

# Stereotypes and Taboo Words in Dictionaries from a Diachronic and a Synchronic Perspective – The Case Study of Croatian and Croatian Church Slavonic

Lazić D., Mihaljević A.

*Institute of Croatian Language and Linguistics, Old Church Slavonic Institute*

*dlazic@ihjj.hr, amihaljevi@gmail.com*

## Abstract

The paper deals with the lexicographic treatment of derogatory and sensitive vocabulary, in particular vocabulary related to social groups, in historical and contemporary Croatian (and Croatian Church Slavonic) dictionaries. The analysis of the dictionary data, motivated by the insights into the relation between dictionaries and society, is conducted to show how dictionaries reflect the worldview of the time, explain the diachronic development of the lexicographic approach to sensitive content, and propose improvements to contemporary descriptions based on social awareness. For that purpose, the treatment of selected lexical items from the following domains is presented: male and female, sexuality and taboo, ethnicity. It is shown that there is a clear distinction in worldview and lexicographic approach between historical and contemporary dictionaries, which is facilitated by the fact that contemporary dictionaries have to balance between political correctness and the corpus. However, the examples given in this paper show that there is still room for improvement.

**Keywords:** social stereotypes; offensive language; critical lexicography; historical lexicography; Croatian

## 1 Introduction

Description of derogatory and sensitive vocabulary – for instance, swear words, vulgar expressions, taboo words, etc. – has always presented a challenge for lexicographers. A type of vocabulary that can be perceived as offensive is the one related to social groups, such as ethnic, religious, gender, age, etc. The offensiveness of such vocabulary may stem directly from its meaning and the intention of the speaker to say something negative about someone or something, or indirectly from stereotypes and prejudices about a group they are grounded in (Schutz 2002: 638).

In recent years, the following factors have to a greater extent shifted such vocabulary into the focus of lexicographers' interest:

- 1) Political correctness has become an important topic in society and consequently in linguistics and lexicography (cf. Atkins & Rundell 2006: 422-430; Cloete 2014). Social awareness regarding gender equality, marginal social groups, different nationalities and religions, people with disabilities, etc. has changed considerably (cf. Allan & Burridge 2006; Mills 2008; Wodak & Benke 2000; Talbot 2005). This has sometimes led to the process of euphemisation and disphemisation and a constant shift of attitude towards certain words or expressions (a well-known and often-quoted example are words like *Negro*, *Black*, *Afro-American*, *African American*).
- 2) Modern dictionaries are usually based on large computer corpora. With higher availability of such corpora, the corpus approach (cf. Tognini-Bonelli 2001) has become the norm in lexicography as an objective approach to language description. However, since actual language usage as attested in corpora is not always polite and politically correct, the question has arisen how sensitive vocabulary, especially vocabulary related to social groups, can be described in a dictionary that aims at being both descriptive and socially responsible.

In addition to being potentially offensive, the vocabulary itself, its usage, as well as its lexicographic description often reflect stereotypical views and values which are culture- and time-specific, and in that sense, the dictionary material can testify to the worldview of a certain society and time.

In this paper, the lexicographic treatment of several sensitive groups of lexical items in historical and contemporary Croatian Church Slavonic and Croatian dictionaries is presented in order to:

- 1) show how dictionaries reflect the worldview of the time;
- 2) explain the diachronic development of the lexicographic approach to such content;
- 3) propose improvements and strategies that could be applied in a modern, socially responsible dictionary.

## 2 Key Concepts and Previous Research

Sensitive lexical items can be defined as lexical items that have “strong connotative values and derogatory implications” (Cloete 2014: 482). The notion of sensitive vocabulary encompasses a diverse group of lexical items. Hartevelde and van Niekerk (1996: 382, 385, 387, 389, 391) proposed the following categorization: 1) racist lexical items, 2) sexist terms and sensitive lexical items which indicate stigmatized sexual phenomena, practices, and preferences, 3) sensitive lexical items which indicate stigmatized physical or mental conditions and phenomena, 4) sensitive lexical items within a social, political, and religious structure, and 5) obscene and vulgar lexical items, abusive language, and swear words. Since these

categories differ in the degree and nature of their sensitivity and offensiveness, a slightly different lexicographic treatment for each category was suggested.

Offensiveness is a concept closely related to sensitivity. Jay (1992: 160-161) defines offensiveness as denoting “the degree to which a certain word or concept possesses negative or aversive properties”, i.e. the degree of negative content. Very offensive words have the potential of becoming taboo words, words we refrain from using. In that sense, sensitivity and offensiveness can be regarded as quasi-synonyms, although offensiveness can imply a higher degree of negativity, thus denoting a somewhat narrower concept. Janschewitz (2008: 1067) relates offensiveness to the reaction of a person who hears (or reads) a word, the extent to which they perceive it as “personally offensive or upsetting”. Tabooness, on the other hand, is the extent to which a word is “offensive and upsetting” in society in general. Sometimes the term *offensive* is differentiated from the term *derogatory*. The former refers to the reaction of the listener or reader to a negative content, and the latter the intentions of the speaker or writer to express a negative attitude towards the referent (Norri 2000: 77). Since offensiveness can be caused unintentionally and not necessarily by a derogatory way of speaking, in this paper, we will mainly use the term *offensive* and the broader term *sensitive*, unless we have the intention of the speaker in mind.

Dictionaries have often been criticized for including offensive vocabulary as well as portraying certain social groups in a stereotypical and/or negative manner, thus codifying and strengthening prejudices that might exist in the society.<sup>1</sup> Having that in mind, a lexicographer can feel urged to omit sensitive content or change linguistic facts, an approach that can be criticized as falsifying reality. However, there is an overall agreement that a dictionary should reflect the real language usage of a certain period (Cloete 2014: 482). Rather than omitting sensitive items that meet the inclusion criteria (e.g. frequency in the corpora), they should be labelled and described properly. In other words, linguistic and lexicographic facts should be distinguished (Bratanić 2005: 39). While language usage does not have to be correct and should be described as such, its lexicographic treatment should be guided by social awareness and should not contribute to strengthening and maintaining the inequality which might be reflected in the language.

Among the elements of lexicographic description which have been discussed with regard to vocabulary sensitivity are the following (Harteveld & van Niekerk 1996; Cloete 2014): 1) choice of headwords, 2) usage labels, 3) metalanguage, 4) references to semantically related items (synonyms, antonyms), where referring to offensive items can be a problem, 5) expressions containing sensitive items, and 6) the choice of illustrative material, such as collocations, editorial usage examples, and citations. Moreover, the lexical treatment of such items will depend on (Cloete 2014: 482): 1) the type of dictionary (e.g. the approach is likely to be different in a school or learners’ dictionary and a dictionary aiming at adult native users), 2) the category of a sensitive item, 3) attitudes within a certain community.

