

Birds-Eye View of Health Studies Comparing the Impact of Gender Bundles: A Descriptive and Historical Analysis of Genders, Social and Linguistic Representations

Bárbara Soares da Silva, Maria Sylvia de Souza Vitale

Postgraduate Program on Education and Health in Childhood and Adolescence, Federal University of São Paulo, São Paulo, Brazil

Email: casadatraducao@gmail.com

How to cite this paper: da Silva, B. S., & de Souza Vitale, M. S. (2024). Birds-Eye View of Health Studies Comparing the Impact of Gender Bundles: A Descriptive and Historical Analysis of Genders, Social and Linguistic Representations. *Open Journal of Social Sciences*, 12, 132-148.

<https://doi.org/10.4236/jss.2024.122009>

Received: January 10, 2024

Accepted: February 19, 2024

Published: February 22, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The research presented was motivated by the need to understand the *representations of human beings* in their identified *genders* mainly during *adolescence* as an important stage of life. However, the work is also motivated by understanding the representations of the human being in other phases of life. The genres with their recognized categories and descriptors were operationalized through the following *lexical elements* added mainly in English due to the search engine (Google Books through association with Brigham Young University) used assuming that language as the search standard. At the same time, the scanning will be carried out in Brazilian Portuguese in the corresponding search engine called *Brazilian Corpus* made available by the *Linguateca* initiative¹. Therefore, these are the genders collected for the research: *male, female, gay, lesbian, homosexual, transgender people, person, transgender, transgender person, transsexual person, two-spirit person, transgender person, transgender person, transsexual person, transgender two-spirit transsexual, gender queer, gender dysphoria, gender identity, gender diverse, gender non-conforming minors, non-conforming gender-conforming, gender-affirming, non-binary, nonbinary, cross gender, gender diverse, nonconforming teen, LGBTQ, nonconforming person, gender minority, gender transition, family, adopted family*, (as well as its appropriate plural forms due to the analysis that shows differences between the representations around the singular form and the plural form of the term). As stages of life, they were, therefore, operationalized through the following lexical

¹Access to Linguateca is through the address: Access to bodies: Brazilian Corpus (linguateca.pt).

items²: *man, woman, adolescent, adolescence, adult, boy, girl, child, elderly, child, young person and adolescent* (as well as their plural forms due to the analysis that shows differences between the representations around the singular form and the plural form of the term). The analysis data consisted of publications made available by *Google Books* in the period 1800 to 2008 (i.e. 208 years), totaling around half a billion words, in addition to the use of *Corpus NOW*, which brings together publications in English and Brazilian Portuguese in this analysis. The NOW (News on the Web) Brazilian Portuguese *Corpus* contains around 1.1 billion words of data from newspapers and magazines on the Internet in four Portuguese-speaking countries between 2012 and 2019. For this, several analyzes were carried out on the ngrams (sequences of adjacent words) formed by these words on *Google Books Ngrams* database. *This is research in full progress and, as the exploration of terminologies in the literature of the area is new and, as it is a study to obtain a postdoctoral degree, it may present new considerations as its development develops.*

Keywords

Genres, Vulnerability, Terminologies, Phases and Stages of Life

1. Brief Presentation of Research History

In 2014, as a specialization student, I observed patients in the life cycle of adolescence in the Adolescent Medicine Sector, of the Discipline of Pediatric Specialties, of the Department of Pediatrics, of the Escola Paulista de Medicina in a Professional Update course for non-Doctors with the esteemed Prof. Dr. Maria Sylvia S. Vitale, to understand what—from the point of view of this phase of life—it could represent for patients in their various vulnerabilities. Then, from 2015 to 2019, I entered the Postgraduate Program in Applied Linguistics and Language Studies at the Pontifical University of Sao Paulo, under the direction of Professor Dr. Tony Berber Sardinha, I found in Corpus Linguistics the *basis* to investigate the various social representations of human beings over time. However, I lovingly needed to return to my home, a place where I discovered myself as a researcher and interested in the issues of adolescence, to delve deeper into topics such as gender identifying words that can often be seen in the linguistic patterns of a Stages of Adolescence person's life.

As the years went by, as I realized the clear vulnerabilities contained in being born a woman in an exclusive society that exalts men, even in the 80s, it was possible to feel that I was part of a minority. Having heard motivational phrases from my parents like, “your life project doesn’t have to be getting married, having children, and washing clothes,” you can be whatever you want to be. Even so, in my heart I sought validation from the external world and expected, from social

²Lexical items mean units of meaning and the entire set of lexical items of a language is called lexicon and is also recognized as a general synonym for the vocabulary of words in commonly recognized slang.

interaction, the need to satisfy the expectations of my environment. It was not easy. I can imagine what it was like for my sister, who after several heterosexual relationships decided to come out as bisexual. *A new world has opened, not a new abyss.*

I began to notice more the minorities, the delays, the looks, of adults, children, adolescents in my routine as a teacher of higher technological education, how the environment was surrounded by oppressed people. Not depressed, oppressed by the standards that until then seemed correct. Among them, I experienced tears, discouragement, I saw people who gave up interacting socially in a work relationship, in a classroom, in so many environments, so as not to be judged for declaring themselves transsexual, homosexual, or bisexual. In the outpatient clinic of the Adolescent Medicine Sector, I was able to observe a welcoming look towards this public. Since 2022, I have continued to carry out my postdoctoral studies in the midst of descriptive research on terminologies focused on genres in the past, present and future. This research can be declared as a new literature because the social declaration that this topic can be discussed in scientific and academic environments is new, and because it is not socio-interactional and interdisciplinary.

