are important because they affect how the ego – the self that functions as a reference-point – can name a certain relative or a group of relatives (Bíró 2004: 33–34).

The Mansi kinship system is based on belonging to a certain clan. Because of exogamy the mother's and the wife's original clan can be considered foreign and therefore less important than the own clan or, in the case of women, the husband's clan (Bíró 2004: 36). This results in certain taboos in the society as well (Sárkány 1989: 346). An important part of the traditional Mansi way of living is the bear cult, which is also accompanied by many taboos. One of them is taboo language that – whilst serving euphemistic purposes – led to lexical differences in the language use of men and women (Laakso 2005: 78).

5.2 Research method

The sources of data are the Wogulisches Wörterbuch (Mansi Dictionary) – edited by Béla Kálmán –, which contains all the words that can be found in the folklore texts collected by Bernát Munkácsi (Munkácsi 1986), and the corpora of the Ob-Ugric Database,[8] which contains folklore texts collected by Kannisto, Rombandeeva and Chernetsov, moreover various newspaper articles from Lūjimā Sēripos.[9] In addition to that, a questionnaire was sent to native speakers with the aim of gaining more knowledge about the realization of male–female opposition in job titles. The informants were two women, both intellectuals, middle-aged and Mansi–Russian bilinguals. Their task was to assign a job title to 75 pictures. 25 occupations were shown, and the aim of the task was to explore how gender-neutral, male and female job titles are expressed. Besides job titles, five other terms were part of the questionnaire: *friend, widow, guest, host, Frenchman*. All the language data (unless it is indicated otherwise) are

---

[8] Source: https://www.babel.gwi.uni-muenchen.de/index.php?abfrage=NM_corpus&subnavi=corpus_pub (latest access: 10.10.2023.)

[9] Source: https://khanty-yasang.ru/luima-seripos (latest access: 10.10.2023.)

from the northern dialect. The Mansi data will be analyzed based on the features discussed in section 3.1.

5.3 Lexeme pairs

Male and female kinship terms are typically lexeme pairs. In Mansi, the previously explained sociocultural factors and the importance of a certain relative are both guiding principles for the ego to refer to an individual family member (Bíró 2004: 33–34). These are important because of marriage customs as the Mansi language expresses potential family relations that result from these customs. Also, as explained above, the father's clan can be considered more close to the ego due to exogamy and this cultural notion manifests in the language too.

Distinction by gender is a significant feature in Mansi, just as in the other Ugric languages (Bíró 2004: 34). As a result, the system of kinship terms can be considered a symmetric one. Terms for the immediate family members are clearly separated: *āś* 'father' ~ *sań* 'mother', *piγ* 'son' ~ *āγi* 'daughter'. It is important to note that in Mansi, *piγ* can mean both 'boy' and 'son', and so does *āγi*, which means 'girl' and 'daughter' as well, meaning that there is no distinction between these two meanings as in English for instance.

In terminology, siblingship overlaps paternal cousinship, which is the result of the previously mentioned potential family relations. If the father of the ego marries his deceased brother's wife, the paternal uncle's children can be labeled not only as cousins but siblings as well. That's the reason, why *kaŋk* 'elder brother', *āpśi* 'younger brother', *uwśi* 'elder sister', and *ēś* 'younger sister' can be used as terms for cousinship, although they are originally used to mark the sex and relative age of siblings only. The words *pānt* 'nephew', 'male cousin' and *ēŋk* 'niece', 'female cousin' can denote not only the children of the paternal aunt but the children of the ego's sister as well. Once again, this is the case of potential kinship relations being part of the meaning. If the ego is a female, then she is a potential wife of her sister's husband, meaning that her nephews and nieces are her potential stepchildren. This is the reason why *pānt* and *ēŋk* can also bear the meaning 'stepbrother' and 'stepsister' (Bíró 2004: 40). These additional meanings provide a widespread use of these lexemes, making them distributionally unmarked as opposed to their symmetrical equivalents on the maternal side of cousinship terms, as presented in the table below.