

*Ksenia Shagal, Merit Niinemägi,
Luan Hammer, Kitti Vojter, Mátyás Béres*

Gendered job titles in genderless languages: A case of Finno-Ugric

I. Introduction

One of the grammatical features shared by all Finno-Ugric languages is that they lack grammatical gender. That is, they do not show verbal and nominal agreement in gender, they lack gender distinctions in pronouns, etc. There is, however, a lexical domain in which gender distinctions are often expressed in a fairly regular and systematic way, namely, job titles. Indeed, many, possibly all, Finno-Ugric languages have gendered variants of job titles that form minimal pairs, such as *kuningas* ‘king’ – *kuningatar* ‘queen’ in Finnish, or *színész* ‘actor’ – *színésznő* ‘actress’ in Hungarian.

Despite the lack of gender distinctions in the grammar, most gender distinctions in the lexicon are expressed by means of the native language material, such as the derivational suffix *-tar* in Finnic languages or the noun *nő* ‘woman’ in Hungarian illustrated in the examples above. However, some of the available distinctions are the result of massive borrowings from languages with grammatical gender. For instance, Meadow Mari has borrowed separate words for ‘salesman’ and ‘saleswoman’ from Russian, *продавец* and *продащица/продащице*, as well as many other gendered job titles.

This article presents a pilot study of the ways in which languages without grammatical gender can express gender in lexical items. It focuses on six Finno-Ugric languages – three national languages (Estonian, Finnish, and Hungarian) and three smaller languages (Aanaar Saami, Meadow Mari, and Northern Mansi) – that differ in their genealogical affiliation, contact history, and sociolinguistic status. We investigate data from native speakers and available corpora in order to discover all the strategies that these languages use for expressing gender in job titles, reveal patterns in their use and formulate generalizations concerning their distribution.

A variety of earlier studies have addressed the representation of gender in Hungarian (e.g., Pete 2000, Nagy & Patti 2006, Kegyes 2008, Vasvári 2015), Finnish (e.g., Tainio 2001, Engelberg 2002, Tainio 2006), and Estonian (e.g., Ross 1996, Olt 2004, Puna 2005, Hasselblatt 2015). Aanaar Saami, Meadow Mari and Northern Mansi, on the other hand, have so far not been subject to research on the expression of gender. An overview of strategies that these languages use to express gender in job titles is, therefore, one of the major contributions of this paper. In addition, the paper aims to contribute to developing connections between Finno-Ugric studies and gender linguistics as outlined in Laakso (2011).

The paper is organized as follows: Section 2 presents our methodology and gives an overview of the data used in the study. In Section 3, we discuss seven strategies of gender expression that we were able to find in the languages under investigation. In Section 4, we summarize and discuss our findings and formulate potential questions for further studies.

2. Methodology and data

Since this is a pilot study, we aimed at using a variety of methods, which would allow us to capture the diversity of the existing strategies. In order to give the process some structure and make it more uniform across the languages, we compiled a list of jobs to serve as a starting point. The list contained 25 jobs which met the following conditions: (a) they were common enough for the respective titles to be familiar to native speakers and to be well represented in the available corpora; (b) they differed in their status, that is, in how often they are normally carried out by men and women, and in whether they are typically seen as masculine or feminine.

These 25 jobs were then used as a basis for a questionnaire. For each job, we found three photos, two of which showed a woman and a man doing the job, and the third featured a person whose sex was unclear (e.g., a person shown from the back) or somehow not emphasized.¹ In

¹ The choice of photos that we could use for the questionnaire was limited to those under a Creative Commons license or those that we could legally obtain from the copyright owners. Due to this, not all of the pictures completely

order to avoid interference with other languages, we did not use translation equivalents of job titles in the questionnaire and had descriptions of the jobs instead. For example, the description we used to elicit possible expressions for ‘teacher’ was ‘a person whose job is teaching, especially in a school’ (the descriptions were mostly based on Lea & Bradbery 2020). For each job, the respondents were asked to list all the words and expressions that they could use to refer to this job in general and all the words and expressions that they could use to refer to a woman and a man doing the job. Additionally, we included five concepts that are not job titles but are likely to have regularly expressed gendered variants, namely, female and male friend, female and male guest, hostess and host, widow and widower, Frenchwoman and Frenchman. The questionnaire was available in two languages, English and Russian, so the speakers could choose one according to their language skills. It was implemented as a Google Form that potential respondents could access online. Both versions of the questionnaire are available at tinyurl.com/fujobs.

As we did not intend to perform any statistical analysis based on the data we received from the speakers, our goal was simply to collect at least a couple of responses per language, so that we could get a general idea of what strategies are available. In addition to the members of our team who speak some Finno-Ugric languages themselves, we received responses from a number of native speakers: five speakers of Finnish, five speakers of Hungarian, three speakers of Mari, two speakers of Northern Mansi, and two speakers of Estonian. In addition, we consulted one native speaker of Aanaar Saami who did not fill in the online questionnaire.

Once we identified the strategies available in each language, we carried out a corpus analysis to investigate how common specific strategies are and how they behave in actual language use. The corpora we used in our study are the following:

matched the description above, and some of them belonged to a cultural context that was different from what the speakers of the target languages might be used to. Nevertheless, judging by the responses we received, the jobs were still recognizable to most speakers.

Estonian: Estonian Web corpus (etTenTen19, www.sketchengine.eu/ettenten-estonian-corpus/); 508 million words crawled and processed from September 2019 to January 2020.

Finnish: Corpus of online forum discussions Suomi 24 (urn.fi/urn:nbn:fi:lb-2020021803); 4.1 billion words from 2001–2017.

Hungarian: Magyar Nemzeti Szövegtár (MNSZ2, clara.nytud.hu/mnsz2-dev/); over 1 billion words of texts from several registers and different varieties of Hungarian (see Oravecz et al. 2014).

Aanaar Saami: Text corpus provided by the SIKOR corpus at UiT The Arctic University of Norway (gtweb.uit.no/korp/); about 1.77 million words, primarily from articles published in periodical publications.

Meadow Mari: Corpus of Literary Mari (corpus.mari-language.com), specifically the Meadow Mari corpus; 57.38 million words from non-fiction texts, fiction texts, law texts, science texts, news texts and Wikipedia texts from 1912–2018.

Northern Mansi: Corpus of the Ob-Ugric Database (OUIDB, www.babel.gwi.uni-muenchen.de/index.php?abfrage=NM_corpus&subnavi=corpus_pub); Northern Mansi section offers 137 texts (the exact number of words is unknown), including mainly folklore texts and several newspaper articles.

As is clear from the descriptions above, the corpora we used differ greatly in size and content, so the data on different languages are not directly comparable. However, they provide a first glimpse into how common different strategies are within each language and allow us to formulate preliminary generalizations and further questions. Due to some technical difficulties in the use of the corpora, such as the impossibility to eliminate homonymy, some of the figures given below should be seen as estimates rather than exact numbers. Nevertheless, we believe that they still reflect the bigger picture and are accurate enough to support our observations.

In the corpus part of our study, we primarily investigated the same jobs that were used for the questionnaire, but we did include some further job titles if they seemed more representative (e.g., occurred more frequently in a specific corpus) or worked better as an illustration of a certain point.

3. Strategies of expressing gender in job titles

In this section, we provide an overview of the strategies that are used in the languages of our sample to express gender in job titles. We primarily aim to focus on strategies that work at the level of single lexemes, including compounds that can be lexicalized to a different extent. However, since it is a widely accepted fact that formulating a cross-linguistically applicable definition of lexeme may be problematic, we mean to be rather inclusive than restrictive in determining our scope. One thing that we do not include at this point are expressions where gender is expressed by a specialized attribute referring to gender, such as *naispuolinen* ‘female’ and *miespuolinen* ‘male’ in Finnish, or *naissoost/meessoost* ‘of male/female gender’ in Estonian.