As mentioned above, sensitive items can vary in their degree and nature of sensitivity and/or offensiveness (cf. Coffey 2010: 1278-1279). Moreover, offensiveness can be context-dependent as an item can be offensive in all or only in some of its senses (cf. Harteveld & van Niekerk 1996: 383). Schutz (2002: 638) pointed out two aspects of offensiveness, direct offensiveness originating from the speakers’ intention to say something negative about someone or something, and indirect offensiveness caused by the stereotype a lexical item is grounded in. Thus, he discriminates between directly offensive items (e.g. *nigger*), indirectly offensive items (e.g. *Dutch treat*), and both directly and indirectly offensive items (e.g. *Jew*, *unmanly*). All these aspects should be kept in mind when describing the usage of an item in a dictionary.

Until now, numerous studies have been conducted in the field of lexicographic treatment of sensitive vocabulary and the social aspect of lexicographic work. They have analysed different grounds of discrimination, such as gender (Fournier & Russel 1992; Russel 2012; Moon 2014, etc.), ethnicity (e.g. Moon 2014), age (e.g. Moon 2014), or illness and disability (e.g. Norri 2019). Some studies regarded certain elements of lexicographic description, such as usage labels (e.g. Norri 2000), or examples (e.g. Fjeld 2015, for Nordic dictionaries). Challenges and lexicographic choices regarding social sensitivity in ongoing dictionary projects (Danish and Swedish) are explained in Jensen et al. (2018) and Petersson & Sköldberg (2020). Additionally, several studies have been conducted on Slovenian lexicography (Gorjanc 2004, 2005; Trojar & Žagar Karer 2013). Few studies have analysed the lexical items connected to sexuality in dictionaries compiled before the 20<sup>th</sup> century (Dykstra 2006; Schweickard 1997; Lebsanft 1997; Radtke 1986). In the Croatian context, the research has merely focused on the presence of gender inequality and gender stereotypes in dictionaries (Bratanić 2005; Dakić 2017; Pišković 2017), while no studies have concerned users’ reactions and expectations regarding the socially sensitive content in dictionaries. Historical Croatian dictionaries have up to now not been studied from the social perspective.

### 3 Corpus and Methodology

In this paper, the analysis of the lexicographic presentation of the sensitive content is conducted from the point of view of historical and contemporary lexicography. Similar entries and entries from the same domains are analysed in historical and contemporary Croatian dictionaries to determine changes in the lexicographic approach to such content. The corpus for our analysis consists of the following dictionaries:

- 1) historical dictionaries: Vrančić (1595), *Dictionarium quinque nobilissimarum Europae linguarum*; Kašić (around 1600), *Hrvatsko-talijanski rječnik*; Mikalja (1649), *Thesaurus linguae Illyricae* – the first modern-type Croatian dictionary; Habdelić (1670), *Dictionar ili reči slovenske*; Jambrešić (1742) *Lexicon Latinum interpretatione Illyrica, Germanica et Hungarica locuples*; Stulli (1801), *Lexicon latino-italico-illyricum*; (1805),

<sup>1</sup> One of the famous examples has been the definition of Bangkok as “a place often mentioned where there are a lot of prostitutes” (*The Herald* 1993). In Nordic context, the inclusion of the entry *grønlanderstiv* (‘drunk as a Greenlander’) in a Danish dictionary is an example that has gained a lot of publicity (Farø & Jensen 2018: 219). A Croatian example that has been criticized is the definition of woman as a being opposite of a man (*Libela* 2013). However, in Croatia, the criticism of that kind is not very common and has almost exclusively come from the activists and minority groups themselves, and rarely from the general public.

*Rjecoslozje*; (1810), *Vocabolario italiano-illirico-latino*; Šulek (1860), *Njemačko-hrvatski rječnik*; the dictionary of the Croatian Academy of Sciences (= ARj) that was compiled from 1880 to 1976 and was based on the corpus from the 12<sup>th</sup> century to the contemporary period; *Slovník jazyka staroslověnského* (= *Slovník*, 1966-1997), a Church Slavonic Dictionary that is being used as the basis for the RCJHR; *Rječnik crkvenoslavenskoga jezika hrvatske redakcije* (= RCJHR, *Dictionary of the Croatian Redaction of Church Slavonic*), an ongoing project of compiling a dictionary of Croatian Church Slavonic based on the corpus from the 11/12<sup>th</sup> century to the 16<sup>th</sup> century;

- 2) contemporary dictionaries: the printed dictionary *Rječnik hrvatskoga jezika* compiled by Šonje (= RHJ; 2000), *Hrvatski jezični portal* (= HJP; <http://hjp.znanje.hr>) – currently the only Croatian comprehensive monolingual open-source web dictionary and thus presumably the one most frequently used by adult native speakers, and *Veliki rječnik hrvatskoga jezika* (= VRH; 2015) – the most recent comprehensive monolingual dictionary, available in a printed and a digitized version. Only general dictionaries for adult native speakers have been analysed; special-purpose dictionaries, such as school dictionaries, are beyond the scope of this study.

The lexicographic treatment of sensitive lexical items will be illustrated by selected entries from the following domains: male/female (sex distinction), sexuality and taboo words, ethnicity. The analysis will focus on the selection of headwords, definitions, usage labels, choice of examples, collocations, idioms, and pragmatic notes.

## 4 The Analysis

### 4.1 Male – Female

#### 4.1.1 Historical Dictionaries

If not stated otherwise, the examples and definitions are from ARj, which has the greatest number of examples and the most elaborate definitions. Older Croatian dictionaries tend not to be socially sensitive towards the stereotypes regarding female identity. In ARj, women are defined as persons whose organism is designed to give birth – “osoba kojoj je organizam udešen za rađanje djece”. In some definitions, it is stated that the word *woman* can be used to denote a man acting like a woman, i.e. someone acting cowardly – “oznaka ili obilježje čovjeka koji se vlada kao žena, koji je plašljiv, kukavica”. The adjective *female* is defined as the opposite of male in strength, courage, and boldness. It is stated that the adjective denotes someone weak, soft, limp, timid, fearful, and cowardly – “suprotan muškom u hrabrosti, smjelosti i snazi, tj. slab, mek, mlitav, bojažljiv, strašljiv, kukavički”. In ARj, it is stated that the words derived from the word *woman* can sometimes be used to denote a man if he is fearful as a girl, soft, sensitive, spoiled, weak, and timid, e.g. “u prenesenom smislu o muškarcu, koji je bojažljiv kao djevojka”; “muškarac, koji je sličan ženi, koji radi i vlada se kao žensko, koji je mek, osjetljiv kao žensko, kome žena zapovijeda, koji voli ženske i trči za njima”; “kao žena, ženski, t. j. razmažen, slab, bojažljiv činiti da tko bude mek i osjetljiv kao ženska”; “mek, slab kao žena”. Some older dictionaries list the most common collocations with the words denoting a woman, such as: *ill-tempered, evil, wild, quarrelsome, disgraced, dishonorable, dishonest, insatiable, lewd, harlot, sinner, adulteress* – “zločudna, zlobna, zlopametna, divlja i grđa, karljiva, osramotjena, nečastna, nepoštena, razpuštenica, zla, bludnica, zlica, nečastnica, grešnica, hotnica, priljubovca, mrska, pogana, nezasična, nepoštena, nikad sita...”. In some dictionaries (e.g. Mikalja), a lot of collocates are connected to the meaning of being married or pregnant. Older dictionaries often give examples, which are also full of stereotypes: a good woman gives birth only if she is married; a new-born is fortunately male; women do not have the brain of a man, but of a child; a man is the head, and a woman is the grass; women have long hair, but a short brain; a woman should be quiet when a man speaks; dogs should bark, and women should be quiet; you should not trust a woman because she changes like the Moon; a man should be a hero, and not act like a woman – “Da bude rodila ne budući za mužem, ne bi bila držana za ženu dobru”; “da j' dite na su sreću muško”; “Mi žene ... neimamo pamet mušku na djetsku”; “čovjek je glava, a žena trava”; “Žena je dugokosa, a kratkoumna”; “Žena jezik za zube kad muž govori”; “Kučka nek laje, a žena nek muči”; “Ne vjeruj ženi, er se kako mjesec mijeni”; “Ta nemoj me ženski udarati, već me udri, čim s' junaci biju”. In the dictionaries, it is often stated that there are jobs and duties fit for a woman and those fit for a man – “poslovi su odijeljeni: čovjek u poļu, a žena u kući”; “ti ženkari lepo šiju ... i druge ženske poslove rade ... samo da ne traže muškaraca”; “ženskadija pravi večeru”; “To u nas radi ženskadija”; “o muškarcu, koji se ponaša kao ženska glava i zna obavljati ženske poslove”. In ARj, domestic violence is depicted as normal: he who does not hit his wife is not a man; you should hit a horse and a woman for them to be obedient – “Ko ženu ne bije, on čovjek nije”; “Ženu i konja udri, ako želiš, da su ti pokorni”. In Šulek's dictionary, it is stated that the woman is more cunning than the devil – “žena je lukavija od vraga”, and in Stulli's, that she is the worst beast – “žena je na svijetu najgora zvir”.

