A corpus analysis of five neologisms from the area of gender and sexuality studies.

Buntak, Ivana

Master's thesis / Diplomski rad

2020

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, University of Zagreb, Faculty of Humanities and Social Sciences / Sveučilište u Zagrebu, Filozofski fakultet**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:131:717977

Rights / Prava: In copyright

Download date / Datum preuzimanja: **2021-05-12**

University of Zagreb

Faculty of Humanities and Social Sciences

Department of English

Academic year 2019/2020

A corpus analysis of five neologisms from the area of gender and sexuality studies

Graduation thesis

Ivana Buntak

Mentor: Marina Grubišić, PhD

Zagreb, September 2020

Abstract

This study deals with five sample neologisms from the area of gender and sexuality studies and their corpus analysis. The aim is to analyse the new words that enter language from those particular areas, as there has been an influx of gender and sexuality neologisms in the recent years. In this study the aim is to analyse these words in an English language corpus, but also in a Croatian language corpus as the thesis is that these neologisms are more frequently used in more registers and with fewer clarifications of the meaning in English, where there have been present tendencies to create a more gender equal and sexuality inclusive language. The objective is also to analyse the linguistic characteristics of these words, mainly their word formation processes, but in order to understand why there is a need in language for these words, a sociolinguistic analysis of the connections between language and sexuality and language and gender needs to be made and explained as well. The study also analyses the context in which the words are used and the collocations they make part of, and finally, the registers in which the neologisms are used. The more frequently the word is found in a corpus and the more registers it is found to be used in, it is considered to be the more accepted in the target language.

Keywords: corpus analysis, neologisms, gender, sexuality, word formation processes

Contents

1. Introduction

In 2018, news media in America wrote extensively about a certain coinage used by the Canadian PM, Justin Trudeau. During a school visit he had made at the time, he happened to interrupt a student asking him a question to tell her to use the noun 'peoplekind' instead of 'mankind', which was the noun she had used. He explained that the coinage 'peoplekind' was "more inclusive" than the commonly used 'mankind'.[1] Many discussions for and against the use of these kinds of word formations followed. As shown by this example, the area of language which is concerned with gender is often productive with new words in order to achieve gender equality and gender and sexuality inclusive language is a very popular topic for debate in the US and the UK.

The area of gender and sexuality is nowadays clearly changing, especially in relation to language (the English language is leading in this change), and while some are in favour of creating a gender equal and sexuality inclusive language, others criticize such endeavours. Even though this study is a linguistic one, it is not enough to study language alone, as language has to be studied in terms of how it relates to the society in which it is used, and the conventions accepted in that particular society. In fact, in studying this topic, one has to consider many areas, as Frederick Attenborough (2014:137) states in the introduction of his study on gender and language:

"Research into the many possible relationships, intersections and tensions between gender and language is diverse. It crosses disciplinary boundaries, and, as a bare minimum, could be said to encompass work notionally housed within gender studies, linguistics, feminist media studies, interactional sociolinguistics, feminist psychology, conversation analysis, media studies and cultural studies."

The aim of the study is to analyse data from two separate corpora based on the search of five words taken from the glossary of gender studies[2] (which includes various types of sexuality) and study the context and frequency of their use, as well as the registers in which they are used, in order to learn whether they can be understood as indicators of a move towards gender equality and sexuality inclusiveness in the language.

---

[1] Canadian leader coins new word: 'peoplekind' (2018). https://theday.co.uk/stories/canadian-leader-coins-new-word-peoplekind (7 September 2020).

[2] Glossary (English). Gender and Sexuality Center. University of Texas.
https://diversity.utexas.edu/genderandsexuality/glossary/english/ (22 September 2020).

2. Theoretical aspects of the study and key terms

2.1. Gender, sexuality and language

The first matter to consider is the matter of gender and sexuality and how it relates to language. Both Croatian and English history of gender studies and terminology need to be considered, if the comparisons are to be made between the two languages.

However, prior to further analyses of the context, the wider setting in which the word appears, outside of language, also has to be studied. The development of gender studies has to be studied as it happened in the US and the UK, but also in Croatia in order to understand whether Croatian and English are languages that are on the way to become more equal and inclusive.

When it comes to English, the changes in the use of language considering gender equality have started in the past century. One of the first steps was the decrease of use of the masculine and feminine 'he' and 'she' and the increased use of 'they'[3] (Balhorn 2004). The origins of the fight for a gender equal language are thus found among feminist activists and theorists. The second feminist wave of the 1970s was highly concerned with terminology and had introduced many changes in language (police officers > policemen, chairperson > chair(wo)man, Ms. as a title for married and single women alike, etc.) (Zimman 2017:87). These examples show that in the fight for gender equality (and later on, sexuality acceptance) language has been made political from the beginning. In the 1990s queer linguistics emerged, when the word 'queer', by then an offensive term referring to homosexual men, has been re-appropriated by the gay community and attained a neutral meaning (Zimman 2017:88). Terminology had expanded to include new words such as 'homophobia' and 'heterosexism' (Zimman 2017:88). Transgender communities also started gaining acceptance in the 1990s. With the recent popularization of social justice in the US and the UK, transgender terminology has also entered all registers of language, such as the newspaper and the academic and the use of the neologisms coming from the language of communities fighting for gender and sexuality inclusion is more supported than ever before.

---

[3] Merriam-Webster chose "they" as its 2019 Word of the Year based on the number of dictionary lookups, and the singular "they" was added to its online dictionary in September. https://www.merriam-webster.com/words-at-play/word-of-the-year/they.

In Croatia, gender studies are not nearly as influential or developed. The terminology is adopted from English in the form of borrowings. Gender or women's studies have only recently been recognized as a scientific area of study in its own right (in 2009). Branka Galić (2009:9) writes that Croatia, like any other country that wants to develop in the contemporary world, aims to achieve gender equality. However, the problem of gender discrimination still persists in Croatia. Whereas in English the problem of terminology concerning 'gender' and 'sex' has been solved some time ago and the two are no longer used synonymously, in Croatia these two words are still used synonymously, which shows that many people in Croatia still recognize only two genders (based on the two sexes – male and female) (Galić 2009:10). According to Galić (2009:17), gender equality has never been as important matter to women in Croatia, as for example, in the US or the UK. She explains it with the fact that Croatia did not have a capitalist society, but a socialist one, and even though many women worked in the socialist country, they were not equal to men in terms of pay, and in Croatia as a developing capitalist country, many women suffered a return to their traditional roles. This explains why a gender equal language may not be as 'popular' in Croatia, as it is in the US and the UK. When it comes to sexuality, Galić (2012:174) writes that there are some positive changes in Croatian society (among younger people with higher education, mainly women), but that society's consciousness remains 'patriarchal and neosexist', with stereotypical views of homosexual groups and women. Earlier research the author mentions in her work showed that in Croatia sexism persisted among men, uneducated people and the older population (Kamenov and Jugović 2009:142; Galić 2004:320). Sexism, in this case, marks patriarchal views of women as of weaker sex, but also intolerance of minor sexuality communities, such as homosexuals.