In some of the descriptions below, we use the term “gender-neutral job title”, which requires a brief clarification. Since Finno-Ugric languages lack grammatical gender, it is natural that all of them feature many job titles that can be used for a person of any gender. Many of these job titles are inherently gender neutral, that is, they do not contain any elements that could be seen as indicative of gender, e.g., *opiskelija* ‘student’ in Finnish is an agent nominalization from *opiskella* ‘to study’, and *туныктышо* ‘teacher’ in Meadow Mari is a present participle of *туныкташ* ‘to teach’. Nevertheless, it is also common that a job title containing an inherently masculine element is used gender-neutrally, e.g., *kalamees* ‘fisherman (lit. fish+man)’ in Estonian. The use of inherently (at least historically) feminine variants as gender neutral is fairly marginal but also attested, e.g., Finnish *karjakko* ‘milkmaid’, Estonian *õde* ‘nurse (lit. sister)’ or the same in Hungarian: *nővér* ‘nurse (lit. older sister)’. All these cases count as gender-neutral job titles for the purposes of this study.

3.1. The gender-neutral job title takes a gendered prefix to refer to either gender

In three languages, Finnish, Estonian, and Aanaar Saami, gender-neutral job titles can take gendered prefixes derived from the words ‘woman’ and ‘man’ to form gendered job titles. The respective prefixes are *nais-* and *mies-* in Finnish, e.g., *naisopiskelija* ‘female student’ and *miesopiskelija*

'male student', *nais-* and *mees-* in Estonian, e.g., *naisnäitleja* 'actress' and *meesnäitleja* 'male actor', and *nissoon-* and *almal-* in Aanaar Saami, e.g., *nissoonkokkâ* 'female chef' and *almalkokkâ* 'male chef'. According to our consultant, in Aanaar Saami, these prefixes can be used freely with various job titles. However, the strategy is not well represented in the corpus, and the only two examples apart from those mentioned above are *nissoonpresident* 'female president' and *nissoonproffheet* 'female prophet', each occurring only once (although the small number of examples might be due to the comparatively small size of the Aanaar Saami corpus). In Finnish and Estonian, on the other hand, the strategy is attested with most job titles and is overall very frequent, although job titles differ greatly as to how frequently these prefixes are used at all and how frequently each of them (feminine or masculine) is used. In what follows, we will address these two issues in turn.

In order to calculate the frequency of a job title's gendered variants, we divided the total number of their occurrences (both feminine and masculine) by the total number of all occurrences (gender neutral as well as gendered formed by prefixes). In Finnish, two words stand out in this respect, namely *pappi* 'priest' and *näyttelijä* 'actor'. For these job titles, 5.54% and 3.85% of all the occurrences are gendered. For a handful of other job titles, gendered versions constitute roughly 1–1.5%, namely *siivooja* 'cleaner' (1.51%), *opettaja* 'teacher' (1.07%), *hoitaja* 'nurse' (1.06%), and *johtaja* 'boss, director' (0.97%). Interestingly, the job titles with a relatively high share of gendered *nais-/mies-* versions almost exactly correspond to the jobs that have widely used female *-tar* versions discussed in Section 3.2.1. The exceptions are *siivooja* 'cleaner', which has 1.51% of gendered *nais-/mies-* forms but for which the form *siivoojatar* is not very common (66 occurrences in our data), and *myyjä* 'salesperson', for which the form *myyjätär* is relatively common (2 636 hits), but the gendered *nais-/mies-* forms constitute only 0.81% of all instances.² This neat correspondence probably shows that these jobs (at least in Finnish) are generally prone to specify the gender of people doing them.

In Estonian, for most job titles, the respective percentages of gendered variants are below 1%, with two exceptions: *näitleja* 'actor'

² This relatively small share can perhaps be explained by the fact that the word *myyjä* is also widely used to refer to non-human sellers, such as companies.

(3.03%) and *lendur* ‘pilot’ (2.11%). Besides gendered variants with *nais-* and *mees-*, *näitleja* ‘actor’ can also take the feminine suffixes *-nna* and *-tar*: *näitlejanna* and *näitlejatar* ‘actress’ (see Section 3.2.1). The total occurrences of the four gendered variants (*meesnäitleja*, *naisnäitleja*, *näitlejanna* and *näitlejatar*) make up 17.09% of all occurrences of *näitleja* ‘actor’. This percentage is much higher than for any other job title, which shows that speakers of Estonian are more prone to indicate the gender of an actor than of any other professional included in our study. In the case of *lendur* ‘pilot’, however, *naislendur* ‘female pilot’ was the only gendered variant attested in the corpus, while *meeslendur* ‘male pilot’ did not occur at all.

To take a closer look at the ratio of job titles with feminine and masculine prefixes, we considered 19 job titles with more than 50 occurrences in Finnish and 20 job titles with more than 10 occurrences in Estonian. In Finnish, these job titles split into three more or less equal groups based on how common each of the gendered variants is. On the one hand, there are six job titles for which the female variant is much more common (female/male ratio from 7 to 12.5), namely *pappi* ‘priest’, *paimen* ‘shepherd’ (predominately in religious contexts), *sotilas* and *varusmies* ‘soldier’, *lentäjä* ‘pilot’, and *metsästäjä* ‘hunter’. These are jobs that are often seen as stereotypically male and/or are usually done by men. On the other hand, there are five jobs where the male variant is more common (female/male ratio under 1), namely *kättilö* ‘midwife’, *hoitaja* ‘nurse’, *siivooja* ‘cleaner’, *palvelija* ‘servant’, *kokki* ‘cook, chef’. These are jobs that are often seen as stereotypically female and/or are usually done by women. For the remaining eight job titles, the female variant is more frequent, but the ratio is not as high (from 1 to 4). These job titles are *näyttelijä* ‘actor/actress’, *myyjä* ‘salesperson’, *opiskelija* ‘student’, *opettaja* ‘teacher’, *lääkäri* ‘doctor, physician’, *johtaja* and *esimies*, both ‘boss, director’ (for the latter two the female/male ratio is close to the upper value, i.e., around 4). These jobs are probably seen as relatively neutral, and they are often done by both men and women.

In Estonian, with three of the 20 job titles, only the feminine prefix *nais-* was used, while variants with the prefix *mees-* were not attested. These three job titles were *kirikuõpetaja* ‘priest’ and the synonyms *piloot* and *lendur* ‘pilot’. For three more job titles, the female/male ratio was

very high (between 9 and 15): *jahimees* 'hunter', *sõdur* 'soldier' and *tuletõrjuja* 'firefighter'. Like in Finnish, the high female/male ratio seems to correlate with the profession being viewed as typically male. In contrast, a low female/male ratio is found with jobs which could be considered typically female, such as *lapsehoidja* 'nanny' (0.1) and *hooldaja* 'caretaker' (0.3). For seven jobs, the female/male ratio was between 0.6 and 1.3. This means that one of the two variants was generally more frequent, but the difference was not as big. For the professions *õpetaja* 'teacher', *müüja* 'seller' and *arst* 'doctor', the female/male ratio was between 0.6 and 0.76, i.e., the male prefix was used slightly more often. For the professions *näitleja* 'actor', *juhataja* 'chairman; manager; head (of an institution)', *direktor* 'director; headmaster' and *ülemus* 'boss', the ratio was between 1.1 and 1.3, i.e., the female prefix was used slightly more often. In addition, there were five jobs with a relatively high female/male ratio (between 2.2 and 7.8). These were *tudeng* 'university student' (2.2), *koristaja* 'cleaner' (4), *juht* 'leader; manager; driver' (5), *vaimulik* 'cleric' (5) and *üliõpilane* 'university student' (7.8). In a nutshell, these numbers resemble the female/male ratios attested for the same job titles in Finnish, with one exception. With the Finnish word *siivooja* 'cleaner', the male variant with *mies-* was more often used than the female variant with *nais-*. As indicated by the female/male ratio of 4, the opposite was the case for its Estonian counterpart, *koristaja* 'cleaner'. For reasons yet to be determined, some (partly) synonymous words behave differently in respect to these prefixes, e.g., *tudeng* and *üliõpilane* both translate as 'university student', yet the latter has a significantly higher female/male ratio.