2. Introduction

The search for word frequencies in PERSPECTIVES OF MULTIPLURALISM AND MULTILINGUISM IN *GOOGLE BOOKS*: The descriptive and historical scanning of genres, deformities and social and linguistic representations was due to the need to understand what lines of agreement exist between the different genres in relation to the life cycles of the Human Being in the age groups that make up a person's journey. In particular, for this research, the focus was always more on adolescents.

In this sense, the challenge was to investigate the most frequent linguistic patterns identified in the English language, beginning with understanding the concept of: 1) *social representation*; 2) *concept of gender and what it implies for current research*; 3) *exploration of life cycles, especially that of adolescence in English terms*; 4) *historical patterns from remote time periods recorded through Google Books in partnership with BYU to more current ones through Corpus NOW*; 5) *research with current and past literature on the concept of gender from a dialogic and linguistic point of view and based on laws, concepts and investigative research in the area.*

As a reference to what is being investigated, the etymology of the term gender brings several openings for research that involves the area being investigated: GENDER—from the Latin genus, means “race or extraction”. The word was first used by Julius Caesar who, when describing his exploration of the region of Gaul, possibly referring to a tribe of Celtic origin, gave it other endings. In the same search you can reflect on the verb and textually DEGENERATE: “to become corrupted, to lose essential qualities”, of the genre mentioned above. This Latin word comes from the Indo-European source gen- or gnê-, “to generate,

beget, cause to be born”.

The objective of the research is therefore to understand what collocations exist in published terminologies in which *adverbs, adjectives and nouns can be found* that “mean something” through the attraction between one word and another, between a term and another term. Collocation so to speak is defined as the “appearance of two or more words within a short ‘distance’ of each other in a text,” according to Sinclair (1991: p. 170). This approach provides the interpretation and understanding of reality to which this research lends itself.

Although it is possible to observe that a diverse number of other words have been generated that are part of the referential vocabulary, the search for understanding in childhood and adolescence is essential in the most diverse areas of knowledge: the discovery of developments that may exist more forward, in a future not so distant from the one we live in today.

This research aims to inform a study whose objective is to investigate frequent patterns of words in the English language identified in the stages of human life in conjunction with the genres identified in these searches, as well as possible vulnerabilities observed in the data collected from *Google Books* that They cover the period from 1800 to 2008, in addition to data collected from *Corpus NOW* (News on the Web), which covers the period from 2012 to 2019.

In addition, in order to seek and identify patterns such as those that emerge from these words, the present study also aims to verify if there are changes in relation to the patterns identified for both genders and ages, that is, the stages of life throughout the ages, twenty-one years of decades of study and also in the period of comparison with the most current database (*Corpus Now*, from 2012 to 2019).

Therefore, we set out to answer the following research questions:

- 1) What are the possible representations of social vulnerability in relation to the collected genres?
- 2) Can possible representations of social vulnerability be identified in relation to the life stages collected?
- 3) Is there a difference between feminine and masculine terminologies when mixed with life stages?
- 4) Is there a difference between the terms in relation to the valuation (positive and negative charge)?

Justification

Bringing to this research the issue of *gender variability* in terminology in addition to the linguistic standardization of the academic core, gradually contributes to a growing and necessary social awareness that can alleviate and soften behaviors seen in relation to transphobia. Furthermore, one of the bases of this study is to analyze the recurring patterns in the different stages of life, where curiously it can be realized that the life expectancy of a trans person is only 35 years and Brazil unfortunately has very high rates of murders due to gender phobias other than those called “expected” by today’s society.

The National Association of Transvestites and Transsexuals (Antra), in collaboration with the institution *Transgender Europe*, confirms that the number of murders due to transphobia reached 132 between October 2018 and September 2019 and 163 in the entire year 2018. These figures place Brazil in first place, it heads the list of countries that kill the most due to transphobia.

Given this, the present study is justified by the clear need to address the issue focused on the various health and education services in the state of São Paulo and in Brazil as a whole. Mainly in Universities, Technical and Technological Education, there is still little dissemination of the topic in study practices in the educational area, mainly. Along with raising awareness, the practice of greater assistance in the search for patterns of recurrence is promoted, proposing new questions regarding the language that should be used, the linguistics used in dysphoria, as well as encouraging new scientific research and extension projects that contribute to the reduction of an oppression that certainly affects all phases of human life in its vulnerabilities and genders identified along the way.

3. Goals

3.1. General Objective

As a general objective, the search deals with what are, possibly, the discoveries in the intersection of information on gender versus age throughout the period stipulated in the research, in addition to verifying in the various publications what are the true confrontations, transformations and perspectives, experienced by these adolescents/young transsexuals/gender variables, throughout the set of stages of life as an explicit description in the publications, these being from periods far from the most current ones.

3.2. Specific Goal

Characterize as research participants in social aspects, the profiles crossed in the data sought, especially those seen in the adolescent age group and who present transsexual/gender variability as their gender;

Understand the process of discovery of sexuality and sexual orientation of transgender/gender variant adolescents/youth researched in *Google Books* through *BYU* and the *NOW* database;

Understand the confrontations experienced with emphasis on the adjective variables of negative evaluation, being: oppression, prejudice and exclusion throughout the proposed terminology;

Investigate the future life prospects, age expectancy of transgender adolescents/young people/with gender variability by cross-referencing information with other age groups and the gender immensity identified throughout the publications.

4. Methodology and Data Collection

4.1. Study Design

The present study addresses an exploratory, empirical, descriptive research with

a qualitative approach.

The qualitative approach was selected for this research because it allows verifying an interpretive analysis between the variables found in the scanning of the hypotheses with the literature. Despite the use of extremely analytical and statistical tools throughout the terminological study, it will be the interpretations of said data that will provide coherence and uniformity to the work. According to Minayo (2014), those who carry out qualitative analyzes separate the different modalities of the instruments applied and the materials collected, highlighting the empirical categories and establishing comprehensive bases for the research topic and/or initial question.