Some gender stereotypes are also visible in (Croatian) Church Slavonic Dictionaries and are mostly influenced by the fact that their corpora mostly consist of Biblical and other religious texts. In *Slovník* and RCJHR, women are depicted as weak, they are often shameless, and they brought destruction to Adam, Joseph, and David. It is stated that men should not listen to women and that women often present themselves as much prettier than they are. It is also stated that the devil often appears in the form of a woman – “muži že takođe sь svoimi živuče ženami êko nemočnêisei veči žen'scêi vzdajuče čьst”; “ženom' bo adamъ iz raê spuen' bi ženom bo pr(a)vdni osip' zatv(o)ren' bê v' tamnici ženomъ d(a)vidъ uriju stv(o)ri ubiti”; “zač ti posluša glasa ženi tvoee ku ti dah' pod' tvoju oblast' i na tvoju volju”; “o gorko i čermernotim' ženom' ke lice svoê pomazuju i lipše se čine nere ih' e bog' učinil”; “mnogo bo krat' dèvl' prihoêše k nemu va obraze žen'sceem”.

#### 4.1.2 Contemporary Dictionaries

Even though the content of the entries related to women and men in contemporary dictionaries clearly shows how the worldview has changed over time – for example, women are no longer depicted as having the primary role of giving birth<sup>2</sup> nor is domestic violence promoted – some stereotypes persist, and they regard, in the first place, what are thought to be typical male/female characteristics, and to some extent the attitude to sexuality of the respective sexes.

The stereotypical view and inequality can be observed in the first place in definitions, collocations, and examples. The following examples from HJP speak for themselves: *male* (*muški*, adj.) – one of the senses: *worthy of a man, a real man* – “dostojan muža, pravog muškarca”; “*male hand*” (*muška ruka*): *a man who in a household does the typical work which is not suitable for a woman* – “muškarac koji u kući obavlja tipične poslove koji nisu za ženu”; *manly/masculine* (*muževan*, adj.): *having all qualities of an adult or an honourable man [manly appearance]* – “koji ima sve odlike odrasla muškarca ili časna muža [*muževna pojava*]”; “*male old wife*” (*muška baba*): *a man with some female characteristics (e.g. talks a lot and the like)* – “muškarac s nekim ženskim osobinama (mnogo priča i sl.)”. In the examples, the positive attitude towards typical masculine traits is emphasized and they are not questioned. Similar sexist definitions can also be found in RHJ, e.g. “*female head*” (*ženska glava*): *a woman with her peculiar way of thinking* – “žena sa svojim osebujnim mišljenjem”.

If the definitions are compared with those in VRH, it is obvious that the latter are more neutral and less emotionally coloured, but the social construct of a typical masculine nature is taken as given: *manly* (*muški*, adv.) – *like a man, in a manner of a man [to act / say / drink / hit someone like a man]* – “kao muško, na način muškoga [*muški postupiti / reći / piti / udariti koga*]”; *manly/masculine* (*muževan*, adj.): *showing characteristics of the male sex [masculine look; masculine attitude; masculine appearance]* – “koji pokazuje odlike muškoga spola [*muževan izgled / nastup; muževna pojava*]”; expressions “*male hand*” and “*male old wife*” are not listed in the dictionary.

Social stereotypes of typical male and female characteristics have been lexicalized in the adjectives *ženskast* and *muškobanjast*, which denote femininity in men (*womanish*) and masculinity in women (*mannish*) respectively, and are normally used disparagingly or mockingly. They can thus be regarded as both directly offensive to the person or the group they are used for, and indirectly offensive to women and men because of the stereotype they are grounded in. However, none of this is indicated in Croatian dictionaries – the words are not even labelled as derogatory/offensive. Moreover, the examples given in VRH reveal further stereotypical view on certain professions which are socially not regarded as masculine/feminine: *a masculine policewoman* – “muškobanjasta policajka”; *For his feminine and somewhat unnatural movements, he could be a ballet dancer* – “Po svojim ženskastim i pomalo namještenim kretnjama mogao bi biti baletni plesač.” It can be argued that such stereotyping is not necessary for describing the meaning of the headwords, neither is it the typical context of their use<sup>3</sup>, and should thus be avoided.

Collocates given for words from the domain men and women often differ in dictionaries, those for women often being related to beauty and emotions and those for men expressing physical strength. For example, in VRH a woman is *energetic, young, pretty, unhappy*, while a man is *brave, unknown, real, threatening, average, middle-aged, loved*. On the other hand, some of the examples in VRH show that an attempt has been made to make it more inclusive and up to date, e.g. in the entry *women* (*žena*), the following examples have been included: *fighting for women's rights* (“borba za prava žena”); *a man trapped in a woman's body* (“muškarac zarobljen u tijelu žene”).