A generalization that can be formulated based on the data from both languages is that speakers are likely to specify the gender if it breaks a stereotype regarding a specific job (e.g., a female soldier or a male midwife), while in neutral contexts they more often feel the need to specify that they are talking about a woman doing a specific job, while male interpretation is seen as default.

3.2. The job title has a feminine suffix to refer to a woman doing the job

3.2.1. The job title takes a (productive) feminine suffix

In two of the languages, Estonian and Finnish, female versions of job titles can be formed using a more or less productive feminine suffix. In this case, the title without a suffix can be either gender neutral, e.g., *müüja* ‘seller’ in Estonian or *johtaja* ‘director, manager’ in Finnish, or masculine, e.g., *kuningas* ‘king’ in both Estonian and Finnish.

In Estonian, two suffixes can be used to form the female variants of job titles: *-nna* and *-tar*. The list below provides an overview of the job titles for which the female variants with *-nna* and *-tar* appeared at least 10 times in the corpus. For each of these job titles, we counted the occurrences of all three variants (the neutral/masculine form as well as the female variants with *-nna* and *-tar*) and calculated the percentage of each variant’s occurrences.

<i>kuningas</i> ‘king’	39 216 (79.98%)	<i>kuninganna</i>	9 805 (20.00%)
		<i>kuningatar</i>	14 (0.02%)
<i>näitleja</i> ‘actor’	43 495 (85.12%)	<i>näitlejanna</i>	5 906 (11.56%)
		<i>näitlejatar</i>	1 701 (3.32%)
<i>õmbleja</i> ‘sewer’	1 604 (94.46%)	<i>õmblejanna</i>	94 (5.54%)
<i>preester</i> ‘priest’	7 910 (97.88%)	<i>preestrinna</i>	131 (1.78%)
		<i>preestritar</i>	25 (0.34%)
<i>müüja</i> ‘seller’	59 143 (99.36%)	<i>müüjanna</i>	314 (0.53%)
		<i>müüjatar</i>	66 (0.11%)
<i>õpetaja</i> ‘teacher’	152 311 (99.78%)	<i>õpetajanna</i>	338 (0.22%)

The female variant with *-nna* was by far the most used in the case of *kuningas* ‘king, male regent’ > *kuninganna* ‘queen, female regent’. The explanation for this might be twofold. Firstly, the use of female suffixes (*-nna* and *-tar*) is the only strategy used with the word *kuningas*. Secondly, as mentioned above, unlike most other job titles, *kuningas* is not gender neutral, i.e., it is only used to refer to a king, not a queen.

For all of the job titles listed above, the female variant with *-nna* was more often used than the variant with *-tar*. The suffix *-nna* mostly occurred with job titles that end in the derivational suffix *-ja*. In addition to the job titles above, this also included scarce occurrences of *juhatajanna* ‘female manager’, *koristajanna* ‘female cleaner’, *hooldajanna*

'female caretaker', *lүpsjanna* 'female milker', and *kirikuօpetajanna* 'female priest'. The suffix *-tar*, on the other hand, typically occurs with stems that do not end in a derivational suffix. In addition to the four words listed above, we attested two more female job titles with *-tar* in the corpus: *kokatar* 'female chef' and *pilooditar* 'female pilot'.

If the original word refers to a member of the aristocracy, such as *kuningas* 'king', then the suffix *-nna* is commonly used for their wife and the suffix *-tar* for their daughter, i.e., *kuninganna* 'queen' and *kuningatar* 'princess'. However, both *kuninganna* and *kuningatar* can also be used to refer to a female regent, that is the female counterpart of *kuningas* 'male regent' (see also Erelt et al. 2007: SM22). Whether or not there is a semantic difference in the case of other job titles, such as *näitlejanna* and *näitlejatar* 'actress', must be left to future research.

In Finnish, the only currently productive feminine suffix used in job titles is *-tar/-tär* (for a detailed overview of Finnish feminine suffixes including also *-kko/-kkö*, *-ska/-skä* and *-nna*, see Kyrölä 1989, 1990). The list of job titles for which feminine variants with *-tar/-tär* are most common (among those with more than 500 occurrences in the corpus) is provided below.

<i>kuningas</i> 'king'	190 882 (93.13%)	<i>kuningatar</i>	14 081 (6.87%)
<i>näyttelijä</i> 'actor'	70 954 (97.70%)	<i>näyttelijätär</i>	1 668 (2.30%)
<i>pappi</i> 'priest'	319 877 (98.96%)	<i>papitar</i>	3 371 (1.04%)
<i>myyjä</i> 'seller'	388 182 (99.33%)	<i>myyjätär</i>	2 636 (0.67%)
<i>hoitaja</i> 'nurse'	166 338 (99.58%)	<i>hoitajatar</i>	703 (0.42%)
<i>johtaja</i> 'director'	323 758 (99.67%)	<i>johtajatar</i>	1 064 (0.33%)
<i>opettaja</i> 'teacher'	442 815 (99.70%)	<i>opettajatar</i>	1 330 (0.30%)

Other female variants with *-tar/-tär*, which occur more rarely, include, for example, *siivoojatar* 'female cleaner', *ompeelijatar* 'female sewer', *metsästäjätär* 'female hunter', and *lääkäritär* 'female doctor'.

As can be seen, the lists for Finnish and Estonian are fairly similar, even though not all of the job titles are cognates. In particular, the words meaning 'female priest' are derived from different stems but are relatively common in both languages, which reflects a general preference to express gender with this occupation (see also Sections 3.1 and 3.6).

3.2.2. The job title is borrowed with a feminine suffix

Several languages feature female job titles that contain feminine suffixes not productive in the respective languages. These are the cases where both the gender-neutral/masculine and the female variant can easily be identified as loanwords. A clear example attested in several languages and featuring basically the same suffix is ‘deacon’ > ‘deaconess’: *diakon* > *diakoniss* in Estonian, *diakónus* > *diakonissza* in Hungarian, and *diakoni* > *diakonissa* in Finnish. Among the jobs on our list, the same suffix is also found in the words *direktriss* in Estonian and in the fairly marginal word *direktrisz* in Hungarian (35 occurrences in our data) meaning ‘female director’. Another example of a word borrowed with a feminine suffix is *kollegina* ‘female colleague’ in Hungarian, a loanword from German. In Aanaar Saami, a borrowed feminine suffix *-târ* can be attested in the word pair *kunâgâs* ‘king’ and *kunâgâtâr* ‘queen’ (cf. Finnish *kuningas* ‘king’ and *kuningatar* ‘queen’, where *-tar* is a productive feminine suffix as discussed in Section 3.2.1).