From this perspective,

The qualitative method is the one that is applied to the study of history, relationships, representations, beliefs, perceptions and opinions, products of the interpretations that humans make about how they live, build their artifacts and about themselves, feel and they think. [...] This type of method, which has a theoretical basis, in addition to making it possible to reveal still little-known social processes related to particular groups, facilitates the construction of new approaches, review and creation of new concepts and categories during the research. It is characterized by empiricism and the progressive systematization of knowledge until the understanding of the internal logic of the group or process under study. Therefore, it is also used to develop new hypotheses, build qualitative indicators, variables and typologies (Minayo, 2014: p. 57).

4.2. Research Database (Selected Corpora)

The data used in this research deals with lists of bigram occurrences found in English-language publications indexed by Google Books. Google Books is a collection of millions of publications digitized by the Google company from library collections around the world. The format of this database, as well as its extension, will be explained in the procedures section. Bigrams are sequences of two words placed next to each other in a text. For example: “Brazilian boy”, “young women” and “Indian men”.

Bigrams are available on the Google Books Ngrams website, which also allows the user to search and produce graphs using these bigrams. Therefore, in the strict sense it is not directly a corpus of texts, since the texts of the publications indexed by Google Books are not made available to users. Google Books only provides bigrams. Thus, the data for this research were bigrams, along with their frequency of appearance in English-language publications indexed between 1800 and 2008 and are available through a registry in association with BYU, Brigham Young University, located in Utah, United States.

What we recognize here as *Corpus* or *Corpora*, is recognized as a Database or a large amount of concentrated information, comes from an area of knowledge

called *Corpus Linguistics*.

Corpus Linguistics can be understood as the study of language based on examples of language used in “real life” (McEnery and Wilson, 1996: pp. 1-2). There are many authors who define *Corpus Linguistics* in a similar way. Some examples of important voices are those of Sinclair (1991), Stubbs (1993), McEnery and Wilson (1996), Biber et al. (1998), Kennedy (1998), Hunston (2002) and Berber Sardinha (2000).

We can highlight four typical characteristics of a corpus-based analysis, as suggested by Biber et al. (1998: p. 5): 1) It is empirical and analyzes current usage patterns in natural texts; 2) uses a collection of natural texts, using specific software, and is carried out in an objective and comprehensive manner, known as a corpus, as a basis for analysis; 3) makes extensive use of the computer for analysis, using interactive and automatic techniques.

It depends on both analytical techniques: quantitative and qualitative. LC, however, is not a new-emerging-methodology for the study of language (as the generativist currents claim): LC presents a new vision of language. The theoretical-methodological position of the CL, according to Leech (1991: pp. 106-107), can be summarized in that which focuses on: linguistic performance; in linguistic description rather than in linguistic universals; in both quantitative and qualitative language models; on empirical research rather than a more rationalist view of scientific research. Regarding the concept of standardization, it's stated that “a standard is a phraseology frequently associated with a word, particularly in relation to the prepositions, groups and clauses that follow that word.” Specifically, it's defined such patterns as all words and structures that are regularly associated with the word and contribute to its meaning.

A pattern can be identified if a combination of words occurs relatively frequently, if it depends on a specific lexical choice, and if there is a clear meaning associated with it. Standardization is further defined as “regularity expressed in the systematic recurrence of coexisting units of various orders (lexical, grammatical, syntactic, etc.)” (Berber Sardinha, 2000: p. 31). Other theorists, such as Hunston (2002), define standardization as: all the words and structures with which they are regularly associated and that contribute to their meaning. A pattern can be identified if a combination of words occurs relatively frequently, if it depends on a specific word, and if there is a clear meaning associated with it. The frequency of co-occurrence between lexical items has allowed corpus linguists to analyze them according to association pattern phenomena known as collocations, coalitions, and semantic prosody.

Firth (1957) already warned of the following: “judge a word by its company.” This principle guarantees that “words do not appear at random in a text” (Sinclair, 1991: p. 110). Collocation is defined as the “appearance of two or more words within a short ‘distance’ of each other in a text,” according to Sinclair (1991: p. 170). There is also, in Hoey (1983), an indication that “collocation, for a long time, has been the name given to the relationship that a lexical item has

with elements that most likely appear in its (textual) context.” The three main definitions of collocation recognized in the area are: 1) Textual: collocation is the co-occurrence of two or more words in a small space of text separated from each other. 2) Psychological: collocation meaning consists of the associations that a word makes due to the meanings of other words that usually occur in the environment. 3) Statistics: collocation has been called the relationship that a lexical element has with elements that appear with significant probability in its (textual) context. Another important concept for LC is that of coalition:

It can be the grammatical company that the word keeps and the positions it prefers (Hunston, 2002).

It is also explained as the “association between lexical and grammatical elements” (Berber Sardinha, 2000: p. 18). For example: the expression “do justice” refers to the concept of coalition, since the word “jus” in Portuguese requires or prefers the presence of the verb “do”. “Jus” is a reduced representation of the word justice, therefore, what we have is “do justice”.

According to Berber Sardinha (2000), there is also another important association in CL, semantic prosody. The name semantic prosody can be attributed to the fact that certain words “prepare the listener, or reader, for the semantic content that is to come, in the same way that prosody, in speech, indicates to the interlocutor what types of sounds are” will come later (Berber Sardinha, 2000: p. 40).

Semantic terms can positively or negatively influence the words that accompany others or even the meaning and meaning of the text. Below are some steps used to begin data analysis for this research.

Portuguese Corpus and the *Contemporary American English Corpus* were taken into account as pre-analysis to continue the analysis via Google Books.