As was already mentioned, the language of gender studies in Croatia is borrowed from English. Borrowings from English are nothing unusual in Croatian – there has been a great influx of English words ('anglizmi') due to globalization in the 21st century (Opačić 2012). The coinage 'hrengleski' refers to the hybrid language, resembling 'spanglish' or 'franglais' which adds Croatian endings to English roots of words.[4] The reason for borrowings in the area of gender studies is primarily the lexical emptiness that exists in Croatian regarding this field, as in order to name a certain concept or subject, one has to take it from English (Reljac-Fajs, Vidović 2015). The prevailing attitudes towards English borrowings used to be negative,

---

[4] The term 'hrengleski' was coined by Marta Medved-Krajnović and Mirjana Vilke in 2006 ('Do you speak Crenglish?' *Jezik i mediji: Jedan jezik - više svjetova.*)

but more recently are more positive / acceptive due to the fact that it is thought that most of the borrowings are only temporarily in language, as well as the fact that the number of these borrowings is overwhelming in many areas, and it is simply not possible to make up Croatian equivalents for them (Reljac Fajs, Vidović 2015:9). This means that borrowings found in gender studies are not alien in Croatian due to their English base, and if they are not accepted in the language, it is because of some other factors, mentioned above (Reljac Fajs, Vidović 2015). Some recent research (e.g. Caswell, Laakso and Prewiitt 2012), which will be referred to later in this study, also shows that there may be a connection between gender equality in language and gendered languages, like Croatian. Countries which do not have a gendered language, e.g. those were English is spoken, show a greater gender equality, and it is suggested that the reason lies in that it is easier to make a revision of those languages, then the gendered ones. This fits with the line of thinking that language is reflective of conventions of culture, but that language itself also shapes our understanding of the world (Boroditsky 2009; Deutscher 2010 as cited in Caswell, Laakso and Prewitt-Freilino 2011).

"Given recent research tying gender in language to gendered perceptions of the world (e.g., Boroditsky et al. 2003), one could infer that when language constantly calls attention to gender distinctions by discriminating between masculine and feminine nouns and pronouns—as is the case in gendered languages—that individuals may be more apt to draw distinctions between men and women. If, in fact, language plays a role in how people organize their beliefs about gender, then it stands to reason that differences in the gendered language systems across different cultures could play a role in societal differences in beliefs, attitudes, and behavioural practices about the role and status of men and women" (Caswell et al. 2011:269).

Other social factors need to be taken into consideration as well when it comes to the establishment of a more gender equal language, such as religion, politics, geography and the standard of living. However, language-related problem may also be encountered, and one such possible problem with adapting English neologisms to Croatian is that people simply do not understand English that well, as even those who can speak the language very well have trouble with understanding a specific meaning of the word (Opačić 2012 as cited in Reljac Fajs, Vidović 2015:8). This may also present a problem when it comes to understanding complex gender and sexuality terminology.

2.1.1. Language and identity – why language matters so much

Now that it has been demonstrated that English has a much richer tradition of gender and sexuality studies and the fight for a gender equal inclusive language, it is important to explain why so much focus is placed on language, when it comes to the fight for recognition (be it gender or sexuality-wise). The answer lies in the relationship between language and identity.

"It is a longstanding observation in socio-linguistics, that language-using, whatever else it accomplishes, is an 'act of identity', a means whereby people convey to one another what kinds of people they are. Clearly, language-using can fulfil this function in relation to sexual identity as it can in relation to other kinds of identity – (e.g. gender, class, ethnicity)" (Cameron and Kulick 2003:11).

Bucholtz and Hall (2002) also find identity to be one of the key aspects of language and sexuality. They recognize the beginnings of *language and sexuality* area of research in the closely related *language and gender studies* research. Nowadays, the two concepts are studied as separate, but it is still recognized that they need to be discussed together. Bucholtz and Hall (2002) emphasize that it is not enough to explore the language of the marginalized sexualities only, but also of the sexuality that is considered to be the norm – *heterosexuality*. The reasoning behind this is that 'heteronormativity' also shapes language and its use, as it is related to the questions of identity. Language is a powerful tool in the fight for identity – it can either affirm or erase one's existence.

In terms of sexuality, the word 'transgender' has been revolutionary in English as it was the first word marking sexuality other than *hetero* and *homo* in the English language. Nowadays, there are many neologisms referring to various types of sexuality, as sexuality has been made more 'fluid' as well, like gender. Zimman (2017) points out that it was indeed language that had a really important role in the changes that have happened in the US, as the awareness outside of the LGBT and transgender communities about transgender issues has grown. He writes about the growing tendencies to use a more inclusive language, the efforts which are made by many institutions and organisations, in particular university campuses. The example of this is the appearance of the pronoun 'they' at college campuses, where students can choose if they want to be referred to as 'they', in case they do not identify as 'he' or 'she'.

"Indeed, trans activism is often centred around linguistic reform. One recent success, for instance, is the introduction of the word non-binary for reference to individuals who do not self identify as either female or male. Similarly, the word cisgender or cis, which has been in wide use within trans communities for well over a decade, has recently entered the general lexicon of a broader (cis) population" (Zimman 2017:85).

Language is at heart of the battle for transgender rights, according to Zimman (2017), as it reflects the fact that the opposite of *transgender – cisgender* is considered to be the norm ('heteronormative'), and that is why the aim of the fight for trans rights is the reform of language – this reform includes the introduction of more transgender-inclusive terms, and the ending of the use of such terms as 'man' and 'woman' when referring to human bodies. This makes the meaning of words (and its change) the emphasis of the fight for a more gender and sexuality inclusive language. However, the analysis of neologisms with regards to their morphological characteristics is also important, as it can show which word formation processes are the most productive in the area, which helps us understand how these new words, necessary to name concepts regarding sexuality and gender, came to be.

2.2. Neologisms

2.2.1. The importance of neologisms for this study

This study focuses on the word formation process of new words in the area of gender studies, which are coined in order to create a more gender equal, but also a more inclusive vocabulary, concerning the naming and acceptance of various types of human sexuality. From the linguistic point of view the aim is to learn what kinds of word formation processes are most frequently found in this area in English, but in Croatian as well – how these neologisms are used in a language which does not make up its own terminology in the area, but takes it from English. The word formation processes of these neologisms are also interesting to look at as this is currently a very productive area of the language, so the processes used to build words in this area may reflect the current trends in word formation processes in other areas of language where changes are occurring as well, mainly the language of various social media which has entered all areas of language, but especially the language of other media, such as popular magazines and newspapers (e.g. the derivative *influencer*, abbreviation *TL;DR*[5], etc.), and a more recent example of medical neologisms entering language due to the coronavirus pandemic – e.g. the compounds *herd immunity, self-isolate*, abbreviation *WFH (work-from-home)*, etc.), also the popular culture references or related words, such as a derivative *bingeable*. The results of corpus analysis are supposed to show how accepted the neologisms analysed are in these two languages by looking at the diversity of registers and sources in

---

[5] Taken from 'New words added to Merriam Webster in 2019'.
https://www.goodhousekeeping.com/life/g25712934/new-dictionary-words/ (25 September 2020).

which these words are used, along with the frequency with which they appear. The thesis is that English is making changes in this regard by accepting these neologisms into the language, whereas other languages, such as Croatian, may or may not follow suit, depending on a variety of social and linguistic factors.

To sum up, the importance of neologisms lies precisely in understanding that language is not fixed, but fluid with constant changes.

2.2.2. Definitions of 'neologism' and a theoretical background

The term 'neologism' has many definitions. Algeo (1991:2) defines it as a form or the use of a form not recorded in general dictionaries. It can either represent a new use of an existing form or be "a shape that has not before been seen or heard in English" (Algeo 1991:2). This definition is explained by the fact that one cannot know for sure how new any word really is. Ingo Plag (2003:65) defines a neologism as "a derivative newly coined in a period." He relates it to the concept of productivity, which he defines as the possibility of creating a new word (Plag 2003:64). Plag (2003) considers counting neologisms in a given period of time as indicative of the productivity of a certain affix.