For the aforementioned languages, this strategy is fairly marginal, but for another one of the investigated languages, Meadow Mari, it is the main way to express gender in job titles. Basically, it seems that in Meadow Mari, the general concept of using gendered job titles is mostly a result of Russian influence. Gender is rarely expressed in native job titles, but for those borrowed from Russian, often both the male and the female version (usually adapted to Meadow Mari phonology) are borrowed, with the male version typically being used to refer to the job in general; see examples below.

<i>санитар</i> ‘paramedic’	352 (63.20%)	<i>санитарка/-ке</i>	205 (36.80%)
<i>актёр</i> ‘actor’	1 554 (69.87%)	<i>актриса/-се</i>	670 (30.13%)
<i>продавец</i> ‘seller’	220 (70.51%)	<i>продащица/-це</i>	92 (29.49%)
<i>слуга</i> ‘servant’	161 (84.23%)	<i>служанка/-ке</i>	30 (15.77%)
<i>почтальон</i> ‘mailman’	646 (88.61%)	<i>почтальонка/-ко</i>	83 (11.39%)
<i>учитель</i> ‘teacher’	1 570 (91.17%)	<i>учительница/-це</i>	152 (08.83%)
<i>студент</i> ‘student’	6 686 (95.24%)	<i>студентка/-ке</i>	334 (04.76%)
<i>лётчик</i> ‘pilot’	719 (98.76%)	<i>лётчица/-це</i>	9 (01.24%)

Even though the male forms are often perceived as default by the Mari speakers, they are still not very likely to be used to refer to female representatives of these jobs. Therefore, the frequency of the gendered ver-

sions to a large extent depends on how often specific jobs are actually performed by women. For instance, for 'pilot', a traditionally male job, the percentage of occurrences of the feminine versions *лётчица/лётчице* is relatively low. On the other hand, for *дояр(ке)* 'milker', *акушер(ке)* 'midwife' and *уборщик/уборщице* 'cleaner', traditionally female jobs, the feminine version is actually the one more often encountered in the corpus. Moreover, *акушерке* 'midwife' was also mentioned as the neutral and male form in the questionnaire by one of the respondents.

<i>дояр</i> 'milker'	136 (09.40%)	<i>доярке</i>	1 312 (90.60%)
<i>акушер</i> 'midwife (male)'	51 (38.64%)	<i>акушерке</i>	81 (61.36%)
<i>уборщик</i> 'cleaner'	7 (14.89%)	<i>уборщица/-це</i>	40 (85.11%)

Interestingly, if certain Russian borrowings in Meadow Mari also allow for other strategies of gender expression, the feminine version borrowed from Russian is in general found more often in the corpus than other options, e.g., *учительница/учительнице* 'female teacher' (gendered borrowing) – 152 occurrences, *учитель ўдыр* 'lit. teacher girl' – 2, *ўдырамаш учитель* 'lit. woman teacher' – 2; *лётчица/лётчице* 'female pilot' (gendered borrowing) – 9 occurrences, *лётчик ўдыр* – 1 'lit. pilot girl', *ўдырамаш лётчик* 'lit. woman pilot' – 1.

3.3. The job title contains a gendered part that can be replaced to refer to another gender

In several languages in our sample, some of the job titles contain a gendered part, typically the one meaning 'man', like in Finnish *palomies* 'fireman', Estonian *jahimees* 'hunter', or Aanaar Saami *suátialmai* 'soldier'. (Among these jobs, many are indeed seen as typically done by men.) Although these job titles are usually perceived as gender neutral rather than gendered, it is usually possible to replace the gendered part they contain by a different one in order to refer to the opposite gender. For example, in Finnish, all these job titles ending in *-mies* '-man' have a feminine version, although they are much less frequent, with feminine/masculine ratio between 0.0025 and 0.0066:³

³ We removed the word pair *kalamies* → *kalanainen* 'fisherman' from the analysis, because in many cases these words were used to refer to a person

<i>postimies</i> 'postman'	2 130	<i>postinainen</i>	14	F/M ratio 0.0066
<i>varusmies</i> 'soldier'	23 378	<i>varusnainen</i>	111	F/M ratio 0.0047
<i>esimies</i> 'boss'	88 047	<i>esinainen</i>	369	F/M ratio 0.0042
<i>palomies</i> 'fireman'	17 897	<i>palonainen</i>	65	F/M ratio 0.0036
<i>sotamies</i> 'soldier'	7 469	<i>sotanainen</i>	19	F/M ratio 0.0025

In Estonian, the picture is similar. For the following job titles the corpus data includes instances in which the gendered part *-mees* is replaced by its female counterpart *-naine* 'woman':

<i>jahimees</i> 'hunter'	11 766	<i>jahinaine</i>	423	F/M ratio 0.0360
<i>esimees</i> 'chairman'	91 466	<i>esinaine</i>	2 758	F/M ratio 0.0302
<i>kalamees</i> 'fisherman'	5 146	<i>kalanaine</i>	16	F/M ratio 0.0031

As can be seen, the feminine/masculine ratio for *kalamees/kalanaine* is much lower than for the other two jobs. A possible explanation for this might be the existence of at least two gender-neutral terms for fisherman, *kalur* and *kalastaja* 'fisher', which do not contain an inherently masculine part and are relatively frequent (5074 and 954 occurrences in the corpus, respectively).

As mentioned earlier, there is good evidence that these job titles are usually seen as gender neutral by native speakers despite containing an overtly masculine element. For instance, both Finnish and Estonian speakers who filled in our questionnaire used the words containing the part *-mies/-mees* to refer both to the job in general and to men and women doing the job. In fact, the respective words with the part *-nainen/-naine* were hardly ever suggested, even to refer to women. The same tendency is observed in the corpus data: judging from the context, words containing the masculine part are widely used with female referents. For example, Kaja Kallas, the current (female) prime minister of Estonia and head of the Estonian Reform Party, is referred to as *esimees* 'chairman, leader' 527 times in our data. This is quite in line with the recommendations of the Estonian Language Institute (Eesti Keele Instituut;

whose astrological sign is Pisces. This is particularly typical of the word *kalanainen*, which explains why it is relatively frequent if compared to the masculine counterpart (*kalamies* 6 962, *kalanainen* 1 411; ratio 0.20). We did not encounter the same problem with the Estonian data, so the respective words in Estonian are analysed normally.

see, e.g., AMSS⁴), which state that *esinaine* ‘female chairman, leader’ should only be used if a woman is the leader of an entirely female organization, association or society as illustrated by example (1). However, the actual use is not entirely in accordance with this recommendation. Even though the Estonian Reform Party has male and female members, Kaja Kallas is sometimes referred to as the *esinaine* of the party, as in example (2) and 87 other examples.

- (1) *Kas sa tead, mis on Naiskodukaitse esinaine nimi?*
 ‘Do you know who’s the **(female) leader** of the Women’s voluntary defence organization?’ (etTenTen19)
- (2) *Reformierakond valis laupäeval Tallinnas toimunud üldkogul erakonna esinaiseks Kaja Kallase 1067 poolthäälega.*
 ‘At their general assembly in Tallinn last Saturday, the Estonian Reform Party chose Kaja Kallas as their **(female) leader** with 1067 votes in favour.’ (etTenTen19)

Another reason to see the words ending in *-mies/-mees* as gender neutral is that they can be modified using other strategies in order to form gendered job titles. For example, in the Finnish corpus, the words *naisesimies* ‘female superior, boss’, *naisvarusmies* ‘female soldier’ and *naispalomies* ‘female firefighter’ had 213, 84 and 15 occurrences, respectively. Moreover, the word *miesesimies* ‘male superior, boss’, which is formally tautological, is attested 55 times; see example (3).