1) Access via registration to the portal <https://corpusdoportugues.org/now> and write in the Placed tab (which would be those placed) the search term of interest—in this case the term “transgender”

The screenshot shows the 'Corpus do Português: NOW' website. The main navigation bar includes 'SEARCH' and 'FREQUENCY'. Below this, there are tabs for 'List', 'Chart', 'Collocates', and 'Compare KWIC'. The 'Collocates' tab is selected. The search term 'TRANSGENDER' is entered in the 'Word/phrase [POS]?' field. Below the search field, there is a table of collocates with a grid of numbers and a 'Find collocates' button. The table shows the following data:

Collocates [POS]	0	1	2	3	4
+	4	3	2	1	0
-	0	1	2	3	4

Below the table, there are buttons for 'Find collocates' and 'Reset'. At the bottom, there are checkboxes for 'Sections', 'Texts/Virtual', 'Sort/Limit', and 'Options'.

Source: Survey data, 1st half of 2022.

2) Search result that took into account two words on the left and two on the right according to transgender *placement*:

HELP	①	★		FREQ	ALL	%	MI	
1	①	★	EUROPE	68	4557	1.49	15.59	
2	①	★	ONG	32	31111	0.10	11.73	
3	①	★	(	22	8156238	0.00	3.16	
4	①	★	ORGANIZAÇÃO	7	250994	0.00	6.53	
5	①	★	ASSOCIAÇÃO	7	277301	0.00	6.38	
6	①	★	INTERNACIONAL	5	349511	0.00	5.57	
7	①	★	GOVERNAMENTAL	4	15703	0.03	9.72	
8	①	★	INITIATIVE	3	859	0.35	13.50	
9	①	★	COMMUNITY	3	1102	0.27	13.14	
10	①	★	HEALTH	3	5124	0.06	10.92	
11	①	★	CENTER	3	22462	0.01	8.79	
12	①	★	EUROPEIA	3	192139	0.00	5.69	
13	①	★	DEFENSE	2	909	0.22	12.83	
14	①	★	BISSEXUAL	2	1229	0.16	12.40	
15	①	★	NÃO-GOVERNAMENTAL	2	3439	0.06	10.91	
16	①	★	LAW	2	3804	0.05	10.76	
17	①	★	BARRAR	2	7440	0.03	9.80	
18	①	★	STORY	2	8393	0.02	9.62	
19	①	★	TENTAM	2	28238	0.01	7.87	
20	①	★	AND	2	37971	0.01	7.45	
21	①	★	PALAVRA	2	81142	0.00	6.35	

Source: Survey data, 1st half of 2022.

3) Lower frequency associated with the word Transgender: “rights”

Source	http://beleza.terra.com.br/maquiagem/transgenero-e-forcado-a-remover-make-para-carta-de-motorista,4e73c37b3f9a6410VgnVCM3000009af154d0RCRD.html
Country/date	BR (14-06-17)
Title	Transgênero é forçado a remover make para carta de motorista

Expanded context

Transgênero é forçado a remover make para carta de motorista Grupo de direitos transgêneros alega que direitos constitucionais de o adolescente foram violados # 17 jun 2014 # 10h20 # atualizado a as 10h24 # comentários # O jovem transgênero Chase Culpepper, de 16 anos, foi forçado a remover a maquiagem para tirar foto de a carteira de motorista, em março de este ano. Agora, com o apoio de o grupo de direitos transgêneros [Transgender Legal Defense & Education Fund](#), o americano quer poder tirar uma fotografia maquiado para o documento. Os dados são de o jornal The Huffington Post. saiba mais # Segundo Culpepper, os funcionários de o Departamento de Veículos Motorizados de a Carolina de o Sul disseram que não parecia como achavam que um menino deveria ser e uma pessoa chamou sua maquiagem de "disfarce". # " Isso é quem eu sou, e minha roupas e maquiagem refletem isso ", disse Culpepper. " O Departamento de Veículos Motorizados não deveria ter me forçado a remover a maquiagem simplesmente porque minha aparência não atende a as suas expectativas de como um rapaz deve se parecer. Só quero a liberdade de ser quem sou, sem o departamento me dizer que, de alguma forma, não sou bom o suficiente. " # em o dia 9 de junho, o grupo de direitos transgêneros enviou uma carta para o

 Translate: --Choose--

 Analyze text (keywords, detailed word sketches, find related phrases)

Source: Survey data, 1st half of 2022.

a) Highest frequency associated with the word Transgender: “Europe”