According to Roswitha Fischer (1998:3), a neologism is a "word which has lost its status of a nonce-formation but is still one which is considered new by the majority of members of a speech community". The principles important for the understanding of neologisms are frequency, coverage, availability and familiarity (Fischer 1998:4). Coverage concerns the particular topics of discourse, availability refers to the retrieval of words (how long it takes to remember a certain word when it is needed) and familiarity is the subjective impression of whether a certain word is familiar to someone. Learnability also matters when learning new words as it refers to the ease or difficulty of learning a new word. The principles of availability and familiarity are, according to Fischer (1998), of great importance when a native speaker learns a neologism, but as they are not easily measurable by lexicographers who add words to dictionaries, they are not as relevant as coverage and frequency. Finally, semantic priming (same semantic field facilitates word recognition) and neighbouring words (many orthographically or phonetically similar words) are also important factors in the comprehension of neologisms.

Plag (2003) makes a distinction between creative and productive neologisms. The concept of productivity is related to the native speakers' ability to produce an uncountable number of

new formations following a set of rules, whereas creative neologisms are unpredictable. Derivation and compounding are considered to be productive word formation patterns, that is, they create productive neologisms.

Word formation processes of neologisms are one of the focuses of this study. The question that needs to be answered before any analyses and conclusions are made is why they matter. Why is it important which word formation processes are productive and produce new words? According to Ingo Plag (2003), productivity of a certain word-formation process is influenced by the usefulness of newly-coined word for the speakers of the language. He lists three main functions of making up new words:

1. the referential function or labelling (giving a name to a new concept or thing (words can be also coined from other similar words when necessary as in ether > etherize);

2. syntactic recategorization (using single complex words instead of longer phrases or whole clauses);

3. expressing an attitude (e.g. fondness of a person).

The first function is the one that relates the most to the neologisms in this study as they were all derived in order to name some new concepts and ideas.

As the words chosen for the study were all created by derivation, more specifically, prefixation, some classification of prefixes should also be made. There are four types of prefixes in English: prefixes expressing quantity (e.g. *uni, poly, bi)*, locative prefixes (e.g. *circum, intra*), temporal prefixes (e.g. *neo* as in *neoclassical*), prefixes expressing negation (e.g. *non, a).*

2.3. Corpus analysis and language registers

Corpora are useful tools when studying language as they contain a myriad of texts from various registers and varieties of language. In his book on building a corpus and performing a corpus analysis, Meyer (2002) emphasizes that in a corpus analysis it is of greatest importance to first establish a linguistic research question and to afterwards make an analysis of it in a corpus. According to him, corpora are well suited to offer information about linguistic forms, their frequency and use, and not so much for sociolinguistic research questions, as the more subjective sociological parameters are more difficult to analyse (Meyer 2002:18). Corpora are

useful in investigating how new words spread and in determining why some words are more successful in spreading than others. They can also help explain the contextualization process, which refers to the diffusion of neologisms across several domains, which means that a new word is accepted into language. As a corpus is a body of texts from different registers / genres, which are written in different styles, definitions of these terms need to be presented. Biber and Conrad (2009) give an overview of the existing definitions of the terms of register, genre and style which sometimes overlap, and make their own definitions accordingly. For this study the focus is on language registers, and a register is a "perspective on text which combines an analysis of linguistic characteristics that are common in a text variety with analysis of the situation of use of the variety" (Biber and Conrad 2009:2). The assumption is that certain linguistic features (e. g. pronouns) are more commonly used for a certain communicative purpose in a certain situation. The difference of register and genre lies in the fact that for the register, unlike for the genre, the focus is on the conventional linguistic structures used to create a certain text in some variety, whereas style is not about functionality of certain linguistic features in certain situations at all, as it is an aesthetic preference of the author. As Biber and Conrad (2009) conclude, registers are best suited for describing text varieties, because they are least specialized and any type of text can be analysed from this perspective.

It is important, as Biber and Conrad (2009) point out, that a register is an approach to text varieties, not a type of texts in itself and that it is the situation that determines the linguistic characteristics used, not vice versa. Registers of text varieties most commonly found are the spoken (conversational) and the written ones, and the latter include academic prose, newspaper writing and fiction. Finally, when it comes to this study in particular, 'hybrid' text varieties should also be introduced. 'Hybrid' is the term Biber and Conrad (2009) use to refer to registers which belong to written language, but are interpersonal, the way spoken language is. This term covers e-mails, blogs, forums, text-messaging and other similar (and new) forms of communication. Usually, the written register is focused on communication of information rather than establishing personal relationship, which is more characteristic for the spoken register. In this hybrid area, it may not always be so.

2.4 Context

The importance of context is considerable when it comes to the native speakers' interpretation and understanding of new words. Plag states that,

'Linguists who work with native-speaker informants often experience that words or sentences are first rejected by informants because speakers fail to make sense of them, and not because the data violate morphological or syntactic rules of their language. Presented with an appropriate context that provides a possible interpretation, the same informants may readily accept the data presented to them.' (1996:774 in Bauer and Renouf 2001:231).

Context is also an important concept when it comes to understanding how gender and language relate. In this study, context is understood as Attenborough (2014) defines it in his study – not as something fixed in space which is not shaped by the words inside of it, but as something which is built through interaction – the words that are used. The context is where people 'bring gender into meaning' (Attenborough 2014:140). It is important as a setting in which the word appears. The information provided in the context (which can be overt or covert) makes it possible for readers to interpret neologisms, and this is called 'contextual anchoring'. Bauer and Renouf (2001) write about the importance of the writer in providing context, as it is the writer's assumption whether it is necessary for them provide some contextual help to the readers to understand a certain word. The writer's 'help' can also be covert or unconscious. Renouf and Bauer (2001:241) interpret covert information as the "material in the context which may aid in the interpretation of the new hapax, even though it is not obviously provided with that aim in mind." According to Plag (2003), hapaxes are the items within the corpora (the language) with the lowest possible frequency, which indicates that the greatest number of neologisms can be found among hapaxes.

There are three types of overt help according to Bauer and Renouf (2001:236): quotation marks, glosses (full explanations, usually in brackets / dashes) and introductory / following phrases (e.g. the introductory 'so called'), whereas there are more types of covert help, including repetitions (of roots or full words), collocation, semantically-related words (synonyms, antonyms), pragmatics (clues in other areas), parallelism (e.g. contrasts), lexical signals (e.g. sense-related words) and lexical field (words from the same topic area). These terms will be exemplified further in the study, as each word gets analysed in its context within the corpus and if any examples are found.

The analysis of context in which neologisms appear is important for this study as it will show how taken words are into language, as the appearance of words without overt explanations of their meaning points to their more significant acceptance into language. If a word is only used with definitions across all registers of language, it shows that even the writers using the word are aware that it is unknown to most speakers of the language / their audience.