- (3) *Lisäksi, kun yrityshierarkiassa mennään ylöspäin, niin miesesimiehiähän siellä alkaa tulla vastaan.*
 ‘In addition, when you go up the company hierarchy, you start to encounter **male bosses** there.’ (Suomi24)

Despite the gender neutrality of the job titles ending in *-mies* ‘man’, in Finnish, for some of these jobs there is yet another gender-neutral option containing the part *-henkilö* ‘person’ instead of *-mies*. There has been quite a lot of discussion around these variants in terms of gender equality in recent years (e.g., www.helsinginuutiset.fi/paikalliset/1805575), and yet, in the corpus from 2001 to 2017 surveyed for this study, they are

⁴ Ametniku soovitusõnastik (www.eki.ee/dict/ametnik/).

roughly as frequent or even less frequent than the feminine forms: *esihenkilö* ‘boss’ – 114 occurrences, *varushenkilö* ‘soldier’ – 69, *palohenkilö* ‘firefighter’ – 35, *postihenkilö* ‘postman’ – 16, *sotahenkilö* ‘soldier’ – 13.

In Aanaar Saami, on the other hand, the inherently neutral version is more common than the feminine one. There is only one job title in our data in which the gendered part *-almái* ‘man’ is replaced by its feminine counterpart, *-nissoon* ‘woman’. This job title is ‘public officer’, which occurs 34 times as *virgealmái* ‘lit. office man’, 29 times as *virgeolmooš* ‘lit. office person’ and only two times as *virgenissoon* ‘lit. office woman’. For another job title, *ovdâalmái* ‘foreman’, the variant with *-olmooš* ‘person’ (*ovdâolmooš*) was even slightly more frequent (10 occurrences) than the variant with *-almái* ‘man’ (8 occurrences). For two other job titles ending in *-almái* ‘man’ found in our data, *puâlualmái* ‘firefighter’ and *suâtialmái* ‘soldier’, no other variant was attested.

Besides job titles ending in *-mees* ‘man’, in Estonian we also find professions made of gendered kinship terms: *(med)õde* ‘nurse’⁵ from *õde* ‘sister’, and *ämmaemand* ‘midwife’, which is a compound of *ämm* ‘mother-in-law’ and *emand* ‘mistress, lady’ originating from *ema* ‘mother’. In Finnish, a cognate word does not appear in job titles but is present in the word *emäntä* ‘hostess’ whose masculine counterpart *isäntä* ‘host’ is derived from the word *isä* ‘father’.

In the Estonian corpus, there are some occurrences where the latter part of the compounds *medõde* and *ämmaemand* is replaced by its male counterparts, i.e., *medvend* ‘male nurse’ from *vend* ‘brother’ and *ämmaisand* ‘male midwife’ from *isand* ‘master, lord’ (< *isa* ‘father’). However, the consulted Estonian native speakers considered both general job titles, *(med)õde* and *ämmaemand*, to be gender neutral and commonly used for men, too. Moreover, the corpus data (etTenTen19) suggest that the use of the gender-specific male counterparts *medvend* and *ämmaisand* is scarce: *medvend* appeared 39 times in the corpus (as opposed to 887 occurrences of *medõde*) and *ämmaisand* only three times (as opposed to 4 146 occurrences of *ämmaemand*).

⁵ While *õde* is the most common word for ‘nurse’, it can also refer to the homonymous kinship term ‘sister’. For that reason, the corpus search focused on the unambiguous term *medõde* instead.

In the corpus, we found no occurrences in which both parts of the compound *ämmaemand* ‘midwife (mother-in-law+mistress)’ would be substituted by their male counterparts. An Internet search did give very few results in which *äiaisand* (from *äi* ‘father-in-law’ and *isand* ‘master’) appeared. However, as illustrated by example (4) below, the word *äiaisand* typically occurred in discussions regarding the job title of a male midwife.

- (4) A: *Kui seda tööd teeb mees, kas siis on tema ametinimetus ämmaisand?*
 ‘If this work is done by a man, is the title of his profession then **male midwife (lit. mother-in-law + master)?**
- B: *Ei, äiaisand ikka!*
 ‘No, of course it’s **male midwife (lit. father-in-law+master)!**
 (foorum.naistekas.delfi.ee/)

The strategy outlined above is also marginally present in Meadow Mari. The word *медшўжар* ‘nurse’, which the native speakers provided as a gender-neutral job title, contains the part *-шўжар* meaning ‘younger sister’. One of our respondents mentioned *медуза* with a part *-уза* meaning ‘older brother’ as the male counterpart, although examples for this (or < *шольо* ‘younger brother’) are not attested in the corpus. The corresponding Russian borrowing *медбрат* (< *брат* ‘brother’) occurred once in the corpus, while the Russian borrowing *медсестра* (< *сестра* ‘sister’) had 445 occurrences. Meanwhile there are 354 occurrences of *медшўжар*, and also two occurrences of *медака* (< *ака* ‘older sister’).

3.4. The gender-neutral job title forms a compound with a gendered noun

All the languages considered in our study can create gendered job titles by adding a gendered noun to a gender-neutral job title. However, the languages differ in how common and regular this strategy is. For Hungarian, for example, it can be seen as the primary strategy. The gender-neutral forms of Hungarian job titles do not contain a gendered part, that is, such words as *tanár* ‘teacher’ or *szakács* ‘cook’ can be used for both men and women. However, by adding the word *nő* ‘woman’ to them it is possible to create a feminine variant.

The use of different job titles ending in *-nő* reflects some of the tendencies that we described for Estonian and Finnish in Sections 3.1 and 3.2.1. In general, the frequency of the feminine version of a job title in Hungarian depends rather on how often the respective job is actually performed by women than on its image of typically female or male. For example, traditionally female jobs, such as ‘housekeeper’ or ‘midwife’ show a high (> 1) female/male ratio, that is, their feminine versions are more common than the neutral ones and are not perceived as marked. Contrarily, for traditionally male jobs, such as ‘soldier’ or ‘hunter’, feminine variants are rarely used, and if they are, they put a significant emphasis on the gender of the word and might even seem unnatural to native speakers. Among the jobs that do not have clear gender preference, ‘actor/actress’ is the one for which a gendered version (feminine in this case) is particularly common.

<i>házvezető</i> ‘housekeeper’	210	<i>házvezetőnő</i>	798	F/M ratio 3.8
<i>szülész</i> ‘male midwife’	822	<i>szülésznő</i>	1 522	F/M ratio 1.8516
<i>színész</i> ‘actor’	99 318	<i>színésznő</i>	46 211	F/M ratio 0.4653
<i>doktor</i> ‘physician’	32 595	<i>doktornő</i>	3 840	F/M ratio 0.1178
<i>tanár</i> ‘teacher’	122 981	<i>tanárnő</i>	10 184	F/M ratio 0.0828
<i>katona</i> ‘soldier’	127 111	<i>katonanő</i>	374	F/M ratio 0.0029
<i>vadász</i> ‘hunter’	12 955	<i>vadásznő</i>	25	F/M ratio 0.0019
<i>pilóta</i> ‘pilot’	46 091	<i>pilótanő</i>	221	F/M ratio 0.0048

Nő is the most frequent component of these compounds in Hungarian, but it is not the only one. Instead, speakers can also use the word *lány* ‘girl’ in some of the job titles. *Diáklány* ‘girl student’ was mentioned by the respondents, and it also has a relatively high number of occurrences in the corpus (3 692). Gendered words with similar meanings are also occasionally used to form compounds in other languages, for example *neiu* ‘maiden’, *tüdruk* ‘girl’ and *poiss* ‘boy’ in Estonian, as in *õpetajaneiu* ‘lit. teacher+maiden’, *müüjatüdruk* ‘lit. seller+girl’ and *kokapoiss* ‘lit. chef+boy’, *neiti* ‘young lady’, *tyttö* ‘girl’ and *poika* ‘boy’ in Finnish, as in *tarjoilijaneiti* ‘waiter+young lady’ *opiskelijatyttö* ‘lit. student+girl’ and *paimenpoika* ‘shepherd+boy’, and *nieidâ* ‘girl, daughter’ in Aanaar Saami, as in *kävppijâsnieidâ* ‘seller girl’ or *puâsuikeččeenieidâ* ‘reindeer herder girl’. In many cases, these words are used to stress the young age of the respective person, as in example (5) below from Estonian.