## 4.2 Sexuality

### 4.2.1 Historical Dictionaries

Words related to sexuality<sup>4</sup> are rarely included in older dictionaries. Sexual relations outside marriage and those with the same sex are described as unnatural and sinful – “objašnava to starim vjerovanjem po kome se smatralo, da su tjelesni odnosi između muža i žene i začecje religijski nečisti i grijeshni”. Sodomy is defined in ARj as unnatural sexual intercourse and in the usage example, it is stated that a man or a woman who has committed sodomy should be burned – “sodomija – nenaravno spolno općeće – sodomija jest, kada muški polj ima činjenje s muškim spolom”; “kadano čovjek sgriješi z' ženom naopako, to jest učini sodomiju”; “sodomska – sedmi grijeh je proti naturi aliti sodomski”; “ako bi se tko naša u grihu nepodobnu, ča se zove grijeh sodomski, ali bi bila muška glava ali ženska ... ima se sažgati”. Prostitutes and mistresses are described as *unclean, sinful, wicked*, and it is stated that they will not go to paradise – “Nisi čista, da bludnica”; “sagriješiti s ženom bludnicom jest blud preprost”; “bludnici i bludnice ... biti će polivani gorućim paklom”; “bludnici ne će ulisti u raj”. Words denoting homosexual men are not attested in older dictionaries and vocabulary related to sexuality and relationships is always described traditionally, having the heteronormative relationships in mind.

### 4.2.2 Contemporary Dictionaries

Due to social changes, today we are more familiar with various forms of sexual orientation and preferences, so it is not surprising that the associated vocabulary is more present in contemporary dictionaries than in historical ones. No signs of reluctance to include such vocabulary have been observed in analysed dictionaries, but several other problems can be

<sup>2</sup> An exception is the definition of the word *woman* (*žena*) in RHJ, which has often been quoted in Croatian publications on sexism in language and dictionaries (cf. Bratanić 2005; Dakić 2017): *a human being of opposite sex than a man, who can give birth to children and take the main care of the upbringing and education of children* [“ljudsko biće po spolu suprotno muškarcu, koje može rađati djecu i preuzeti glavnu brigu za uzgoj i odgoj djece”].

<sup>3</sup> The nouns *policajka* and *baletan* do not appear as common collocates of the adjectives *muškobanjast* and *ženskast* in *Croatian Web Corpus – hrWaC* (accessed 26/02/2021).

<sup>4</sup> “Central aspect of being human throughout life and encompasses sex, gender identities and roles, sexual orientation, eroticism, pleasure, intimacy and reproduction.” (European Institute for Gender Equality 2021).

discussed: usage of labels, cross-references, and definitions of headwords related to sexuality.

In contemporary dictionaries, there are several words denoting a homosexual man, varying from neutral through colloquial to derogatory and vulgar. However, sometimes labels are missing or do not correspond to actual usage (usually they are too mild). For example, the word *peder* is labelled only as colloquial, while it can be argued that it is often used disparagingly and can be perceived as offensive. Moreover, unlike vulgar words (discussed in the next section), which are usually not listed as synonyms / cross-referenced within neutral entries, the neutral entry *homoseksualac*, for example, contains several synonyms in HJP and VRH ranging from the colloquial *homić* to the very offensive *dajguz* (literally a “butt giver”), but the stylistic value of the synonyms is not indicated. It can be discussed whether and how the offensive synonyms should be listed within neutral entries. While it can be useful for the user to get a list of similar words to choose from in text production, an uncritical listing of offensive expressions contributes to the negative view of social groups and could cause public disapproval. A possible solution could be to include appropriate labels both when the expressions appear as headwords and when they are listed within other entries.

A traditional, heteronormative view often occurs in descriptions of vocabulary related to sexuality and relationships in general. Even though some improvement can be noticed, e.g. “*male virgin*” (*djevac*) is defined as *the one who is living as a virgin, who has renounced or is deprived of the touch of a woman* – “onaj koji živi u djevičanstvu, koji se odrekao ili je lišen dodira žene” in HJP, but as *a man without sexual experience* – “muškarac bez spolnoga iskustva” in VRH, in many cases, such as in definitions and examples provided for words *girlfriend* (*cura*), *boyfriend* (*dečko*), *lover* (*ljubavnik/ljubavnica*), etc. the meaning and usage are described having heterosexual relationships in mind.

A dissimilar social attitude towards female and male sexuality is also reflected in dictionaries, and it can be observed in the first place in the inventory of expressions included in a dictionary and their definitions, which is also more prominent in HJP than in VRH. In HJP, expressions like the following can be found: *secondary virginity* (*drugo djevičanstvo*), defined as *the condition of a married woman who, according to social and economic reasons and customs, has an older husband and is left without an erotic life in her vital years* – “stanje udate žene koja prema društvenim i ekonomskim razlozima i običajima ima starijeg muža i u vitalnim godinama ostaje bez erotskog života”; “*mental female prostitute*” (*duševna prostitutka*), defined in one of the senses as *a woman who takes advantage of a man and keeps him hoping he will be successful; one who promises or gives hope of an intimate relationship she does not intend to get into* – “ona koja iskorištava muškarca i drži ga u nadi da će postići uspjeh; ona koja obećava ili daje nade u intiman odnos u koji ne misli ući”. In the examples, female sexuality, age, etc. are portrayed as a means of taking advantage of men, gaining social and financial security, etc. There are no male counterparts in the dictionary and definitions feel outdated and one-sided since there is no comment on the social context they have arisen from or their usage today. Moreover, there is hardly any evidence of their usage in available Croatian corpora (e.g. *Croatian web corpus – hrWaC*; accessed 26/02/2021), so it is not unexpected that the expressions do not appear in the newer VRH. However, some vulgar expressions referring to female sexuality and character are to be found in both dictionaries, for example, the very vulgar “*cold cunt*” (*mrzla pizda*), meaning “a frigid woman” and “*wolf with a pussy*” (*vuk s pičkom*), defined as *a very determined, enterprising, strict woman* – “Vrlo poduzetna, odlučna, oštra ženska osoba”. These expressions are also very scarcely attested in contemporary corpora, so it is questionable whether they should be included in a dictionary.

### 4.3 Taboo Words

#### 4.3.1 Historical Dictionaries

As examples of taboo words<sup>5</sup>, we have chosen three word-formation clusters – *kurac* and its derivatives – an offensive word for penis, *pička* and its derivatives – an offensive word for vagina, and *jebanje* and its derivatives – an offensive word for sexual intercourse. Older Croatian dictionaries (e.g. Vrančić and Kašić) generally do not list those words. In Jambrešić’s bilingual Latin-Croatian dictionary, the words *mentula* and *penis* are listed, but without their Croatian equivalents.<sup>6</sup> In the entry *penis*, it is stated that the translation can be found in the entry *mentula*, but the definition found there is also in Latin – *membrum pudendum viri*, “shameful male body part”. In ARj, although it is a monolingual dictionary, and the entries are in Croatian, some taboo headwords are defined only by their Latin equivalents or have a Latin definition: *kurac* – *mentula*; *kurat* – *mentulatus*; *kurcati se* – *penis vocabulo abuti*; *kurcoglavac* – *senecio vulgaris L.*; *kurčev* – *mentulae*; *kurčevit* – *ut mentula*; *kurčiti se* – *penem imitor*. Some derivatives have a Croatian definition, but a part of the taboo word is censored: *kura* – *hyp. ...ac, kurcovina* – *augm. od ...cov, kurčekanja* – *augm. ...ac, kurčenje* – *djelo kojijem se ko ...či*. It is interesting to note that the headword is not censored, but the word derived from the same root in the definition is. Sometimes there is a combination of the censored Croatian definition followed by a Latin definition, e.g. *kurcanje* – *djelo kojijem se ko ...ca, penis creber usus in loquendo*. In most of these entries, there are no examples of usage, although this dictionary usually gives many examples. In one of the entries where the example is given, the taboo word is also censored: *kurcov* – *ti si već ...cov*. One word is defined in German: *kurcokret* – ‘*ein komisches wort fur celer*’. Words denoting vagina are found only in ARj and are similarly defined. Only Latin translations are given when defining