19-01-02 BR	O POVO Online	de mortes de pessoas homossexuais em 2017. # Outro estudo feito por a ONG Transgender Europe, o Brasil também é o país que mais mata
19-02-13 BR	EL PAÍS Brasil	travestis e transexuais perderam a vida de forma violenta (citando dados de a ONG Transgender Europe). # O Grupo Gay de a Bahia é uma as
19-02-20 BR	UOL	canal criado para receber informações sobre violações a os direitos humanos), por o Transgender Europe e por o GGB (Grupo Gay de a Bahia)
19-04-21 BR	Meu Timão	como o país onde mais transexuais são mortos, de acordo com dados de a Transgender Europe (TGEU). Entre 1º de outubro de 2017 e 30 de s
19-05-10 BR	G1	em o local por um homem não identificado. # Um levantamento de a ONG Transgender Europe divulgou em o ano passado uma pesquisa feit
19-05-17 BR	Hypeses	, 455 pessoas foram mortas vítimas de homofobia. Os números são de a ONG Transgender Europe. # 34 mil curtidas e outros 9 mil compartili
19-05-24 BR	EBC	número de mortes de pessoas trans em o mundo, segundo a organização não governamental Transgender Europe (TGEU), com sede em a Ale
19-06-05 BR	Jornal Extra	redes sociais entre os defensores de as causas LGBT. Segundo dados de a ONG Transgender Europe (TGEU), o Brasil ocupa o primeiro lugar e
19-06-09 BR	Jornal Extra	Brasil continua a ser o país onde mais transexuais são mortos, segundo a ONG Transgender Europe. Porém, acredito em a força de a arte e de
19-06-10 BR	Emails Estadão	de Stéfano se reflete em números. De acordo com um relatório de a ONG Transgender Europe, de novembro de 2018, o Brasil é o País onde m
19-06-14 BR	EL PAÍS Brasil	denúncias de violações de direitos humanos), por o GGB e por a associação Transgender Europe. # O Brasil também é o país mais perigoso de
18-10-05 PT	Expresso	pessoas transgênero em o mundo inteiro (facto atestado por um estudo de a ONG Transgender Europe, publicado em o final de 2016). " É im
18-09-12 BR	EXAME.com	menos 868 travestis e transexuais mortos entre 2008 e 2016, segundo a ONG europeia Transgender Europe. Outro dado mostra que 90% de
18-09-12 BR	Catraca Livre	mata travestis e transexuais em o mundo, de acordo com dados de a ONG Transgender Europe. O reflexo de tamanha violência é refletido em
18-09-30 BR	G1	' É um primeiro passo', diz ativista # Cianán Russell, de a Transgender Europe, uma de as principais ONGs de o mundo de defesa de os direitos
18-10-04 BR	Jornal Extra	os dar, foi incível. # De acordo com uma pesquisa realizada por a Transgender Europe (TGEU) de 2017, em dez anos, em os 71 países
18-10-24 BR	Jornal O Globo	a transfobia entre 2008 e 2016. Foram 868 em o Brasil, segundo a Transgender Europe. Já que me expuseram, quero mostrar que estão mata
18-11-11 BR	Diário da Amazônia	o mundo, com 868 pessoas assassinadas entre 2008 e 2016, segundo a ONG Transgender Europe. Apesar de os dados, a mulher afirma que n
18-06-22 PT	SAPO 24	comunidade internacional de ativistas de os Direitos Humanos e aclamada por organizações europeias como a Transgender Europe e a ILGA
18-01-17 BR	Jornal O Globo	o que as trans norte-americanas. # De acordo com o Relatório Mundial de a Transgender Europe, de um total de 325 assassinatos de transgê
18-05-06 BR	UOL	de acordo com o Projeto de Monitoramento de Assassinatos de Pessoas Trans de a organização Transgender Europe. Um estudo publicado e
18-05-09 BR	Globo.com	a o ano anterior. Os dados corroboram outros relatórios de entidades internacionais como a Transgender Europe, que apontam o Brasil como
18-05-12 BR	RIC/PMI - Prefeitura de Indaiatub	em o mundo. A constatação é feita com base em levantamento de a ONG Transgender Europe (TGEU). # Conforme a Associação Nacional de T
17-11-09 PT	Sol	o Brasil lidera o ranking mundial de assassinatos, segundo dados de a organização não-governamental Transgender Europe. # Ao longo de a
17-10-21 BR	Diário de Pernambuco	Brasil é o líder mundial de violência contra trans, segundo a organização não governamental Transgender Europe, rede europeia de organiz
17-10-21 BR	CartaCapital	país que mais assassina pessoas trans em o mundo, segundo relatório de a organização Transgender Europe em 2016. # Só em o ano passad
17-12-18 BR	O POVO Online	cenário de violência. Segundo o CGL, o Brasil é considerado por a associação Transgender Europe um de os países que mais mata transexu

Source: Survey data, 1st half of 2022.

4) Choice to make use of *Corpus of Contemporary American English*:

Corpus of Contemporary American English

SEARCH FREQUENCY CONTEXT ACCOUNT

ON CLICK: ☒ CONTEXT ☐ TRANSLATE (77) ☐ ENTIRE PAGE ☐ GOOGLE ☐ IMAGE ☐ PRON/VIDEO ☐ BOOK (HELP)

HELP ☐ ☒ ALL FORMS (SAMPLE): 100 200 500 FREQ

1 Se ☒ TRANSGENDER 3565

0.234 seconds

Source: Survey data, 1st half of 2022.