- (5) *Seal olid kaks nooremapoolset meessoost tohtripoissi.*
 ‘There were two youngish male **physicians (lit. doctor+boys).**’
 (etTenTen19)

Other words that frequently appear in such compounds are those meaning ‘aunt’ and ‘uncle’. In colloquial Estonian, the words *tädi* ‘aunt’ and *onu* ‘uncle’ are commonly used instead of *naine* ‘woman’ and *mees* ‘man’ to refer to an unknown adult person. This is especially typical for (but not limited to) the communication with small children. Both, *tädi* and *onu* are also frequently used in compound nouns, amongst others with common professions. In the Estonian Web Corpus (etTenTen19), the following compounds with *-onu* and *-tädi* could be attested:

<i>arstitädi</i> ‘physician+aunt’	201	<i>arstionu</i> ‘physician+uncle’	70
<i>kokatädi</i> ‘chef+aunt’	191	<i>kokaonu</i> ‘chef+uncle’	13
<i>koristajatädi</i> ‘cleaner+aunt’	83	<i>koristajaonu</i> ‘cleaner+uncle’	4
<i>müüjatädi</i> ‘seller+aunt’	50	<i>müüjaonu</i> ‘seller+uncle’	5
<i>tohtritädi</i> ‘doctor+aunt’	5	<i>tohtrionu</i> ‘doctor+uncle’	26
<i>õmblejatädi</i> ‘sewer+aunt’	1	-	

Just like the words *tädi* ‘aunt’ and *onu* ‘uncle’ by themselves, the compounds listed above are often used by children or in contexts that involve children, e.g., when addressing children (6) or when talking about a person who works at a kindergarten or school (7).

- (6) *Tule kooli arstionuna või arstitädina või hoopis doktorina.*
 ‘Come to school as a **male physician (lit. physician+uncle)** or **female physician (lit. physician+aunt)** or rather as a **doctor.**’
 (etTenTen19; originally a call for pupils to dress up as a physician for the school’s “Health Day”)
- (7) *Muidugi kõik lasteaias teadsid, et neil ei ole tööl mitte kokatädi, vaid kokaonu, aga mina seda ilmselgelt ei teadnud ja sellest aru ei saanud!*
 ‘Of course everybody in the kindergarten knew that they did not have a **female chef (lit. chef+aunt)** working there, but a **male chef (lit. chef+uncle)**, but I clearly did not know nor understand that.’ (etTenTen19)

The corresponding words can also function in a similar way in Hungarian and Finnish. In Hungarian, the expressions *doktor néni*, *doktor bácsi*,

'female/male doctor' and *tanító néni*, *tanító bácsi* 'female/male teacher' featuring the words *néni* 'older lady, aunt' and *bácsi* 'sir, uncle' are very common in contexts where children are addressed. In Finnish, the words *täti* 'aunt' and *setä* 'uncle' can form compounds with various job titles, such as, for instance, *opettajatäti* 'lit. teacher+aunt', *lääkäri-setä* 'lit. doctor+uncle', and occasionally even *esimiestäti* 'lit. boss+aunt' or *palomies-setä* 'lit. fireman+uncle', where the original job title already contains a gendered part (see Section 3.3).

In Hungarian, Estonian, Finnish and Aanaar Saami, many of the gendered compounds described above are lexicalized, or at least the order of the elements in them is fixed. In Meadow Mari, on the other hand, this strategy, which was the one most frequently used in the questionnaire, is more flexible. The gendered noun (usually *ўдырамаш* 'woman', *пӧръен* 'man', *ўдыр* 'girl', or *рвезе* 'boy') can be placed either before or after the job title, e.g., *ўдырамаш туныктышо* 'lit. woman teacher', *пӧръен туныктышо* 'lit. man teacher', *ургызо ўдыр* 'lit. tailor girl', *ургызо рвезе* 'lit. tailor boy'. Though usually *ўдырамаш* and *пӧръен* are placed before the job title and *ўдыр* and *рвезе* after the job title, also gender-neutral job titles followed by *ўдырамаш* or *пӧръен* were mentioned by the respondents, e.g., *эмлыше пӧръен* 'doctor man', *ургызо пӧръен* 'tailor man', or *сарзе ўдырамаш* 'soldier woman'. The respondents also mentioned one word (not exactly a job title) with which other gendered nouns are used, namely *оза* 'host, landlord' > *озавате* 'lit. host+wife' (*озавате* was also mentioned as the neutral version by one of the speakers). In the corpus, also *озакува* 'lit. host+old woman' and its male counterpart *озакугыза* 'lit. host+old man' are attested.

In the corpus, this strategy is not very common – most of the gendered forms are attested less than 10 times, and for most of the job titles the gendered forms make less than 1% of all the occurrences (gender-neutral versions and gendered versions in total). Exceptions are *новар* 'cook' (2.2%), *тунемше* 'student' (2.9%), *кўтўчӧ* 'herder' (3.28%), *шолтышо* 'cook' (6.4%), and *оза* 'host, landlord' (12.29%). The high frequency of gendered versions for *оза* might be partly explained by the fact that *озавате*, *озакува* and *озакугыза* appear to be relatively fixed terms (compared to other gendered words in our study), which is also reflected in them being traditionally spelled as single words. This also holds true

for *тулар* ‘matchmaker’ and its gendered terms *туларвате/тулар вате* ‘lit. matchmaker+wife’ and *тулармарий/тулар марий* ‘lit. matchmaker+husband’ (18.12% of the total occurrences are gendered). However, not all the gendered expressions that can be written in one word are that common. For example, for the word *тан* ‘friend, lover’, the gendered versions *танўдыр/тан ўдыр* ‘lit. friend+girl’, *танвате/тан вате* ‘friend+wife’, *ўдырамаш тан* ‘lit. woman+friend’, *тан рвезе* ‘lit. friend+boy’, and *танмарий/тан марий* ‘lit. friend+husband’ constitute only 0.58% of all the occurrences, and for *йолташ* ‘friend’, the versions *йолташўдыр/йолташ ўдыр* ‘lit. friend girl’, *йолташрвезе/йолташ рвезе* ‘lit. friend boy’ make up only 3.82%.

For most job titles for which this strategy is used, the female versions can be found in the corpus about as often as the male versions. The only exceptions are *туныктышо* ‘teacher’, *повар* and *шолтышо* ‘cook’ and *оза* ‘host’, which favour feminine versions, and *кўтўчө* ‘herder’, which is more common with the masculine noun *рвезе* ‘boy’, as shown below.

<i>туныктышо</i> ‘teacher’	<i>туныктышо ўдыр</i>	88
	<i>туныктышо пөръен</i>	2
	<i>ўдырамаш туныктышо</i>	12
<i>повар</i> ‘cook’	<i>пөръен повар</i>	10
	<i>повар ўдыр</i>	13
<i>шолтышо</i> ‘cook’	<i>шолтышо ўдыр</i>	8
<i>кўтўчө</i> ‘herder’	<i>кўтўчө ўдыр</i>	8
	<i>кўтўчө рвезе</i>	30
	<i>озавате</i>	2 316
<i>оза</i> ‘host’	<i>озакугыза</i>	44
	<i>озакува</i>	157

As mentioned earlier, for most of these words, gendered variants are generally more common in Meadow Mari than for others. In other words, if speakers of Meadow Mari often specify gender with a specific job title, there is also a relatively clear tendency as to which gender is specified.