<sup>5</sup> On the history, definitions, and features of taboo words see Allan & Burridge 2006; Jay 1977, 1992, 2000, 2009; Jay, Caldwell-Harris & King 2008; Janschewitz 2008. The term *taboo words* describes “the lexicon of offensive emotional language. A taboo is a ‘ban or inhibition resulting from social custom or aversion’ (The American Heritage Dictionary of the English Language, 2000). Taboo words are sanctioned or restricted on both institutional and individual levels under the assumption that some harm will occur if a taboo word is spoken.” (Jay 2009: 154). “Taboo words represent a class of emotionally arousing references with respect to body products, body parts, sexual acts, ethnic or racial insults, profanity, vulgarity, slang, and scatology (Jay, 1992, 2000).” (Jay, Caldwell-Harris & King 2008: 83).

<sup>6</sup> On the difference between the two and on Latin sexuality vocabulary in general see Adams 1982.

the following words: *pica* – *cunnus*; *pičkar* – *fututor*, *amans cunni*; *pičkaroš* – *fututor*, *amans cunni*.<sup>7</sup> The word *pička* is defined by both German and Latin equivalents: *pička* – *Scham*, *cunnus*. The word derived from the same root *pican* is defined in Croatian as the boy who likes to play and be friends with girls and is spoiled – “dječak koji je razmažen i rado se s djevojčicama igra i druguje”.

In older dictionaries, male genitalia can also be found as entries *ud* and *udo* (‘extremity/limb’) with adjectives *dishonest*, *shameful*, *ashamed*, *secretive*, *childbearing*, and genitalia of both men and women as entries deriving from the words denoting shame and disgrace: *sram* and *stid*, usually defined as the thing that women and men hide.

Words derived from the root *jeb-* are found only in ARj and Stulli. In ARj, they are often defined by Latin equivalents, e.g. *jebač* – *futuens*. Two words have both German and Latin equivalents: *jebac* – *der hurer*, *fututor validus*; *jebaonica* – *das bordell*, *lupanar*. Some words have Croatian definitions with taboo words censored in the definition: e.g. *jebane* – *djelo kojijem se ...e*. The word *jebičina* is defined as *augm. od jebica*, without the taboo word being censored. Only one example is given also in its censored form – “Taman laže, vsi mu je.li majku!”. The equivalents in other Slavic languages are given in uncensored forms although they are very similar to the Croatian word, while the uncanonical forms of these Slavic words are censored – “usporedi novoslov. jebati... češ. jebati, praes. ...am i ...u, polj. jebać, ...ę”.

### 4.3.2 Contemporary Dictionaries

The words and their derivatives discussed above appear as headwords in two contemporary Croatian dictionaries – HJP and VRH – while reluctance to include them can be noticed in RHJ. In the latter, only neutral words such as *vagina* (as well as its Croatian synonym *rodnica*), *vulva* (as well as *stidnica*) and *penis* (as well as *udo*) are found, while their vulgar synonyms are not mentioned either within the entry of a neutral headword or as a separate headword. The headword *jebati* exists in the dictionary and is labelled as vulgar, but its description is deficient since only the following definition is provided: *to have sexual intercourse with a woman*. However, the word can denote any type of sexual intercourse, and that has been taken into account in the other two analysed dictionaries, where the gender of those involved is not specified. Thus, it is evident that the definition in RHJ reflects a traditional view by mentioning only heterosexual intercourse and associating men with an active role. Furthermore, other senses of the word *jebati* (‘to bother someone; to ignore someone/something’), its usage as a swearword as well as numerous expressions and derivatives it appears in have been omitted in RHJ, so it can be concluded that the lexicographic description is not up to date and that some aspects of language usage – especially colloquial, have been disregarded.

In the remaining two dictionaries, where vulgar words for genitalia and sexual intercourse are listed as well as their various derivatives, their stylistic value is indicated by means of:

- 1) a stylistic label *vulg.* (‘vulgar’), which is found in all entries discussed here;
- 2) (rarely, non-systematically) an explanation which is:
  - a. a part of the definition, e.g. one of the senses of the word *kurac* is defined in the following manner (HJP, VRH): *the word which as a filler often fills a pause in a sentence and which is orthographically expressed by a dash (hyphen) or a comma placed according to the intonation, in texts usually abbreviated to k... – “riječ kao poštapalica u rečenici često popunja stanku koja se pravopisno izražava crtom (povlakom) ili zarezom postavljenim po intonaciji, u tekstovima obično kraćena k...”;*
  - b. added separately, as additional information, e.g. within parentheses in the entry *pizda* (HJP), after the senses have been listed: *the word is very rude and inappropriate in polite communication, in texts it is usually abbreviated to p... – “riječ je vrlo nepristojna i neprikladna za iole pristojan način izražavanja, nalazi se u tekstovima obično kraćena p...”;*
- 3) cross-references: there is a tendency to include references to neutral entries in the vulgar ones, but not vice versa; however, some exceptions exist, e.g. the neutral entry *penis* includes references to both neutral/scientific *falus*, colloquial *pimpek*, and vulgar *kurac* in VRH. In HJP, only neutral words (*spolovilo*, *udo*) are listed as synonyms (note the difference in the synonyms given in the two dictionaries); similarly, no vulgar entries denoting sexual intercourse (*jebati*, *ševiti*, *fukati*...) are referenced to within either vulgar or neutral entries. It can be discussed whether colloquial and vulgar synonyms should be listed at least within entries of a similar style, if not all of them; it could be a useful information for language production.

## 4.4 Ethnicity

### 4.4.1 Historical Dictionaries

In older dictionaries, ethnonyms<sup>8</sup> are often defined depending on their social and historical background. The ethnic groups more closely related to Croatia are described in more detail and with more examples and stereotypes, e.g. in ARj, the Turks are described as warriors and are sometimes described negatively as enemies, non-believers, and liars – “Turci nas su oplinili i požgali”; “u Turčina nigda vire nije”; “laže ka Turčin”; they are also described as people who smoke a lot and drink coffee and wine – “Turci vino piju”; “pije kavu ka Turčin”; “puši ka Turčin”. Most stereotypes are attested in ARj in entries derived from the words denoting Gypsies. Gypsies are often described as people who *deceive*, *cheat*, *tell fortune* and *wonder*, *steal* and *lie*, and are *lazy* – “cigančiti – cigančiti je osobito prošiti ili iskati navaľujuć, ne odstupajuć, kao što čine Ciganke”; “ciganiti = varati, prosjačiti”; “svit Cigane vrlo kori od svi ľudi da su gori, jer su lini od kolina i lupeži od starina”; “još ciganski i dlane gledate”; “prijateľa ki ukani, još je gorši neg Cigani”. It is stated that calling

<sup>7</sup> On the etymology and the lexemes denoting genitals see Reinhart 1994.

