UNIVERSITY OF TARTU
DEPARTMENT OF ENGLISH STUDIES

PRODUCTION OF COLLECTIVE MEMORY IN *THE NEW YORK TIMES'* OBITUARIES AND ITS *OVERLOOKED* SERIES: A CORPUS-BASED STUDY

MA thesis

KAROLIINE LAANEOJA
SUPERVISOR: *Prof*. RAILI MARLING

TARTU
2024

ABSTRACT

Obituaries are a genre of commemoration where individual lives become a part of the collective memory through how the events and people they represent are portrayed. The *Overlooked* series published in *The New York Times* aims to give obituaries to people who did not receive one at the time of their death for different reasons. This thesis aims to compare the obituaries of *The New York Times* with its *Overlooked* series to find out what are the similarities and differences between the two obituary cycles, what the vocabulary used can tell us about the production of collective memory, and how it differs in the two obituary types. For this goal, corpus-assisted discourse analysis is used. The corpus tool Sketch Engine is used, with two corpora created for the aims of this thesis. One corpus consists of the regular obituaries published in *The New York Times* from January 1, 2021, to December 31, 2021 (1,388,794 words), and the second of *The New York Times*' *Overlooked* series up to and including December 2023 (241,153 words).

The introduction discusses the history of the obituary genre and *Overlooked* together with who are the subjects of obituaries. The first chapter provides an overview of previous research into the obituary genre, its role in shaping collective memory, previous research on obituaries as a discourse, and how media has attempted to use affirmative action strategies to create a more inclusive collective memory. The second chapter introduces the methods used, Corpus-Assisted Discourse Analysis, the corpora used, and the key findings of the corpus-assisted analyses of the vocabulary used in *The New York Times'* obituaries and its *Overlooked* series. The conclusion summarizes the overall findings of the thesis.

TABLE OF CONTENTS

INTRODUCTION

The obituary recounts a given person's life after their death. Thus, the obituary is a genre of both commemoration and chronicling and to an extent reflects different societal values at any given time (Hume 2000: 11). Hume (2000: 11) has claimed that obituaries can link memories of individual lives with generational and collective ones. Obituaries can reveal different values through commemoration, representing an ideal which contributes to the understanding of history, as some people and events are given more significance through publication than others (Hume 2000: 13-14). According to Hume (2000: 150), obituaries also help to fulfill a given society's need for a shared memory through their presence in the mass media and their reflection of shared cultural values.

The obituary as it is known today – a death announcement including a short biography – started appearing in the 18th century in London (Fowler 2007: 4) in *The Gentleman's Magazine*. They were later also published in *The Daily Universal Register* (later known as *The Times*), where, for a brief time, deaths were listed in order of precedence, starting with the members of the House of Lords (Fowler 2007: 5-6). Occasional individual obituaries were granted separate entries, but these became a regular feature only in 1879, remaining inherently linked with class (Fowler 2007: 6). While the status of the people commemorated continues to be central to this day, it is no longer necessarily tied to social class, but derives from a person's achievements. This is one example of how obituaries also reflect changing social values.

The obituary tradition moved from the UK to other English-speaking countries, including the United States, where they now appear regularly in daily newspapers. For example, *The New York Times* has a regular obituary section with a specialized staff. It describes its obituaries process as follows:

> There is no formula, scoring system or checklist. One thing to remember is that it is not our intent to honor the dead; we leave the tributes to the eulogists. We seek only to report deaths and to sum up lives, illuminating why, in our judgment, those lives were significant. The justification for the obituary is in the story it tells. (McDonald 2022)

The newspaper, that is, makes an indirect reference to collective memory, as it seeks to focus on lives that are "significant" and that tell a story that fits the broader narrative of the period. These criteria can fit very different people: the person's life has to carry a newsworthy story to be granted an obituary. Newsworthiness can be measured by comparing events and stories to official and unofficial values or norms of the country and period (Lester 1980: 985). Values shift, and hence, it can be assumed that what was newsworthy and thus also memorable a hundred years ago need not be so. Events or people considered do not just present reality or tell us what was or is important; they also forge it (Lester 1980: 985). In other words, by writing about certain kinds of people in the obituary section, newspapers also create a collective memory of what matters, which can influence people's perception of events, stories, and people.