Finally, this strategy is also occasionally used in Northern Mansi, but is mainly restricted to loanwords, such as *χōn* ‘king, tsar’, *kūśaj* ‘host, boss’, or *pūp* ‘priest’, which have a general meaning when used on their own. In gendered compounds, these word forms are followed by a gendered noun, which is primarily either *nē* ‘woman’ or *χum* ‘man’, but in

some cases also *ēkwa* ‘woman, elderly lady’ or *ōjka* ‘old man’, e.g., *pūp ēkwa* ‘female priest, nun’, *pūp ōjka* ‘male priest, monk’,⁶ *pirikaska nē* ‘female seller’, *χōn nē* ‘queen’.

3.5. The first part of a compound is related to the job and the second is a gendered noun

In the strategy considered in the previous section, gendered nouns are attached to existing job titles in order to form gendered versions. In this section, we address the cases where gendered nouns are an essential part of job titles, that is, where they combine with a job-related expression to form an inherently gendered job title. This is the most widely used strategy in Northern Mansi. Job titles which use this strategy are built up by more or less the same method. If there is an end product or an object of the working act, it stands in the first place, and it is followed by a present participle. This first part of the title is then always followed by a noun, which is predominantly *χōtpa* ‘person’ or *māχam* ‘people’ to refer to the job in general. For example, *tēnut pājtnē χōtpa* ‘chef, cook’ can be analyzed as *tēnut* ‘food’ + *pājtnē* ‘cooking’ + *χōtpa* ‘person’. Instead of *māχam* and *χōtpa*, gender-specific nouns can also be used, as in *χul alāšlan χum* ‘fisherman (lit. fish-catching man)’ or *χanti lātəη χańiştan nē* ‘female Khanty language teacher (lit. Khanty-language-teaching woman)’. The nouns which are most likely to be found as a gender-specifier are *nē* ‘woman’ and *χum* ‘man’, although *āyi* ‘girl’ and *piγ* ‘boy’ are also often used when the job title describes a young person, e.g., *kolledžit χańiştaxtən āyi* ‘female student (lit. in a college + learning + girl)’ or (*armijat*) *lusitan piγ* ‘male soldier (lit. [in the army] + serving + boy)’.

Interestingly, such gendered versions can be used by native speakers even if a completely different word exists for the job in general. For instance, one of our respondents used the word *saltat* ‘soldier’ (< Russian *солдат*) to describe a soldier in general, while for the gender-specific terms she preferred the compound strategy mentioned above. Similarly, the gender-neutral word *pirikaska* ‘seller’ (< Russian *приказчик* ‘counterman’) was provided as a general term, but the corresponding gendered versions were *lāpkat rūpitan nē* ‘lit. in a store+

⁶ One of our respondents used these terms to refer to a nun and a monk, although originally the word *pūp* means ‘priest’.

working+woman' and *lāpkat rūpitan χum* 'lit. in a store+working+man'. This can be seen as an indication that at least this particular speaker prefers this strategy of expressing gender to the one discussed in Section 3.4.

In the rest of the languages in our sample, this strategy is fairly marginal. In most cases, the first part of the compound is a place with which the job is associated. This is the case, for example, with *köögitüdruk* 'lit. kitchen+girl' in Estonian, *kassaneiti* 'lit. cash register+girl' in Finnish, *konyhás néni* 'kitchenaid (lit. kitchen-ADJ)+aunt' in Hungarian, *kävppinissoon* 'female seller (lit. shop+woman)', and *asuntolasiäsá* 'lit. dormitory+aunt' in Aanaar Saami, and *унагудо рөезе* 'hotel boy' in Meadow Mari. One of our Meadow Mari respondents also mentioned the word *тулүдыр* 'lit. fire+girl' meaning 'female firefighter', but this is not attested in the corpus.

Gendered parts of the resulting compounds demonstrate some variation. For instance, in Estonian, the second part of a gendered compound starting with *kassa-* 'cash register, counter' can be *neiu* 'maiden' (25 occurrences), *tädi* 'aunt' (23), *poiss* 'boy' (4), or *tüdruk* 'girl' (3). For postman, the compound nouns consist of *post* 'post, mail' and the respective gendered noun: *postitädi* 'lit. post+aunt' (62), *postionu* 'lit. post+uncle' (10), and *postineiu* 'lit. post+maiden' (3). There were also 208 occurrences of *postipoiss* 'lit. post+boy', but in most of them, *postipoiss* occurred as a proper noun, that is, as a part of a newspaper's name. To refer to someone working in the kitchen, either a chef or a kitchenaid, a compound noun consisting of *köök* 'kitchen' and a gendered noun can be used. In the corpus, such compounds were *köögitüdruk* 'lit. kitchen+girl' (160 instances), *köögitädi* 'lit. kitchen+aunt' (21) and *köögipoiss* 'lit. kitchen+boy' (5).

In colloquial Estonian, similar compound nouns are also formed with male or female first names. A plumber, for example, can be referred to as *torujüri*, a compound which consists of *toru* 'pipe' and the male given name *Jüri* (see Mäearu 2002 for more examples). In regard to the job titles analysed in this study, this strategy could be attested in *köögikata* 'female chef, kitchenaid' (46 instances in the corpus), which consists of *köök* 'kitchen' and the female given name *Kata*. However, the examples found in the corpus suggest that *köögikata* is commonly used to refer to housewives and, more often than not, has a rather pejorative meaning. In

her article Mäearu (2002: 31) also mentions alternative names for other job titles included in the present study, such as *karjakusti* (< *kari* ‘herd’ and the male given name *Kusti*) and *lehmalisuus* (< *lehm* ‘cow’ and the female given name *Liisu*), which refer to a male and a female shepherd, respectively. However, these compounds were not attested in our corpus data and shall therefore not be discussed further. The same strategy is also found in the Finnish word *maajussi* (< *maa* ‘land’ and the male given name *Jussi*), which is a slightly derogatory term for ‘farmer’.

3.6. Borrowings with different stems

Among the occupations that were systematically considered in this study, there was one that in several languages had different stems for female and male representatives – ‘a member of a religious community of women/men who serve God and often live together in a monastery/convent’, that is, ‘nun’ and ‘monk’. The respective words are *nunn* and *munk* in Estonian, *nunna* and *munkki* in Finnish, and *nunná* and *munkki* in Aanaar Saami. Hungarian has two options for each, *apáca* or *nővér* for ‘nun’ and *szerzetes* or *barát* for ‘monk’, although the word *szerzetesnő* with a gendered part *-nő* ‘woman’ also exists (see Section 3.4).

Most of these words are ultimately borrowed from Indo-European languages, where the same strategy is used in this case; cf. Swedish *munk* ‘monk’ and *nunna* ‘nun’. Additionally, the existence of two distinct versions for this occupation might reflect the fact that religious titles in general tend to be gender-specific across languages. For example, in Finnish, even though there is a gender-neutral word for ‘priest’, *pappi*, the gender of the priest is very often specified using different strategies outlined in the earlier sections. This word is among the job titles that most frequently take the *-tar* suffix (see Section 3.2.1), and it has by far the highest percentage of gendered uses with the *nais-/mies-*prefix (see Section 3.1).