<sup>8</sup> I.e., the names of ethnic groups. For more see Koopman 2016.

someone a Gypsy is an insult – “poruga čovjeku koji laže i vara.”, but it is not noted that the definitions of Gypsies as fraudulent are based on a stereotype.

In addition to negative stereotyping, some positive stereotyping can also be found, mostly in the entries connected to Bosnia and Croatia. Bosnian women are *pretty* – “lepše djeve Bosankiće; oženio sam se Bošnjakićom lijepom divojkom”, Croatians are *good, famous, proud*, etc.

#### 4.4.2 Contemporary Dictionaries

In contemporary dictionaries, ethnonyms are approached more cautiously and are either not included at all as headwords or, when they are, they are defined neutrally, e.g. with regard to their geographical origin, and their usage is not exemplified. On the other hand, prejudices and stereotypes tend to be revealed in related words and expressions, such as derivatives.

Prejudices about social groups often give impetus to the development of secondary meanings as well as the formation of derivatives and expressions, which are usually more or less derogatory given the fact that they are often based on characteristics that are considered socially unacceptable or undesirable. They are even indirectly offensive to the social group which the underlying stereotype regards. The example of that is the colloquial/offensive word *Gypsy* (*Ciganin*) and its derivatives like *ciganski*, *cigančiti*, *ciganija*, *ciganluk*, etc., which exist in abundance both in the language itself and in dictionaries and denote something deceitful, messy, or dishonourable either in their secondary or even primary and only sense. However, such words are not always labelled as derogatory, offensiveness to the group is never indicated, and the stereotype is rarely commented on.<sup>9</sup>

In addition to the negative, some positive stereotyping can also be found, as in the expression *Slavic soul*, defined in the following way: *according to the established positive prejudice, a peaceful human nature, a magnanimous person* – “po uvriježenoj pozitivnoj predrasudi nesebična, miroljubiva narav čovjeka, široka duša” (HJP, VRH). Negative stereotyping can both regard the group one belongs to (e.g. expressions *Croatian envy* (*hrvatski jal*) and *Croatian silence* (*hrvatska šutnja*) in HJP and VRH) and other groups, such as neighboring nations (e.g. RHJ, HJP, VRH: “*Bosnian pot*” (*bosanski lonac*) – regarding complicated political circumstances; *srbovati* – defined as *to express Serbian national feeling intrusively* – “nametljivo izražavati srpske nacionalne osjećaje”).

In the entries related to ethnic groups, stereotypes are commented on more frequently than in other semantic groups (probably because they are perceived as especially sensitive), even though not very often, for example in the definition of “*Bosnian pot*”: *intricate political circumstances typical for Bosnia (according to prejudices outside of Bosnia)* – “zamršene političke prilike tipične za Bosnu (prema predrasudama izvan Bosne)” or *balkanština: primitivism and dishonest actions in public, cultural and political life, which are according to preconceptions in Western Europe considered typical of the Balkans* – “primitivizam i nečasni postupci u javnom, kulturnom i političkom životu što se po preduvjerenjima u Zapadnoj Europi smatra tipičnim za Balkan” (HJP). Sometimes, expressions like *allegedly* are used, as in the example *Croatian silence: allegedly the conformism common for Croatian public and politicians* – “navodno uobičajeni konformizam hrvatske javnosti i političara”.

## 5 Conclusion

In historical dictionaries, many stereotypes have been attested, and it is obvious that dictionary compilers were not aware of the potential offensiveness of some entries. Many entries also reflect the worldview from a certain period and prevailing stereotypes. In defining the words denoting women and men, characteristics that are stereotypically perceived are given without hesitation, and many stereotypes can be found in both definitions, collocations, and examples. Being a woman is thus described mostly negatively: women are inferior to a men, they should be beaten, keep quiet, obey their husbands, and give birth to children. The headword inventory connected to sexuality is limited. Words denoting sexual relations are usually omitted, especially in dictionaries from the oldest period (e.g. Vrančić and Kašić). Sexual intercourse is defined as something sinful, and, if it is between two men, unnatural. The intercourse between two women is not commented on. The headword inventory of taboo and vulgar words is also limited. If included, the headwords are usually defined in Latin (and sometimes in German). Although the headwords are not censored, the vulgar words in the rest of the entry usually are. In the entries with vulgar headwords, usually, no usage examples are given. The ethnic groups geographically or historically closer to Croatia are described in more detail. The entries reflect historical relations (e.g. Turks are described as warriors and enemies) or prevailing stereotypes (e.g. Gypsies are described as lying, lazy, thieves, etc.).

In contemporary dictionaries, the stereotypes are more subtle than in historical dictionaries – which implies a change in the worldview over time, but also a change of the lexicographic approach towards socially sensitive content. Moreover, the dictionary material has shown that social awareness is not equally present in the treatment of vocabulary in all domains, which can be due to the fact that some of them (e.g. ethnicity), are perceived as more sensitive than others. An overview of the results of our analysis is presented in Table 1.

Although a progress can be noticed when the contemporary Croatian dictionaries are compared with historical dictionaries, and even within the group of contemporary dictionaries – the newest one being more socially sensitive than the older ones – the examples presented in this paper show that there is still room for improvement. The analysis has revealed the following elements that should be revised: 1) definitions should be checked for subjectivity (*worthy of a real man*), outdated perceptions (*typical work not suitable for woman*), sexism (*woman with her peculiar way of thinking*), etc.;

<sup>9</sup> An exception is the entry *ciganluk* in HJP, defined as *an ugly act of a kind that is according to the prejudice attributed to Gypsies* [“ružan postupak kakav se prema predrasudama pripisuje Ciganima”].

a comment of a stereotype lexicalized in an item can be considered (as it is sometimes done in entries concerning ethnic groups, but rarely other, e.g. *womanish*, *mannish*); 2) definitions and examples should in some cases be more inclusive, e.g. words denoting relationships and sexuality are often approached from a heteronormative perspective; 3) unnecessary stereotyping should be avoided in examples and collocations (*masculine policewoman*), 4) lists of collocations and expressions should be revised and updated, 5) labels should reflect the actual usage (they are sometimes missing or are too mild); both direct and indirect offensiveness could be considered and indicated by labels or other means, 6) synonyms / cross-references should be reviewed – for offensive items, neutral items should be given and offensive items, if listed within other entries, should be labelled. Modern e-dictionaries, often published online, have innovative features which can be useful for discovering and describing socially sensitive content. For example, the absence of space limitation allows for a more detailed description, explanatory notes, more examples to illustrate different contexts, etc. The Internet as a medium makes it easier to edit the dictionary data and enables communication with users, who can provide useful information on potentially sensitive content.