3.Methodology

Because of the limited scope of the study, five neologisms were taken from a glossary of neologisms in gender and sexuality studies[6], as it is frequently updated with new words. The neologisms are: *pansexuality* (n.)*, demisexual* (adj.)*, cisgender* (adj.)*, non-binary* (adj.) *and transphobia* (n.). These neologisms were chosen in specific as some of the others tested returned no results at all in Croatian corpus.[7] All of the neologisms chosen for the study had at least one result in the *hrWac*. As formal translations of these words do not yet exist, they were searched in the *hrWac* as borrowings of the English words – the prefixes remained the same, only the bases were adapted. Thus, in Croatian these terms are the following – *panseksualnost, demiseksualan, cisrodan, nebinaran* and *transfobija*. As Croatian is a gendered language, unlike English, sometimes the search had to be done by searching the word by its female or neutral gender as the word in male gender showed no results (e.g. *cisrodan (m.) – cisrodna (f.), cisrodno (n.)*). The neologisms were first explained by their meaning, then their word formation process was analysed, as well as the lexical items they consist of. This was done in order to learn which word formation process is most frequently found. The corpus analysis consisted of finding results for these neologisms in the corpora of both languages, the collocations they are part of and the register (the sources) in which they appear. Ten per cent of the results found in the English corpus were analysed, as some neologisms had too many results to be analysed in this study, whereas in Croatian no such selection had to be done as the results were too scarce and frequently overlapped. However, some of the results found in English were found to be repetitive, so every result was checked in order to not choose those that were very frequently found.

iWeb corpus is a corpus of English language websites from the UK, US, Canada, Ireland, Australia and New Zealand. It is used for linguistic research and it contains 14 billion words

---

[6] The glossary is maintained by the University of Texas' Gender and Sexuality Center and available at https://diversity.utexas.edu/genderandsexuality/glossary/english (14 September 2020).

[7] As mentioned earlier in the text, a number of words from the glossary were entered into the hrWac, but were not chosen for the study as they returned no results at all. Those words are: *aromantic (hrv. aromantičan), nonconforming (hrv. rodno nepotvrdan), genderqueer (hrv. rodno nenormativan, rodno nastran), heteroromantic (hrv. heteroromantičan).* Even though all of the analysed examples are adaptations and do not have a universally accepted translation / adaptation, some of these examples are assumed to not have any results in the corpus because their translation is more complex (as in *genderqueer*) and the few authors who use this word either leave the English word or create their own translations (*rodno nenormativan, rodno nastran*). The more recent, specific terms, such as *aromantic* and *heteroromantic* returned no results, as well as *non-comforming*, which, like *genderqueer,* has a more complex, non-economical adaptation, that is, it cannot, like *demisexual* or *pansexuality*, just use the same prefixes as in English with a Croatian base.

from about 22 million webpages. *iWeb* corpus can be browsed for word frequency lists, collocates, etc., but it can also be searched by individual words or by phrases. Unlike, for example, *COCA*, it only brings results found on the web, so it is not practical to explore the fiction register. It also does not have instances of the spoken language. The benefit of this corpus is that it does not have entries in American English exclusively, like *COCA*, but includes other English-speaking countries as well.

hrWac is a Croatian web corpus. It currently contains 1,9 billion tokens (total amount of word types) collected from the .hr domain. It also contains collocations for words, a thesaurus which includes synonyms of words and frequency lists. As *iWeb*, it can be searched by typing individual words into the search box. That is how the research for this paper was conducted – five words chosen for this study were entered into both corpora, as adaptations of the English words were entered into *hrWac*, and ten percent of the results found by the corpora were analysed by the context in which they appear, words they appear with, frequency with which they appear in the corpora and the register in which they are found. The difference in the size of the two corpora needs to be taken into account when comparing the number or results for the words searched, as the English corpus contains almost 14 times more words than the Croatian one.

4.Corpus analysis – iWeb and hrWac

As expected, the corpus of English texts, which contains 14 billion words, found more results than the Croatian corpus. Also, most of the words in the research were coined in English and, because of a number of other social factors concerning gender and sexuality ideology, some of which are mentioned in the study, are currently more used in English than in Croatian.

Each of the five words that were analysed returned some results, which varied in number both in English and Croatian corpus, but in the *hrWac* the results for all examined entries were scarce. The first word that was analysed is *pansexuality*.

4.1. Pansexuality

According to the *Merriam-Webster* dictionary, the meaning of this word is the following; 'sexual desire or attraction that is not limited to people of a particular gender identity or sexual orientation'.

The word formation process used to create this word is derivation. In the process, the prefix *pan-* was added to the base *sexuality* (n.), which is also a derivative noun (*sexual + ity*). The prefix *pan-* comes from Greek and its meaning is 'all', according to the *Oxford Dictionaries*. The word *pansexuality* has 226 occurrences in the *iWeb* corpus. Its Croatian translation, *panseksualnost*, returns only three results in the *hrWac*.

The sentences in Croatian occur in the following contexts:

1. **Panseksualnost** - *neke queer osobe često govore da ih privlače osobe u mnogim rodnim varijetetima , dakle mnogim varijetetima*
   Pansexuality – some queer people tend to claim they are attracted to people of various sexual identities, i.e. many identities
2. *Postoje još brojne seksualne orijentacije prema kojima se ljudi identificiraju, npr.* **panseksualnost**, *autoseksualnost, ambiseksualnost…*
   There are many other sexual orientations people identify with, e.g. pansexuality, autosexuality, ambisexuality…
3. *Marlene i Madonna su, prema njoj, istinito ili lažno odavale dojam panseksualnosti.*
   Marlene and Madonna gave off either a genuine or fake air of pansexuality, according to her. (neon.com)

These examples show that where the word is used its meaning has to be overtly explained, because it is recognized by the writers that it is not widely known. The first two sources are from *Wikipedia* and a website belonging to the LGBT community, respectively. The newness of the word is emphasized by the fact that even among the LGBT community, the meaning has to be overtly explained in the context. In the third example, however, the word 'pansexuality' is used without an explanation of its meaning, as a pop culture reference. The source no longer exists, but it used to be a web magazine for women, without any special LGBT rights' promotion. The conclusion when considering other results is that this is not a case of definition not being used (that is, of the writer assuming the audience will understand the word without their help), but rather of an article on pop culture being automatically translated from English, without considering the readers.

In the *iWeb*, the situation is a bit different as there are more entries, but the most frequent entries are also those where the word *pansexuality* collocates with words which carry overt information such as 'word', 'term', and a lot of the entries are, in fact, definitions of the term. Ten per cent of the results were taken in order to analyse the phrases in which the word appears and the collocations of the word. The following phrases occurred most frequently.

1) *Pansexuality is/ means / is supposed to be/ shouldn't be confused with;*
2) *the idea / concept / definition of / quality of / issue of pansexuality;*
3) *refer to pansexuality/ explain pansexuality / understand pansexuality / learn about pansexuality / define pansexuality;*

Some collocations frequently encountered are the following: *embrace one's pansexuality, accept one's pansexuality/ talk about one's pansexuality / be open and proud about one's pansexuality;*

The examples in, analysed according to Bauer and Renouf's (2001) classification of contextual help, 1) represent overt information, as the verbs are employed as introductory to the following clarifications of the neologism. The examples in 2) classify as covert information, as these nouns are used as lexical signals for the neologism (especially nouns like 'idea' and 'concept'). The examples in 3) are verbs that have the role of lexical signals as covert information in the context or serve as introductory phrases as overt information in the context.

From the analysed results, the variety of English in which the neologism is most frequently found is American English (various American newspapers, such as *azcentral, usatoday*). Interestingly, there seem to be quite a few examples of use without any contextual help in the newspaper register (represented by more mainstream media than the closed LGBT community websites and blogs). This seems indicative of the word's acceptance into the language, mainly the US variety of English. However, a closer look into these texts show that they all overlap and refer to the same event that happened, as a certain American celebrity 'came out as' pansexual and thus made the neologism trendy.

> He tells the British tabloid The Sun that he thinks it's " great " that Miley Cyrus has spoken openly about her **pansexuality**, but he won't cop to that himself. (abcnewsradio)

Various websites offering therapy and psychological help (e.g. *psychologistanywhereanytime*) also use this neologism, and they always use the help of glosses to better illustrate the meaning.