Corpus of Contemporary American English

SEARCH FREQUENCY CONTEXT ACCOUNT

CLICK FOR MORE CONTEXT HELP SAVE TRANSLATE ANALYZE

1	2012	BLOG	...uality.wordpress.com	40 percent of all homeless youth identify as lesbian, gay, bisexual, or transgender. # Upon coming out to their parents, 26 percent of LGBT people repo
2	2012	BLOG	...uality.wordpress.com	is a study by Jody L. Herman entitled "The Cost of Employment Discrimination against Transgender Residents of Massachusetts," that found that discr
3	2012	BLOG	...uality.wordpress.com	Residents of Massachusetts, "that found that discrimination against the state's approximately 33,000 transgender residents costs the state millions of c
4	2012	BLOG	...uality.wordpress.com	only does the Commonwealth suffer lost income tax revenue because of discrimination, but each transgender person who loses a job may become elig
5	2012	BLOG	...uality.wordpress.com	, and public-assistance expenditures to replace lost income and insurance coverage. # The National Transgender Discrimination Survey found that tran
6	2012	BLOG	...uality.wordpress.com	replace lost income and insurance coverage. # The National Transgender Discrimination Survey found that transgender people also face discrimination
7	2012	BLOG	womanist-musings.com	in Denver, Colorado, and an author, public speaker, and trainer on transgender issues. He blogs at Tranifesto. In his ideal world, no one would
8	2012	BLOG	dailykos.com	? No. For one thing, it actively discriminates against people with disabilities, transgender people, and, to a lesser but still serious extent, women. And
9	2012	BLOG	ohio.edu	serving more than 11 years in the Ohio University Lesbian, Gay, Bisexual, Transgender Center, director Mickey Hart left his position in August to become
10	2012	BLOG	crossdreamers.com	men and women who dream about being the other sex, crossdressing, crossdressing and transgender issues August 27, 2012 The Other Side of your T
11	2012	BLOG	crossdreamers.com	crossdressing and transgender issues August 27, 2012 The Other Side of your Transgender Soul - from Dostoyevski to Buffy the Vampire Slayer # The A
12	2012	BLOG	crossdreamers.com	a woman or a man is changing. That has repercussions for how crossdreamers and transgender people see themselves. The changing anima # In my pr
13	2012	BLOG	crossdreamers.com	see themselves. The changing anima # In my previous posts in my series on transgender psychology I argued that Carl Jung's concept of the anima (the
14	2012	BLOG	crossdreamers.com	the animus (the unconscious masculine side of women) may be useful in a transgender context, but not in the way most people think. # Far too often
15	2012	BLOG	crossdreamers.com	stereotypical feminine personality traits are ridiculed. # This is probably why male to female transgender are more likely to be made fun of than the fen
16	2012	BLOG	crossdreamers.com	bodied persons are to suppress the feminine ones. # This means that male bodied transgender persons who have typical "feminine" personality profil
17	2012	BLOG	crossdreamers.com	often associated Cybele with Dionysus. In Rome Cybele's priestesses were male to female transgender. # Jack, Do "gynephilic" crossdreamers and tran
18	2012	BLOG	crossdreamers.com	. # Jack, My opinion is that all this constant brouhaha over transsexual versus transgender is simply a product of the fantaic western mindset to reduce
19	2012	BLOG	crossdreamers.com	sex-reassignment surgeries and at the same time, they start maligning the other groups of transgender people are "autogynephilics", or "effeminate

Source: Survey data, 1st half of 2022.

According to the image above, the previous illustration of the preliminary results using the COCA database consists, *Corpus of Contemporary American English*: The theme begins in 1993 and ends in 2012.

Most of the information transmitters represent blogs and sometimes smaller and larger newspapers, dating from the first decade of the 2000s onwards. And below is the first screen of Google Books after previously registering on the portal:

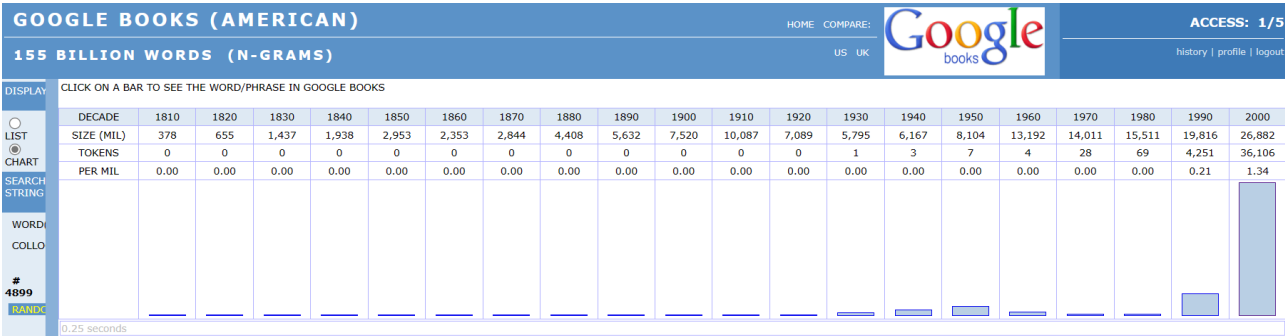


Start with which corpus?

Corpus	Size (words)
American	155 billion
British	34 billion
Spanish	45 billion

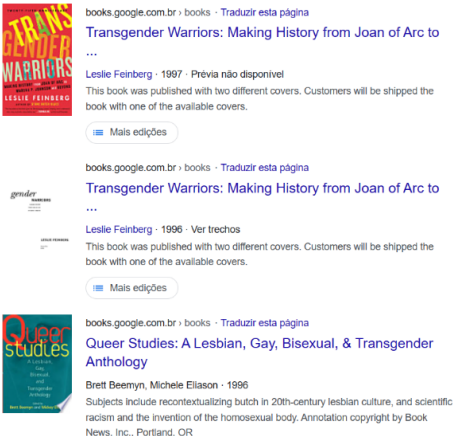
[Compare to standard Google Books interface]

Source: Survey data, 1st half of 2022.



Source: Survey data, 1st half of 2022.

Above it is possible to find the initial results after choosing the word *transgender* and choosing two placed on the right and two placed on the left. Since the 1990s, it is possible to find recorded books over the Internet that mention the transcriber “Transgender”.



Source: Survey data, 1st half of 2022.

In comparison, the term LGBT first appears in the “Internet world” only after 1970, and still very timidly.

WORD(S)	CHARTS	TOTAL	1810	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000
1 LGBT	G	48130																	1	11	3515	44603

Source: Survey data, 1st half of 2022.

Pesquisar páginas em Inglês ▾ Qualquer visualização ▾ Qualquer documento ▾ 1 de jan. de 1970 – 31 de dez. de 1979 ▾



Nenhuma imagem disponível

books.google.com.br > books · Traduzir esta página

LGBT Danmark collection of newspaper and magazine clippings ...

1973 · Prévia não disponível

This collection complements other parts of the LGBT Danmark collection of newspaper and magazine clippings on LGBT topics.

[Mais edições](#)



Nenhuma imagem disponível

books.google.com.br > books · Traduzir esta página

LGBT Danmark collection of newspaper and magazine clippings ...