	Historical dictionaries		Contemporary dictionaries	
	characteristics	examples	characteristics	examples
male/female	<ul style="list-style-type: none"> <li>– typical male/female characteristics mostly taken as given</li> <li>– stereotypes found in definitions, collocations, and examples</li> <li>– being a woman often perceived as negative or inferior to being a man</li> <li>– potentially offensive items usually not labelled as such</li> </ul>	<ul style="list-style-type: none"> <li>– woman – a person whose organism is designed to give birth; someone acting cowardly</li> </ul>	<ul style="list-style-type: none"> <li>– typical male/female characteristics sometimes taken as given</li> <li>– stereotypes found in definitions, collocations, and examples</li> <li>– potentially offensive items sometimes not labelled as such</li> <li>– progress has been made in the newest dictionary</li> </ul>	<ul style="list-style-type: none"> <li>– “male hand”: <i>a man who in a household does the typical work which is not suitable for a woman</i></li> <li>– masculine policewoman</li> <li>– pretty woman vs. brave man</li> <li>– fighting for woman’s rights</li> </ul>
sexuality	<ul style="list-style-type: none"> <li>– the headword inventory is very limited</li> <li>– sexual relations (especially of the same sex and outside marriage) described as unnatural, shameful, and sinful</li> </ul>	<ul style="list-style-type: none"> <li>– sodomy – unnatural sexual intercourse</li> </ul>	<ul style="list-style-type: none"> <li>– the headword inventory is more inclusive</li> <li>– offensive headwords and cross-references are not properly labelled</li> <li>– different view on male and female sexuality</li> <li>– definitions of words denoting relationships and sexuality should be more inclusive</li> </ul>	<ul style="list-style-type: none"> <li>– “male virgin”: <i>the one who is living as a virgin, who has renounced or is deprived of the touch of a woman</i></li> <li>– “mental prostitute”: <i>a woman who takes advantage of a man and keeps him hoping he will be successful</i></li> </ul>
taboo words	<ul style="list-style-type: none"> <li>– the headword inventory is very limited</li> <li>– no Croatian definitions are given</li> <li>– equivalents in Latin (and German)</li> <li>– vulgar word censored</li> </ul>	<ul style="list-style-type: none"> <li>– penis – Latin translation: <i>mentula</i></li> <li>– fucking – the act of f...ing</li> </ul>	<ul style="list-style-type: none"> <li>– listed without hesitation in 2 out of 3 dictionaries</li> <li>– vulgar expressions labelled as such</li> </ul>	<ul style="list-style-type: none"> <li>– pizda (‘cunt’), vulg.: <i>the word is very rude and inappropriate in polite communication, in texts it is usually abbreviated to p...</i></li> </ul>
ethnicity	<ul style="list-style-type: none"> <li>– larger entries with more examples for the ethnic groups historically closely related</li> <li>– stereotypes found in definitions, collocations, and examples</li> </ul>	<ul style="list-style-type: none"> <li>– Turks – warriors, enemies, non-believers...</li> </ul>	<ul style="list-style-type: none"> <li>– treated more cautiously than other domains</li> <li>– stereotypes are sometimes commented on</li> </ul>	<ul style="list-style-type: none"> <li>– “Bosnian pot”: <i>intricate political circumstances typical for Bosnia (according to prejudices outside of Bosnia)</i></li> </ul>

Table 1: The treatment of vocabulary from the domains male/female, sexuality and taboo words, and ethnicity in historical and contemporary dictionaries – main features.

## 6 Sources

- ARj = *Rječnik hrvatskoga ili srpskoga jezika* (1880-1976). Zagreb: Jugoslavenska akademija znanosti i umjetnosti.
- Habdelić, J. (1670). *Dictionar ili reči slovenske*. Accessed at: [http://croqip.ffzg.hr/default\\_hr.aspx](http://croqip.ffzg.hr/default_hr.aspx) [15/03/2021].
- HJP = *Hrvatski jezični portal*. Accessed at: <http://hjp.znanje.hr> [15/03/2021].
- Jambrešić, A. (1992). *Lexicon Latinum interpretation Illyrica, Germanica et Hungarica locuples*. Zagreb: Hrvatski filološki institut.
- Kašić, B. (1999). *Hrvatsko-talijanski rječnik s Konverzacijskim priručnikom*. Zagreb: Kršćanska sadašnjost.
- Mikalja, J. / Gabrić-Bagarić, D., Horvat, M., Lovrić Jović, I. & Perić Gavrančić, S. (2011). *Blago jezika slovinskoga*. Zagreb: Institut za hrvatski jezik i jezikoslovlje.
- RCJHR = *Rječnik crkvenoslavenskoga jezika hrvatske redakcije* (2000-2018). Zagreb: Staroslavenski institut.
- RHJ = Šonje, J. (2000). *Rječnik hrvatskoga jezika*. Zagreb: Leksikografski zavod Miroslav Krleža – Školska knjiga.
- Slovník = Slovník jazyka staroslověnského* (1966-1997). Praha: Akademia.
- Stulli, J. (1801). *Lexicon latino-italico-illyricum*. Budapest: Regia Universitas Pestana.