> While traditionally bisexuality has been defined as' an attraction to both males and females', it commonly encompasses **Pansexuality**,' an attraction where the gender of the partner is of little or no relevance' (i.e. to male, female, and any other gender identity).

Definitions and explanations of the term are often found in LGBT websites and portals as well, but that is due to the fact there are still discussions among members of the community on what the proper definition of the word is.

> Either **pansexuality** is a distinct pattern of attraction to bisexuality (something I'm happy to get behind), or else **pansexuality** and bisexuality are simply two ways of saying the same thing- in which case a hell of a lot more justification is needed for the assertion that they are two entirely different orientations, as she claims.

To sum up these examples, the neologism 'pansexuality' seems to be taken by the registers of the US variety of English other than those belonging to the LGBT community (particularly various pop culture magazines and tabloids) due to a number of famous American musicians identifying with it. However, as it is much less frequently used in other varieties of English, this may be an example of using and recognizing the word that is 'trendy' in a given period of time. Even among the LGBT communities and their support groups, the meaning of the neologism is still subject to discussions. Some even go so far as to claim pansexuality is not even real as it threatens identities of people identifying with some other sexualities, such as bisexuality.

*Polysexuality is the same as **pansexuality**.*

***Pansexuality** means that you are attracted to all genders.*

*Polysexuality is almost like **pansexuality**, but you are not attracted to all genders, only a few…*

*Some define being **pansexua**l as including being attracted to trans people.*

*As I understand **pansexuality** is a branch that comes from bisexuality…*

One can also notice, in some of these examples, that other similar terms, such as bisexuality, are used in order to help define this new formation by saying what it does not refer to, in contrast to these other terms. This an example of covert context of using simultaneously semantically related words and parallelism in order to help readers understand the meaning.

It is also important to mention that no example of uses of the neologism with any negative connotations was found. It is either used neutrally or positively as in the examples listed in 4).

4.2. Transphobia

The second neologism analysed is *transphobia (n.)*. The word is made out of prefix *trans-* (Latin) and the noun *phobia* (Greek). In the iWeb it has 2289 results. The meaning, as described in *Merriam Webster* is, irrational fear of, aversion to, or discrimination against transgender people. The first use of the word was recorded in 1992. The word has a negative meaning and it is used to express disapproval or criticism of practices and attitudes which are seen as being harmful or offensive to transgender people.

The corpus analysis finds the following verbs to collocate with *transphobia*: *eliminate; exhibit; confront, face; study; respond to; fight; tackle; eradicate; experience.* Most of these words express the negativity of transphobia and present it as undesirable in the society.

When it comes to noun phrases, they are the following; *the issue of, the cause of, the threat of, the result of, accusations of, claims of, expressions of, campaign against.* These nouns act as lexical signals and they are also used to offer readers contextual help, not so much with the definition of the neologism, but to denote the negativity of the term (as in 'the threat of transphobia'). Some of these words are neutral introductory phrases.

Some of the typical collocations found are: 'run from transphobia', 'accuse somebody of transphobia'. The meaning is covertly indicated as negative in the collocations, as is overtly in the common phrases.

Some of the adjectives frequently used to describe the word *transphobia* are *rampant, deeply entrenched, over the top, ignorant, embarrassing*. These are also lexical signals, analysed according to Bauer and Renouf (2001), intentional or unintentional from the writer, and they emphasize to the readers the danger and depth of the presence of *transphobia* in the society. The neologism is used with other neologisms from the same area which also have negative meanings, such as *homophobia*.

> Sexism, racism, classism, ableism, homophobia, **transphobia**, and hate speech have no place here. (feministing.com)

> We stand in solidarity with folx who fight for justice and work towards freeing all people from the social constructs of oppression, including but not limited to anti-Blackness, heteronormative patriarchy, homophobia, **transphobia**, racism, colonialism, sexism, ableism, capitalism, and imperialism. (chicagoist)

As for the variety of English where the word most frequently appears, it is the US and the UK variety of English. A significant number of results comes from the .uk domain in specific, but

the neologisms has spread to all varieties of English, as there are results from the Australian, Canadian and Irish varieties of English as well. The word is also used across all registers of English – most significantly the newspaper (tabloids, broadsheets and magazines) and the academic register (colleges and universities, mostly across the UK and the US).

> Chandler's transgender father is often referred to as her birth name' Charles' rather than her new name' Helena Handbasket'. Incorrect labels are used when the cast are talking about her, lots of jokes are made about her gender and whether she has a penis, and there is definitely a bigger issue throughout the sitcom with **transphobia**. (Cosmopolitan.uk)

The example above shows that the word is accepted by the speakers of the language, as the word is used without overt clarifications in this popular magazine for women. Many other instances of such use were found in newspaper and academic writing, as well as in the written conversational register, such as blogs). The writer of the article obviously regards the contextual information to be enough even for those readers who may not be aware of the neologism's meaning.

A significant number of results was also found in the Croatian corpus, as there are 264 results found for the Croatian equivalent of *transphobia - transfobija*. Interestingly, this is the only neologism that has a significant score in the hrWac. As it is also the only neologism that carries a negative connotation, as its meaning is 'dislike of or prejudice against transgender and transsexual people', it may be significant that so far the only word which has gained wider recognition in Croatian texts is the word used to describe a negative feeling. This may be reflective of the attitudes to gender equality and LGBT acceptance in Croatia, which are still overwhelmingly negative. This assumption may also be further confirmed by how rarely the other neologisms, coined to represent various types of sexuality and explored in this study, are used.[8]

Ten per cent of the results found were taken for the research.

The following examples show the most frequent phrases and collocations found:

> *internalizirana transfobija* – internalized transphobia;

---

[8] Most of the results are from 2013, which may be related to the fact that it was the year in which the government wanted to make sex education obligatory in schools, which caused the Catholic conservatives, as well as lots of common people in Croatia to fight against that and they blamed the LGBT community for that decision. This caused many discussions in the media about hateful feelings towards these groups, which explains the increased usage of English borrowings – *transphobia, biphobia, homophobia*, etc.

borba protiv homofobije i transfobije / suzbijanje / eliminiranje / proces iskorenjivanja – the fight against homophobia and transphobia, elimination, eradication, erasure;

odupirati se transfobiji / obračunati se s – fight transphobia ;take measures against

homo / bi / transfobija – homophobia / biphobia / transphobia ;

javna osuda transfobije i homofobije – public disapproval of homophobia and transphobia ;

problem homofobije i transfobije u društvu / Hrvatskoj - problems of homophobia and transphobia in society / Croatia ;

homofobija i bifobija u Crkvi - homophobia and biphobia of the Catholic church

As stated above, all of these examples have negative connotations and indicate there is a problem in Croatian society with transphobia and other phobias towards people of different sexualities and sexual preferences than what is deemed as normal. Language is negative as it is a reflection of the attitudes of the society. The registers most frequently found are the hybrid register (blogs, web forums) and the newspaper register (e.g. *večernji.hr, antenazadar, glasistre*). The use of the word is also, unlike in English, used in order to criticize not what the neologism stands for, but the fact that a word with such a meaning exists in the first place by some sources, usually in support of or belonging to the Catholic Church, which is still very influential in Croatia. The example below illustrates this, as the word 'perversion' is used to denote the attitude towards those taking part in the fight against transphobia.