4. Summary and discussion

In this paper, we have provided an overview of the strategies that six Finno-Ugric languages use to express gender in job titles. The data on the strategies are summarized in Table 1 below: “++” means that the strategy

is common in a language, “+” stands for “common but possibly somewhat marked”, “(+)” means “fairly marginal”, and “-” means that the strategy is neither mentioned by our respondents nor attested in our corpus data.

Strategy	EST	FIN	HUN	Saami	Mari	Mansi
Gendered prefix	++	++	-	(+)	-	-
Productive feminine suffix	++	++	-	-	-	-
Borrowing with a feminine suffix	(+)	(+)	(+)	(+)	++	-
Replacing the gendered part	+	+	-	(+)	(+)	-
Job title + gendered noun	+	(+)	++	(+)	(+)	+
Compound with a gendered noun	+	(+)	(+)	(+)	(+)	++
Borrowings with different stems	(+)	(+)	(+)	(+)	-	-

Table 1. Expressing gender in job titles in Finno-Ugric languages

As is clear from the table, the languages differ greatly as to how many strategies they use and how frequently. Estonian and Finnish each have a wide selection of strategies ranging from very common and productive to fairly marginal ones. Aanaar Saami also uses almost all the strategies (which, judging by some specific gendered nouns, might to some extent be a result of its contact with Finnish), even though none of them is very productive. Each of the remaining three languages has one productive strategy and one or more that are quite infrequent. Among the strategies, the only two that we were able to find in all of the languages in question are compounds with a gendered noun, where the first part either is a job title itself or is somehow related to the job.

An important observation that can be made concerning the distribution of strategies is that different job titles behave differently in terms of how often they specify the gender and how common each of the gendered variants is. The preferences regarding the frequency of gendered forms in general seem to be mostly language-specific, with occasional overlaps. For example, the speakers of Estonian, Finnish, Hungarian, and Meadow Mari all tend to specify the gender of actors/actresses (notably, English also has a widely used lexicalized female variant for this job), but for the rest of the jobs there is hardly any clear tendency.

As for the frequency of specific gendered variants, it is mainly dependent on two factors: (a) how often the respective job is actually done by women and men, and (b) whether the job is seen as stereotypically

masculine or feminine. The first factor is crucial in the situation when the feminine variant exists in binary opposition to a masculine one which is not normally perceived as gender neutral. For example, in Meadow Mari, loanwords from Russian form such pairs, and *лётчик* ‘male pilot’ is much more frequent than *лётчица/лётчице* ‘female pilot’, while *доярке* ‘female milker’ is much more frequent than *дояр* ‘male milker’ (see Section 3.2.2). The second factor is relevant in the situation when a gender-neutral form is available, and thus, any gendered variant can be seen as marked. In this case, a speaker is more likely to use a gendered variant if the reality does not match the existing stereotype. For example, in Finnish and Estonian, gender-neutral words meaning ‘nurse’ or ‘nanny’ are more likely to take male prefixes, while with gender-neutral words meaning ‘soldier’ or ‘hunter’ female prefixes are notably more frequent.

As already mentioned in the introduction, this project is seen as a pilot study, which means that it might actually raise more questions than it is able to answer, both within individual languages and on a more abstract level. For instance, it is yet to be discovered what determines the choice of a specific strategy if several options are available. If specifying the gender is optional in a language, what factors motivate the use of the gendered variant? How stylistically neutral are different strategies and what are the stylistic properties of each of them? Can the strategies be combined, and what are the rules regulating their interaction? These and many other questions remain for further research.

References

- Engelberg, Mila. 2002. The communication of gender in Finnish. In Marlis Hellinger & Hadumod Bussmann (eds.), *Gender across languages. The Linguistic representation of women and men*, vol. 2, 109–132. Amsterdam: John Benjamins.
- Erelt, Mati, Tiiu Erelt & Kristiina Ross. 2007. Eesti keele käsiraamat, 3rd edn. Tallinn: Eesti Keele Sihtasutus. www.eki.ee/books/ekk09/.
- Hasselblatt, Cornelius. 2015. The representation of gender in Estonian. In Marlis Hellinger & Heiko Motschenbacher (eds.), *Gender across languages* 4, 125–151. Amsterdam: John Benjamins.
- Kegeyes, Erika. 2008. Sexusgrammatik und Gendersemantik. Ausdrucksformen des Geschlechts im Ungarischen. In Erika Kegeyes &

- Ágnes Hussar (eds.), *Genderbilder aus Ungarn. Ergebnisse der ungarischen Genderforschung*, 73–90. Hamburg: Kovač.
- Kyrölä, Katriina. 1989. *Feminiinijohdokset suomen kielessä*. Turku: Turun yliopisto Master's thesis.
- Kyrölä, Katriina. 1990. Kaunotar elää, laamannitar on kuollut. *Kielikello* 2/1990. www.kielikello.fi/-/kaunotar-elaa-laamannitar-on-kuollut.
- Laakso, Johanna. 2011. Some basic issues of gender theory from a Finno-Ugric point of view. In Hanneke Boode & Edit Zsadányi (eds.), *Gender Perspectives on Hungarian and Finnish Culture*, 7–24. Studia Fenno-Ugrica Groningana 6. Maastricht: Shaker Publishing.
- Lea, Diana & Jennifer Bradbery (eds.). 2020. *Oxford Advanced Learner's Dictionary*, 10th edn. Oxford: Oxford University Press. www.oxfordlearnersdictionaries.com/definition/english/.
- Mäearu, Sirje. 2002. Pindejürist prügikastikustini. Eesnimelised isikunimetused. *Oma Keel* 2002/1. 29–31.
- Nagy, Marcell & Viviana Patti. 2006. Lexical gender in Hungarian: Reflections of social structures in an electronic corpus. In Eva Maria Thüne, Simona Leonardi & Carla Bazzanella (eds.), *Gender, language and new literacy*, 182–195. New York: Continuum Press.
- Olt, Kerli. 2004. Sooliselt markeeritud sõnavara trükiajakirjanduses. In Reet Kasik (ed.), *Tekstid ja taustad III: lingvistiline tekstianalüüs*, 90–107. Tartu Ülikooli eesti keele õppetooli toimetised 28. Tartu: Tartu Ülikooli Kirjastus.
- Oravecz, Csaba, Tamás Váradi & Bálint Sass. 2014. The Hungarian Gigaword Corpus. In *Proceedings of LREC 2014*, 1719–1723.
- Pete, István. 2000. Férfinyelv-e a magyar? *Magyar Nyelvőr* 124. 198–115.
- Puna, Kerli. 2005. Mees- ja nais-alguliste liitsõnade kasutuse dünaamika. In Reet Kasik (ed.), *Tekstid ja taustad IV: tekstiliigivaatlusi*, 82–99. Tartu Ülikooli eesti keele õppetooli toimetised 29. Tartu: Tartu Ülikooli Kirjastus.
- Ross, Kristiina. 1996. Naine eesti keele grammatikas. *Vikerkaar* 11–12. 102–106.
- Tainio, Liisa. 2001. *Puhuvan naisen paikka. Sukupuoli kulttuurisena kategoriana kielenkäytössä*. Suomalaisen Kirjallisuuden Seuran Toimituksia 854. Helsinki: SKS.
- Tainio, Liisa. 2006. Gender in Finnish language use: Equal, unequal and/or queer? *Wiener elektronische Beiträge des Instituts für Finno-Ugristik*. webfu.univie.ac.at/texte/tainio.pdf.
- Vasvári, Louise O. 2015. Gender trouble in a grammatically genderless language: Hungarian. Marlis Hellinger & Heiko Motschenbacher (eds.), *Gender across languages*, vol. 4, 203–225. Amsterdam: John Benjamins.