## 7 References

- Adams, J.N. (1982). *The Latin Sexual Vocabulary*. London: Duckworth.
- Allan, K., Burridge, K. (2006). *Forbidden words: taboo and the censoring of language*. Cambridge: Cambridge University Press.
- Atkins, S.B.T., Rundell, M. (2008). *The Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press.
- Bratanić, M. (2005). Mjesto žene u rječniku. In D. Stolac, N. Ivanetić & B. Pritchard (eds.) *Jezik u društvenoj interakciji*. Zagreb – Rijeka: Hrvatsko društvo za primijenjenu lingvistiku, pp. 37-46.
- Cloete, A.E. (2014). The treatment of sensitive items in dictionaries. In R.H. Gouws et al. (eds.) *Dictionaries. An International Encyclopedia of Lexicography. Supplementary Volume: Recent Developments with Focus on Electronic and Computational Lexicography*. De Gruyter Mouton, pp. 482-486.
- Coffey, S. (2010). Offensive items, and less offensive alternatives, in English monolingual learners' dictionaries. In A. Dykstra, T. Schoonheim (eds.) *Proceedings of the XIV EURALEX International Congress*. Leeuwarden, pp. 1270-1281.
- Dakić, M. (2017). Kada ćemo postati žene u rječnicima hrvatskoga jezika? In *Jat: časopis studenata kroatistike*, 1(3), pp. 106-123.
- Dykstra, A. (2006). J.H. Halbertsma: 'Sexual Language and the Lexicon Frisicum (1872)'. In *Dictionaries: Journal of the Dictionary Society of North America*, 27, pp. 21-35.
- European Institute for Gender Equality (2021). Sexuality. In *Glossary & Thesaurus*. Accessed at: <https://eige.europa.eu/thesaurus/terms/1379> [11/06/2021].
- Farø, K.J., Jensen, L.V. (2018). Den Danske Ordbog på nettet: en undersøgelse af version 3.0. In *LexicoNordica*, 25, pp. 215-232.
- Fjeld, R.V. (2015). Om ordbokseksempler og stereotypisering av kjønn i noen nordiske ordbøker. In C. Sandström et al. (eds.) *Perspektiv på leksikografi, grammatik och språkpolitik i Norden*. Helsinki: Institutet för de inhemska språken, pp. 35-65.
- Fournier, H.S., Russell, D.W. (1992). A study of sex-role stereotyping in the Oxford English Dictionary 2E. In *Comput Hum*, 26, pp. 13-20.
- Gorjanc, V. (2004). Politična korektnost in slovarski opisi slovenščine – zgolj modna muha? In Stabej, M. (ed.) *Moderno v slovenskem jeziku, literaturi in kulturi: zbornik predavanj / 40. seminar slovenskega jezika, literature in kulture*, 28. 6.–16. 7. Ljubljana: Center za slovenščino kot drugi/tuji jezik pri Oddelku za slovenistiko Filozofske fakultete, pp. 153-161.
- Gorjanc, V. (2005). Neposredno in posredno žaljiv govor v jezikovnih priročnikih: diskurz slovarjev slovenskega jezika. In *Družboslovne razprave*, 21(48), pp. 197-209.
- Harteveld, P., van Niekerk, A.E. (1996). Policy for the Treatment of Insulting and sensitive Lexical Items in the Woerdeboek van die Afrikaanse Taal. In M. Gellerstam et al. (eds.) *Proceedings of the 7th EURALEX International Congress*. Göteborg, pp. 381-393.
- Janschewitz, K. (2008). Taboo, emotionally valenced, and emotionally neutral word norms. In *Behavior Research Methods*, 40(4), pp. 1065-1074.
- Jay, T. (1977). Doing research with dirty words. In *Maledicta: The International Journal of Verbal Aggression*, 1(2), pp. 234-256.
- Jay, T. (1992). *Cursing in America*. Philadelphia: John Benjamins.
- Jay, T. (2000). *Why we curse: A neuropsychosocial theory of speech*. Philadelphia: John Benjamins.
- Jay, T. (2009). The Utility and Ubiquity of Taboo Words. In *Perspectives on Psychological Science*, 4(2), pp. 153-161.
- Jay, T., Caldwell-Harris, C. & King, K. (2008). Recalling Taboo and Nontaboo Words. In *The American Journal of Psychology*, 121(1), pp. 83-103.
- Jensen, J., Lorentzen, H., Nimb, S., Svendsen, M.-M.M. & Trap-Jensen, L. (2018). Thai piger, muskelhunde og fulde svensker: nedsættende ord, stereotyper og ligestilling i Den Danske Ordbog. In *Nordiske Studier i Leksikografi*, 14, pp. 141-151.
- Koopman, A. (2016). Ethnonyms. In C. Hough (ed.) *The Oxford Handbook of Names and Naming*, pp. 251-262.
- Libela* (2013). Jezik pokazuje rodnu neravnopravnost. Published: 09/10/2013. Accessed at: <https://www.libela.org/sa-stavom/4326-jezik-pokazuje-rodnu-neravnopravnost/> [15/06/2021].
- Lebsanft, F. (1997). Sprachtabu und Euphemismus in der französischen Sprachgeschichte. In G. Holthus, J. Kramer & W.

- Schweickard (eds.) *Italica et Romanica*, Festschrift für Max Pfister zum 65. Geburtstag. Band 3. Tübingen: Max Niemeyer Verlag, pp. 111-131.
- Ljubešić, N., Klubička, F. (2016). *Croatian web corpus – hrWaC 2.1*. Accessed at: <http://hdl.handle.net/11356/1064> [26/02/2021].
- Mills, S. (2008). *Language and sexism*. Cambridge – New York: Cambridge University Press.
- Moon, R. (2014). Meanings, Ideologies, and Learners' Dictionaries. In A. Abel, C. Vettori & N. Ralli (eds.) *Proceedings of the XVI EURALEX International Congress: The User in Focus*. Bolzano, pp. 85-105.
- Norri, J. (2000). Labelling of derogatory words in some British and American dictionaries. In *International Journal of Lexicography*, 13(2), pp. 71-106.
- Norri, J. (2019). Treatment of words for illness and disability in monolingual English dictionaries. In *International Journal of Lexicography*, 33(3), pp. 227-250.
- Petersson, S., Sköldbeg, E. (2020). To discriminate between discrimination and inclusion: a lexicographer's dilemma. In Z. Gavriilidou, M. Mitsiaki & A. Fliatouras (eds.) *Proceedings of XIX EURALEX Congress: Lexicography for Inclusion, Vol. I*. Democritus University of Thrace, pp. 381-386.
- Pišковиć, T. (2017). Perpetuiranje rodnih stereotipa u hrvatskim rječnicima. In *Romanoslavica*, 52(2), pp. 343-363.
- Radtke, E. (1986). Konstanz und Wandel in der Beurteilung von Sexualia in der Geschichte der Lexikographie. In *Osnabrücker Beiträge zur Sprachtheorie*, 35, pp. 107-117.
- Reinhart, J. (1994). Slav. Ружь "membrum virile". In P. Vavroušek (ed.) *Iranian and Indo-European Studies. Memorial Volume of Otakar Klíma*, pp. 219-223.
- Russel, L.R. (2012). This is What a Dictionary Looks Like: The Lexicographical Contributions of Feminist Dictionaries. In *International Journal of Lexicography*, 25(1), pp. 1-29.
- Schutz, R. (2002). Indirect Offensive Language in Dictionaries. In A. Braasch, C. Povlsen (eds.) *Proceedings of the 10<sup>th</sup> EURALEX International Congress*. Copenhagen: Center for Sprogteknologi, pp. 637-641.
- Schweickard, W. (1997). Tabu und Euphemismus in der italienischen Lexikographie. In M. Lieber, W. Hirdt (eds.) *Kunst und Kommunikation. Betrachtungen zum Medium Sprache in der Romania*, Festschrift zum 60. Geburtstag von Richard Baum. Tübingen, pp. 303-310.
- Talbot, M. (2005). Gender Stereotypes: Reproduction and Challenge. In J. Holmes, M. Meyerhoff (eds.) *The Handbook of Language and Gender*. Oxford: Blackwell Publishing, pp. 468-486.
- The Herald* (1993). Dictionary withdrawn in war of words over 'Bangkok prostitutes' entry. Published: 06/07/1993. Accessed at: <https://www.heraldscotland.com/news/12730658.dictionary-withdrawn-in-war-of-words-over-bangkok-prostitutes-entre/> [15/06/2021].
- Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. Amsterdam – Philadelphia: John Benjamins Publishing Company.
- Trojar, M., Žagar Karer, M. (2013). Družbena občutljivost v terminoloških slovarjih. In A. Žele et al. (eds.) *Družbena funkcijskost jezika (vidiki, merila, opredelitve)*. Ljubljana: Univerza v Ljubljani, Filozofska fakulteta, pp. 457-463.
- Wodak, R., Benke, G. (2000). Gender as a Sociolinguistic Variable. In F. Coulmas (ed.) *The Handbook of Sociolinguistics*. Oxford – Malden: Blackwell Publishers, pp. 127-151.

### Acknowledgements

This paper was written within the research projects *Croatian Web Dictionary – MREŽNIK* (IP-2016-06-2141) and *Retro-digitization and Interpretation of Croatian Grammar Books before Illyrism* (IP-2018-01-3585), financed by the Croatian Science Foundation.