*...zahtijevajući cenzure i kaznene progone za sve koji drukčije misle. Svjetski dan borbe protiv homofobije i **transfobije** podsjeća upravo na takvu perverziju.(zupa-sveti-nikola)*

*...demanding censorship and criminal prosecution for all those who think differently. International Day of fight against homophobia and **transphobia** is reminiscent of such perversions.*

4.3. Non-binary

Another derivative analysed, *non-binary[9]*, is an adjective. The meaning of this word is 'not relating to, composed of, or involving just two things.' However, in the area of gender studies,

---

[9] In English, there is another formation which refers to the same meaning as *non-binary*, and that is another neologism which developed from *non-binary – enby* (adj). According to *Dictionary.com*, *enby* developed from the phonetic pronunciation of *NB*. *NB* is an abbreviation (initialism) of *non-binary*. The origin is attributed to *Tumblr*, a social media platform known for being active when it comes to fighting for gender equality and transgender rights. Why the phonetic pronunciation is used in word spelling instead of 'NB' is not completely clear, but one LGBT and social justice blog commentary offers a possible explanation, 'I use "enby" because

its meaning denotes that a person does not identify with being of either of the two genders (male and female), but as of being of some other gender instead, or has no gender at all. Lexicological analysis of the word shows that it is the result of the word formation process of derivation, as the negative prefix *non-* (Latin) was added to the base *binary* in order to create this adjective. The word is polysemous, as it has a pre-existing meaning in the Information Technologies area, but in the *iWeb* there were scarcely any results found for this meaning.

Alternatively, this word can also be spelled without a hyphen as *nonbinary*, but that spelling has only 140 results in the corpus, which shows that the hyphenated version is more in use with 2322 results. It is frequently used along with the following nouns, as their modifier (adj.): *non-binary people, individuals, pronouns, identities, gender*, etc. It is often combined with the adjective *trans*, which is a clipping of *transgender*, so that a person is marked as *non-binary trans*, or using the full word *transgender*, as *transgender non-binary*. However, it is not exclusively *transgender* people that are classified as *non-binary*. Some of the verb phrases which commonly collocate with this adjective are: *identify as, come out as, recognize, be / decide to be*. The number of results in the corpus, particularly the number of entries which have no explanations of the meaning of the word, and the variety of registers in which it is used without definitions, show that *non-binary* has already entered mainstream use and is deemed to be understandable to the general public. Some of the examples that confirm this are the following:

A guardian can be a mom, a dad, a **non-binary** parent, a grandparent, an aunt, an uncle, a pet owner…' (newtondailynews.com)

'Black and Brown performers, queer **non-binary** performers, alternative artists and designers-these are all centred in the Brooklyn scene.'(dianerehm.org)

In these examples, the covertly expressed information help identify the neologism (using the same lexical field or pragmatics – clues in the other areas, as in the second example). The employment of lexical signaling also helps clarify the meaning of the neologism, as in the first example (mom, dad).

As for the register, the adjective is mostly found in the academic writing and the newspaper register, as well as the hybrid register of blogs and forums (the conversational written register).

"NB" already means several things, including "non-black" as in NBPOC ("non-black people of color)'. (http://www.anamardoll.com/2018/02/storify-why-i-use-enby-and-not-nb.html). *Enby* is thus a creative neologism derived from another neologism. The word has 22 entries in the iWeb and it is mostly used by the LGBT community (positively) in their own discussion boards and forums.

The meaning of this word is defined in a fluid manner, that is, as if it is not a strict category of sexuality but a spectrum. This is shown by the examples where it is used in such phrases as 'be non-binary enough'.

Croatian equivalent of *non-binary, nebinaran*, has only four results in the *hrWac*. However, only two entries have their meanings actually related to gender identities, because the word is also used in computing, as it is used in English as well. The word is used by the *queer.hr* website, which is a blog committed to queer people and all that may be of interest to them. The word is used as an adjective and it is a literal translation (adaptation) of the English word. It is used to describe a society ('non-binary society') and to denote gender identities ('non-binary gender identities').

> *Blogerica MilitantBarbie opisala je svoje iskustvo sudjelovanja na radionici o **nebinarnim** rodnim identitetima gdje se govorilo o arodnosti.* (queer.hr)

> Blogger MilitantBarbie described her experience of taking part in a workshop on **non-binary** gender identities…

> *...ideja o gotovo bespolnim, normama nesputanim bićima donosi viziju nekog alternativnog **nebinarnog** društva.* (queer.hr)

> …the idea of some genderless unrestrained beings offers a vision of an alternative, **non-binary** society.

According to the *Teaching Transgender Toolkit,* the definition of *non-binary* is 'a continuum or spectrum of gender identities and expressions, often based on the rejection of the gender binary's assumption that gender is strictly an either/or option of male/men/masculine or female/woman/feminine based on sex assigned at birth. Words that people may use to express their nonbinary gender identity include "agender," "bigender," "genderqueer," "genderfluid," and "pangender."' (Green, Maurer 2015). This shows that in spite of this neologism's acceptance, it does not yet have a universally accepted meaning – it is unclear whether it denotes a person which does not identify as male or female (as they are transgender or intersex, for example), or also persons who are male or female, but who simply reject the assumption that there are two genders out of ideological reasons.

4.4. Cisgender

Cisgender is another word created by derivation, and it consists of a Latin prefix *cis-* and the noun *gender*, which is the root of this word. Interestingly, the prefix was taken for its 'older', Latin meaning, 'on this side of', as since the late 19th century *cis* has been an adjective, a standalone word, with a meaning in chemistry. Nowadays, the same way *trans* can be used as an individual noun standing for *transgender*, so is *cis* (the clipped form) used to stand for a cisgender person. The clipped form is even used to create new words, as in a neologism blend 'cishet' (someone who is cisgender and heterosexual). The definition of the neologism in the glossary is the following, 'denoting a person whose gender identity is the same as the sex assigned to them at birth'. Its Croatian translation, *cisrodan*, has only one result in the *hrWac*, where it is used as an adjective, modifying the noun *people* ('cisgender people'). It was used by *iskorak.hr*, which is a Croatian organization dealing with the protection of human rights, especially the rights of gay, lesbian, bisexual and transsexual people.

Seksualna orijentacija kod transrodnih osoba varira kao i kod cisrodnih osoba.

Sexual orientation of transgender people varies the same way as with cisgender people.

As for the *iWeb, cisgender* has 1488 results in that corpus. The adjective *cisgender* commonly occurs as a modifier of the following nouns: *men, women, allies, friends, male, female, person, actor, actress, individuals, body, boy, population, norms*, etc. The verbs it commonly collocates with in the corpus are: 'identify as', 'pass as', 'include (cisgender people)', 'be cisgender', 'get/be called cisgender'. The adjective is used by various academic sources, such as *Routledge Textbooks* and news sources, such as *The Huffington Post* and *Metro Newspaper UK,* but mostly it is used by the LGBT community's blogs and websites and feminist websites. For example, *Routledge* uses this neologism when referring to the reading suggestions for diversity, asking students to write about their 'status as cisgender or transgender'. Other educational institutions, such as *Vassar* school for women's studies also use the neologism. It is also used in certain political contexts, mainly on the websites of the political parties fighting for the rights of minorities. The word is frequently used in opposition to *transgender,* which is its antonym, this is a form of expressing covert information in the context in order to signalize the meaning. Sometimes, the word is clipped to *cis*, which then represents a standalone word with the same meaning as *cisgender*. Out of the 1488 results found in the *iWeb* corpus, ten per cent was analysed in order to find the context and phrases within which this term appears.

…obstacles " that can come with courtship, particularly when the other partner is a cisgender (nontrans) man. (salon.com)

It is common, as illustrated in the example above, to still include a clarification of the meaning (here in the form of a gloss) in the more mainstream media, which shows the assumption of the writer that not everyone may be aware of the neologism's meaning.