Nikoline Werdelin · 1975 · Prévia não disponível

Names of cartoonists and their comic strips represented include: Alison Bechdel (Dykes to watch out for = Dykes på væggen), Claire Bretecher (Den nye mand), Howard Cruse, Matt Groening (Life in Hell = Livet i helvede), Ralf König, Olaf ...



books.google.com.br > books · Traduzir esta página

Mathematics for Health Professionals - Página 266

Ida Marie Martinson, G. R. Kepner · 1977

ENCONTRADO NO TEXTO – PÁGINA 266

Source: Survey data, 1st half of 2022.

Above, by clicking on the element found, we can read the news of the time or view the book scanned by *Google Books*. Procedure 3.3 will clarify the additional steps followed for data analysis.

4.3. Procedure in the Google Books N-Gram Database

The data for this research comes from the *Google Books* N-Gram database. An n-gram is a sequence of words, so a unigram is one word, a bigram is a sequence of two words, a trigram is a sequence of three words, a quadrigram is a sequence of four words and a staff, so five words. *Google Books* offer these five options. On the *Google Books* N-Gram web address you will find several language options for n-grams: English, Spanish, Russian, etc. For the English language, there are the following versions: 1) English version 20120701 2) English version 20090703 3) One Million English version 20090705 4) American English version 20120701 5) American English version 20090715 6) English version British 20120701 7) British English version 20090715 8) English fiction version 20120701 9) English fiction version 20090715 The *Google Books* search interface provided by Brigham Young University (via the online *corpus project* by Pro-

fessor Mark Davies) searches more quickly and efficiently than directly through *Google Books* Ngram.

However, this interface does not provide access to the full “English 20120701” collection, but only to American English (items 4 and 5 above) and British English (items 6 and 7 above). Since the option for the American database is almost five times larger than that of the British database, we opted for the American database, which corresponds to the “American English (155 billion)” option in the Google interface BYU Books. This option actually refers to point 5 above, American English version 20090715 and not point 4, American English version 2012071.

The 2009 and 2012 versions differ not only in relation to the fact that the 2012 version has three more years of publication, but mainly due to the fact that the 2012 version considerably increased the number of publications and therefore the number of words from previous years. For example, in the 2009 version, there were just over 3 billion words indexed by 2005; in the 2012 version, that number more than tripled, to more than 10 billion words. The same thing happened, to a different extent, in relation to the other years. In the “Total” line the totals from 1810 to 2009 appear.

As you can see, the base size doubled between the 2009 version and the 2012 version. When verifying a brief comparison between the number of words in *Google Books* version N-Grams 2009 and 2012, American English. It is important to note that the database names on the *Google Books* N-Gram website and the BYU interface reflect the place of publication and not the exact variety of English. Therefore, publications listed under the “American English” option are not necessarily written in American English or by native American English authors, just as those in British English do not exclusively reflect British English or British authors. This indexing itself is based on data recorded by *Google Books* from the libraries where the publications were automatically scanned and may contain inaccuracies, as there is no information on whether the data was subsequently manually verified. Furthermore, even if the bibliographic record in the database is reliable, a book published in Great Britain may have been written by an American author and vice versa, calling into question the representativeness of the text as a copy of one of these variants.

5. Final Considerations and Remarks

To answer the questions of this ongoing research, computational tools were used, available on the Internet as mentioned above, and an *interpretive-qualitative analysis of the data was carried out*, since, despite the effective use of the tool, the ability to extract social representations occurs satisfactorily with human eyes. Regarding the first research question, the results showed an excellent range of representations that are still being explored.

The evaluative representation of the human being was the most constant among the terms, carried out through the application of adjectives such as “dif-

ficult", "good", etc. Another very common representation is that of superiority, a more accentuated type of evaluation, carried out through adjectives such as "older" or even "European" together with transgender, as used to illustrate this study, etc. Physical representations ("attractive", "beautiful", "heartthrob", "handsome", etc.) were also very common. But it was in the age grading where a representation ("early", "older", etc.) focused on clinical conditions ("autistic", "crazy", "deaf", "depressed", "sick" was very present., etc.), as well as types of behavior ("aggressive", "angry", "bored", etc.), social issues (e.g. "disadvantaged", "homeless", "unemployed", etc.) related with the example shown in this work, such as in association with the term transgender or inferiority (e.g. "desperate", "primitive", etc.) and gender identification (e.g. "bisexual", "heterosexual", "homosexual", etc.).

This set of representations shows characteristics frequently attributed to human beings from the historical point of view and the locations implicit in the published texts and their biases, as far as we have been able to verify, this is the first description of this type in the literature, in which We seek to cross-reference data and even seek in future analyzes to understand whether: "to be transgender is to be an adult", such as "to be a woman is to be an adult", or "to be a man is to be an adult". This analysis requires a statistical scan through patterns and probabilities, but we are not far from this—on the contrary—if there is space for this research to take place, there will be lexical space for minorities to recognize themselves in all stages of human life.

Vitale et al. (2019) lead us to reflect on the issue of adolescence in its entirety, seek a practice based on evidence, demystifying the work with this life cycle in its various aspects, whether physical and/or emotional health, legal or social. It also suggests that thinking about it means being closer to the topic, from the most common problems of adolescence to the most current and controversial issues, ranging from drug use, school difficulties, dentistry, sports, nutrition and diet, the rights and duties of adolescents, violence, fashions, sexuality, sexual and reproductive rights and intervention programs. Its *interdisciplinary and multiprofessional nature* highlights the depth of the different perspectives of the different specialists, favoring a unified approach for better care, monitoring and understanding of gender and its confrontations.

The analysis of this present research seeks to detail the representations of each term and showed unexpected considerations: that apparently synonymous terms have different representations and that morphological forms of the same term can also have separate representations. In relation to the first point, the analysis may indicate a mixture of representations of different nuances: "adolescents" attribute greater emphasis to gender identification; adolescent/s, mental aspects; and adolescent/s, to marital status. But everyone seeks evaluation, coloring this phase of life in a not very favorable light. Regarding the second point, the analysis shows differences in the considerations between the representations regarding the singular and plural form of the terms. For example, gender identification

is more commonly conveyed through the plural rather than the singular form of “adolescent”. This confirms once again the finding of *Corpus Linguistics* that the language in use avoids true synonyms, since each form tends to take on a different role (cf. Sinclair, 1991).