The neologism is frequently used with other words denoting 'privilege' to the LGBT and minority communities, especially in the US.

Those with privilege (Christian, able-bodied, cisgender, heterosexual, white, male, etc. (collegian.com)

The variety of English most commonly found in the results is American English (US) and its newspaper register (*The Nation, Salon, The Advocate*, etc.), as well as LGBT community's websites, but the neologism is found to be in use across all varieties of English.

4.5. Demisexual

Demisexual is another derivative adjective, constructed of a prefix *demi-*, which also comes from Greek, and the base *sexual*. The meaning of *demi* is half, as in *demi-god*. This word is used to describe a person who only experiences sexual attraction after forming a strong emotional attachment and, on the sexuality spectrum, falls somewhere near asexual, which denotes people experiencing no sexual attraction at all. However, as this neologism is mostly used by the LGBT community on their blogs and forums, it does not seem to have a definite meaning yet.

I am often kind of **demisexual**: not completely… (emptyclosets.com)

I'm a bit complicated so bear with me I am a **demisexual**, androgynous female that is attracted to females... whew! (lgbtchat.net)

…**demisexual** is, according to some, a variant of asexuality (lgbtchat.net)

For example, In the Croatian corpus, this word hits only one result. The corpus result is the following excerpt.

Npr. čovjek tako može biti demiseksualan, što bi značilo da ga privlače tek osobe s kojima ostvari prethodnu emocionalnu vezu. (tportal)

A person can thus be demisexual, which means they are only attracted to people to whom they had previously established an emotional attachment.

This shows that it this neologism is not really used (almost) at all in Croatia yet.

In the *iWeb*, there are 226 results for 'demisexual'. LGBT portals and websites such as *emptyclosets.com* and *asexuality.org* are the most frequently found sources for this word, but there are also some more mainstream sources, such as *huffingtonpost.com* or *irishexaminer.com*, also using this word. However, in these sources, the word is mentioned along with a gloss (explanation) which shows that its meaning is not considered to be widely recognized by the writers of the articles in this media.

Experts also identify another category, **demisexual** (someone who only experiences sexual attraction to a person after forming a strong emotional bond with them). (irishexaminer)

Most commonly, 'demisexual' is used as part of 'identify as' phrase, as well as 'come out as.' It is often modified by other neologisms or used in the same context as other similar words, which is a pragmatic manner of expressing covert contextual information – e.g. 'a heteroromantic demisexual' or a 'sapiophilic demisexual', and in even in such complex collocations as 'transexual demisexual polyamorous aromantic.'

For this word, like *pansexual*, it is shown by the corpus analysis that it is not yet accepted in the language (contextualized and lexicalized), and it seems to have been created as a form analogous to 'homosexual' and 'heterosexual', as there is an entire group of such neologisms (*polysexual, asexual, bisexual, pansexual*, etc.). They are developed on the LGBT websites and blogs and used by this community. As a form created by analogy, it may be indicative of a current trend in this area of language in naming new concepts. This will be further analysed below. This neologism, like *pansexuality*, is rarely used outside of the LGBT community (not counting the mocking uses of the word outside of the LGBT forums) and the chances of it leaving that area and spreading over the other registers of language are scarce, due to its complexity, conflicting definitions and necessity for it in the language. Whereas a neologism like transphobia or cisgender has some grounds on which to be accepted into language, the same cannot be said for 'demisexual'. It may only be taken as a passing trend, as in the case of pansexuality in the US.

5. Discussion and results

The aim of this study was to explore the neologisms in the area of gender studies by the means of a comparative corpus analysis of English and Croatian. Furthermore, the focus was on the word formation processes in the area of gender studies and how they relate to the word formation processes present in the other areas of language nowadays (such as social media). The area of gender studies was chosen for the analysis because there has lately been a surge of words such as 'pansexual' and 'non-binary' in English, which may point to English becoming a more inclusive and equal language. The question of why there is a need for these neologisms in the language has been answered by the analysis of the scope and variety of use of these words in English and Croatian. Analysis of the context and register of use points to the need in language for a certain word.

In a recent study of most frequent word formation processes, affixation was found to be the most common word formation process in English, followed by compounding (O'Dell 2015). In this study, all of the analysed neologisms are derivatives. Most of the affixes are found to be neoclassical, which is, according to O'Dell (2015:95), common in the other areas of English nowadays, as well[10]. There are several features that are in common to all of the words in the study. They have prefixes which have their roots in Greek (*trans-, pan-, demi-*) and Latin (*non-, cis-*). They were all created by the word formation process called derivation, which is where a prefix is added to a base word at the beginning of the word or a suffix at its end. Pius Ten Hacken (2011) calls these affixes 'neoclassical'. In fact, he considers the neoclassical word formation system to possibly be a different kind of system than the other word-formation systems in English, which has not been researched enough. The emergence of neoclassical roots and affixes is explained by Ten Hacken as being the result of the intention to create "more general concepts" (2011:100). However, this is not explained in further detail. He makes two observations of the neoclassical word formation:

1.   Neoclassical word formation is available for naming new concepts.

2.   Neoclassical word formation is not applied unintentionally (Ten Hacken 2011:86).

---

[10] O'Dell (2015) gives an example of the use of Latin prefixes *de-* and *inter-* in words such as *debug, declutter, defriend* and *interleading, intersectionality*. (95)

To go back to creativity and productivity, this indicates that the resulting neologisms are results of productivity. The reason for the creation of neoclassical words is the intention to name new concepts. By the principle of productivity, new words are created in this process in analogous fashion. As Ten Hacken (2011:86) explains, "each new form is produced by a kind of analogy."

In gender studies, for example, neologism 'cisgender' followed the neologism 'transgender' and neologisms 'polysexual, 'asexual', 'demisexual', etc. followed 'heterosexual' and 'homosexual'. The comparison with the word formation processes in other areas of language, mainly of the language used in social media, has shown that the complexity of new words may nowadays be intentional, as in the world where people are constantly bombarded with various information, one has to think of a means of capturing their attention. Words in gender studies follow suit in this regard as they attract attention with their form. In 2014, *The Atlantic* published an article about whether 'cisgender' will survive as word and be added into dictionaries by lexicographers like its antonym 'transgender'. As referenced in the article, new words ought to be easy, that is, 'user-friendly' – 'easy to pronounce, read and spell'.[11] The argument against 'cisgender' is that it is not user friendly since prefix *cis-* is not as familiar in English as *trans-* is.

Another thing discovered is that some of these words are created in analogies, from the other words from the area, which shows how productive this area of language can be (infinite ways to refers to the various types of sexuality and gender can be produced).

Not all neologisms in this area are derivative analogies, however. There are also many compounds and abbreviations used mainly by the LGBT communities for now. For example, the aforementioned abbreviation *enby* (*non-binary > NB > enby*). This word was also analysed in the *iWeb*, and its scarce use is mostly found on LGBT boards and forums. It can illustrate how people communicate on these boards – within the community where it is assumed that everyone is aware of the meanings of the terminology that is used. The statements commonly found are the following: 'send some **enby** positivity out there', 'sending all my **enby** siblings good wishes', etc., that is, they do not, among themselves, take time to clarify the meaning in any way, be it overtly or covertly. When it comes to the overt and covert contextual clues, as defined in Bauer and Renouf (2001), the fact that it is often

---

[11] The author of the Atlantic's article 'Will Cisgender Survive?' is Paula Blank, professor of English. https://www.theatlantic.com/entertainment/archive/2014/09/cisgenders-linguistic-uphill-battle/380342/ (25 September 2020).

overt contextual help that is found in the texts using these words (outside of the LGBT community's websites) shows that the author recognizes this word as new and its meaning as foreign to their readers (or to the authors themselves as well). Glosses, quotation marks and brackets are very frequently found, especially with words with fewer entries in the corpus (*demisexual, pansexuality*). In Croatian, overt contextual help is found more frequently than in English across all registers, including texts belonging to the LGBT community. Covert contextual help (which may be unconscious on the part of the writer) is more frequently found in English, with collocations such as 'come out as (*demisexual, enby, non-binary*, etc.)', which make it easier for the readers to understand the meanings of these neologisms. Lexical signals, semantically related words and repetitions are also frequently found. For example, with neologism *pansexuality*, often either the root *sexuality* is repeated in the text or other types of sexuality are mentioned (*bisexuality, homosexuality, transsexuality*, etc.) as semantically related words.

The question answered in this study is also the question of the importance of language in the fight for gender and sexuality equality. Language is shown to be extremely important to the LGBT groups when it comes to their identity. Language can either erase or affirm their existence and it is the first step in their fight for recognition and equal rights. This is determined to be the reason for the creation of all these neologisms. The corpus study made it possible to study these words by the context in which they appear and the words they collocate with, as well as the registers in which these words are used. The register analysis is extremely important as it shows the frequency (and variety) of use of a certain neologism. In the fight for a gender equal and sexuality inclusive language, it is important that these neologisms are used not only in the LGBT community's websites and blogs, but also in the other registers, like the newspapers and magazines, as well as in the academia, where it has been important to offer LGBT students the words with which they can identify. As shown in the previous chapter with corpus analysis of these neologisms, in English they are used across all register (with more or less contextual clues), whereas in Croatian they are still mostly being introduced and explained by the LGBT community to their members. The only neologism which has more results is the one carrying negative connotations, *transphobia*, and it may be the proof that the use of these neologisms can, to some degree, point to the state of gender equality in language and society in general.

However, even in English, it remains questionable whether people really use non-sexist, inclusive and equal language in their daily lives in conversation, etc. The problem of

analysing the acceptance of these words is that other than the hybrid registers of blogs, forums and chats, we do not have the examples of real talk in the corpus. In the academia and journalism, people are encouraged and demanded to use this kind of language, but it remains uncertain whether people who use this kind of language where it is prescribed, do the same in their everyday conversations. For now, I consider that people should make an effort to use this kind of language in their everyday speech as well, even if it sounds forced at first, because the use of these words has to be made habitual, not just among the members of the LGBT communities – it should be widespread.

6.Conclusion

To sum up, this study's aim was a corpus analysis of the five neologisms from the area of gender and sexuality studies, in order to learn whether these words are accepted in language. The languages compared were Croatian, which is a gendered language, and English, which is more gender neutral as it does not have grammatical gender. The detailed analysis of the neologisms' word formation processes, meaning, context of use and collocations has shown that some of them have been taken up by the speakers of mainly US and UK varieties of English, whereas in Croatian, there is still long way to go. This study has shown that some of these neologisms do signify that English is turning into a more equal, inclusive language (*cisgender, non-binary, transphobia*), whereas others, like *demisexual* and *pansexuality*, are more interesting in terms of their formation processes, than because of their meanings, as their use is shown to be the result of the trendiness of accepting various sexuality and gender identities, more than of the need to name the concepts in language. Lexical analysis of word formation processes of these neologisms showed that the dominant word formation process is neoclassical derivation, that is, a number of Latin and Greek prefixes are used to build new words. Context-wise, it was shown that in both English and Croatian, the authors still offer contextual clues (overt or covert) in order to clarify the meaning. That indicates the awareness that these neologisms' meanings are recognized to still be foreign to some.

7. References

Attenborough, Frederick (2014). Words, Contexts, Politics. *Gender and Language* 8 (2), 137-145.

Algeo, John (1991). *Fifty Years Among the New Words: A Dictionary of Neologisms 1941-1991*. Cambridge: Cambridge University Press.

Balhorn, Mark (2004). The Rise of Epicene *They. Journal of English Linguistics*, 32 (2), 79-104.

Biber, Douglas and Conrad Susan (2009). *Register, Genre and Style*. Cambridge: Cambridge University Press.

Bucholtz, Mary and Hall, Kira. Theorizing Identity in Language and Sexuality Research. *Language in Society* 33 (4), 469-515.

Bauer Laurie and Renouf Antoinette (2001). Contextual Clues to Word Meaning. *International Journal of Corpus Linguistics* 5 (2), 231-258.

Caleffi, Paola Maria (2015). The 'hashtag' a new word or a new rule? *SKASE Journal of Theoretical Linguistics* 12 (2), 46-69.

Caswell, Andrew, Laakso Emmi K. and Prewitt-Freilino, Jennifer L (2012). The Gendering of Language: A Comparison of Gender Equality in Countries with Gendered, Natural Gender, and Genderless Languages. *Sex Roles* 66, 268-281.

Fischer, Roswitha (1998). *Lexical Change in Present-Day English*. Tubingen: Gunter Narr Verlag.

Galić, Branka (2004). Seksistički diskurs rodnog identiteta. *Socijalna ekologija* 13 (3–4), 305-324.

Galić, Branka (2009). Društvena uvjetovanost rodne diskriminacije. *Rodna ravnopravnost i diskriminacija u Hrvatskoj: istraživanje 'Percepcija, iskustva i stavovi o rodnoj diskriminaciji u Republici Hrvatskoj'*. Zagreb: Ured za ravnopravnost spolova Vlade Republike Hrvatske. 9-28.

Galić, Branka (2012). Promjena seksističkog diskursa u Hrvatskoj? *Socijalna ekologija* 21 (2), 155-177.

Green, Eli R. & Maurer, Luca (2015). *The Teaching Transgender Toolkit*. New York: The Center for Sex Education.

Kamenov, Željka and Jugović, Ivana (2009). Percepcija, iskustvo i stavovi o rodnoj (ne)ravnopravnosti u obitelji. *Rodna ravnopravnost i diskriminacija u Hrvatskoj: istraživanje 'Percepcija, iskustva i stavovi o rodnoj diskriminaciji u Republici Hrvatskoj'*. Zagreb: Ured za ravnopravnost spolova Vlade Republike Hrvatske. 120-143.

Lehrer, Adrianne (2003). Understanding Trendy Neologisms. *Italian Journal of Linguistics* 15, 369-382.

Meyer, Charles (2002). *English Corpus Linguistics: An Introduction*. Cambridge: Cambridge University Press.

O'Dell, Felicity (2015). Creating New Words: Affixation in Neologisms. *ELT Journal* 70 (1), 94-99.

Opačić, Nives (2012) *Hrvatski ni u zagradama*. Zagreb: Hrvatska sveučilišna naklada.

Plag, Ingo (2003). *Word-formation in English*. Cambridge: Cambridge University Press.

Reljac Fajs, Emilija and Vidović, Ester (2015). *Usvajanje anglizama i njihova adaptacija u hrvatski jezični sustav. Jezik, kultura i književnost u suvremenom svijetu : zbornik radova sa znanstvenog skupa s međunarodnim sudjelovanjem Međimurski filološki dani II održanoga u Čakovcu 26. travnja 2013*. 215-223.

Ten Hacken, Pius (2011). Neoclassical word formation in English and the organization of the lexicon. *10th International Conference of Greek Linguistics,* Komotini: Democritus University of Thrace, 78-88.

Zimman, Lal (2017). Transgender language reform: some challenges and strategies for promoting trans-affirming, gender-inclusive language. *Journal of Language and Discrimination* 1 (1). 84-105.