In response to the second research question, the analysis seeks to verify whether there may be differences between the representations both from a historical and gender point of view. From a historical point of view, one can notice a slight increase in representations of life stages as “flexible” categories that can be graduated, using adjectives such as “young”, “old” and even “adolescent”.

Compared to the beginning of the 19th century, mainly when the search for collocations with respective adjectives, nouns or verbs begins, there is an increase in representations of gender identification that go beyond the binary or non-binary classification through “homosexuals”, “lesbian”, “bisexual” and related adjectives left and right. Also considered in the searches was an increase in ethnic-racial representations, such as “black” and a decrease in others, such as “colored”.

If it were possible to summarize the initial findings of this research, it could be suggested that *the conception of a human being can be flexible throughout their life*, going from initially being a figure that is classified by their physical appearance, then by their nuances of age, then by his potential, and other times by a conflictive figure, marked by clinical issues or by the identification of behavioral traits, ending in a figure recognized for his failed occupation. And finally, in relation to the third research question, the analysis shows possible considerations of variation between gender terms and the initial results suggest that masculine generic terms (man and men) tend to have a more positive evaluation than all other terms feminine. This is not surprising, but at the same time, feminine terms in initial searches seem to be represented more positively than masculine terms (child and children). Regarding age progression, the analysis suggests, although still expanding, in general, that *adulthood is more positively represented than childhood*, which in turn is more positively represented than adolescence.

The intersection of this data with the intended genres is still under analysis and qualitative and interpretive considerations must be shared further in the advancement of this research. In all cases, however, the best-rated item tended to be in the singular, while the worst-rated item tended to be in the plural.

What this research seeks to offer, *for the first time*, is an overview of how human life is represented historically, in terms of its age stages and its differentiation between gender exploitation and its perceived vulnerabilities.

Research tends to show the impossibility of broad generalizations: each term tends to present its own range of representations, as if they represented niches of experience of social groups that distinguish them from other terms.

The generalization that we can make, from the initial considerations, is that the passage of life can be marked by a constant evaluation of the human being in relation to others, in relation to himself, in relation to society, in terms of a more

or less finite relationship of respective representations, with a notable evaluative bias.

It remains to be understood to what extent the assessment brings new terminologies that truly represent these social groups linguistically or not, clinically speaking. Finally, we consider at the present moment the great discovery possible from this study, the possibility of exploiting that, in a short or medium term of time, it will be possible to observe that new terminologies can be associated with genders with respect to the age of a being human.

If earlier in history an adult individual was recognized as “a boy” or a child—that was because people lived less, and at fourteen years old, for example, that child was already doing the tasks of an adult, he already had the responsibilities of an adult and He even married as an adult and perhaps even had children.

In a Latin American country such as Brazil, unfortunately, it sounds like a reality in some corners of the country, but what we really seek to exploit is that, the issue is, that today, while we say *woman to a woman*, or man to a man, We recognize that woman and that man as adults and not as adolescents—and so to speak, with the support and help of statistical studies such as, the use of the SAS Edition (SAS on demand for Academics), SPSS (IBM, on Predictive Analysis Classification and Regression), use of *Sentiment Analysis* by Hamilton et al. (2016) in addition to other *scripts* to be developed in association with the area of Software Engineering using *Python language*, *Power Shell* among others, in association with the areas of Psycholinguistics, Applied Linguistics and Education and Health for these challenges, we intend to find possible new terms that society will use to recognize a transsexual, for example, as an adult.

This is one of the great possible discoveries of our research and since this is new literature in the area of knowledge, we will treat it with all the attention that the topic deserves as this research is developed.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- Berber Sardinha, T. (2000). Corpus Linguistics: History and Problems. *Delta*, 16. <https://doi.org/10.1590/S0102-44502000000200005>
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511804489>
- Firth, J. (1957). *Studies in Linguistic Analysis* (2nd ed.). Special Volume of the Philological Society.
- Hamilton, W. L., Leskovec, J., & Jurafsky, D. (2016). *Diachronic Word Embeddings Reveal Saint*.
- Hoey, M. (1983). *On the Surface of Discourse* (p. 211). George Allen and Unwin.
- Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge University Press.

<https://doi.org/10.1017/CBO9781139524773>

- Kennedy, G. (1998). *An Introduction to Corpus Linguistics*. Longman.
- Leech, G. (1991). The State of the Art in Corpus Linguistics. In A. Karin, & B. Altenberg (Eds.). *English Corpus Linguistics: Studies in Honor of Jan Svartvik*. Longman.
- McEnery, A. M., & Wilson, A. (1996). *Corpus Linguistics (Edinburgh Textbooks in Empirical Linguistics Series)* (p. 209). Edinburgh University Press.
- Minayo, M. C. de S. (2014). *The Challenge of Knowledge: Qualitative Research in Health* (14th ed.). Hucitec.
- Sinclair, J. Mch. (1991). *Corpus, Concordance, Collocation*. Oxford University Press.
- Stubbs, M. (1993). British Traditions in Text Analysis: From Firth to Sinclair. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), *Text and Technology: In Honour of John Sinclair* (pp. 1-33). John Benjamins. <https://doi.org/10.1075/z.64.02stu>
- Vitalle, M. S. S., Silva, F., Pereira, A., Weiler, R. M. E., & Niskier, S. R. (2019). *Adolescent Medicine: Fundamentals and Practices*. Editora Atheneu.