The New York Times, besides its usual obituaries, has started publishing a series titled *Overlooked*. First published in 2018 for International Women's Day, the obituaries editor of *The New York Times*, William McDonald, calls it the series of *The Times* "acknowledging that many worthy subjects were skipped for generations, for whatever reasons" (McDonald 2018). He speaks of the vital part of choosing a person who will receive an obituary: the person's newsworthiness and achievements, which McDonald argues often go hand-in-hand with fame (McDonald 2018). He calls the obituaries desk a "rearview mirror, reflecting the world as it was, and not as we might wish it to have been" and uses this to justify the current lack of women and people of color as subjects of obituaries (McDonald 2018). This descriptor of the current obituaries as a "rearview mirror" shows that the obituaries we now have reflect who had the possibility and privilege in society back in the day – mainly white upper- and upper-middle-

class men (McDonald 2018). Indeed, historically, for example, during the 19th century, Native Americans were rarely granted obituaries, and the rare persons covered were represented in such a way as to portray their inferiority to the whites or were only portrayed because of their connection to the whites (Hume 2000: 134-135). Similarly, the deaths of African Americans were not written about in mainstream media and appeared only when a newsworthy narrative could be found (Hume 2000: 136). Others who have been historically mostly excluded from obituaries include the poor, social outcasts (i.e., prisoners or mentally ill), and most working-class people (Hume 2000: 136-137).

While the genre historically gave attention to minorities when newsworthiness could be found for the white readership, now changes are being made to the genre to ensure that the collective memory covers a more diverse range of people. The *Overlooked* series acts as a vessel to give retroactive, retrospective obituaries to subjects not previously highlighted, creating newsworthiness and bringing these people back to fill the gaps in the historical narrative of the time when they died. The *Overlooked* series stands in contrast to the conventional obituaries cycle in *The New York Times*, as it talks about people a considerable time after their death, as informed by today's perspective.

The present thesis aims to compare the obituaries of *The New York Times* with its *Overlooked* series obituaries using corpus-assisted discourse analysis to find out the similarities and differences between the two obituary cycles. More broadly, the thesis is interested in what the vocabulary used can tell us about the production of collective memory and how it differs in the two obituary types. In order to fulfill this aim, the thesis first introduces the obituary genre and its role in the shaping of collective memory, as well as previous research on obituaries as a discourse and ways in which media has attempted to use affirmative action strategies to mitigate previous prejudice. The term covers strategies to increase the number of representatives of

marginalized groups who have been underrepresented in a field (Fullinwider 2018). While affirmative action has mostly been debated in higher education and hiring, media discourse is also a place where attempts are made to increase diversity and create a more inclusive collective memory. This will be followed by the introduction of the method used, Corpus Assisted Discourse Analysis, and the process of data collection and analysis. The results will be compared to the main findings of the literature review.

1 THE OBITUARY GENRE, *OVERLOOKED*, AND COLLECTIVE MEMORY: LITERATURE REVIEW

The professionally-written obituary has different criteria, such as which people are chosen to be written about after their death due to limited newspaper space and how these subjects are represented. As a genre responsible for shaping public memory about different people, the way this memory is conveyed and produced is relevant for understanding the broader value system of a culture. The following literature review will first cover research into obituary studies, then research into collective memory, and then discuss previous methods used to research obituaries and collective memory.

Fowler (2007: 110) speaks of the dichotomy that the obituary column editors face: while there is a mission to broaden the content beyond those of a certain status, they must also maintain the cultural canon. Obituary editors often avoid representing certain ethnic groups or genders just to maintain equal representation, as this may skew the understanding of the past when inequalities were widespread, but are also aware of the impact this omission may have for following generations (Fowler 2007: 110). What obituary editors are looking for are people who have had significant achievements or, when it comes to non-traditional subjects of obituaries, fascinating life stories (Fowler 2007: 116). The editors admit that the choice is often very subjective and mostly do not attempt to represent every group equally (Fowler 2007: 115-116). They are aware of women getting fewer obituaries than men because the women who have died belong to the generations that were not as active in the public sphere as today or did not hold higher positions within the labor market that usually merit attention in an obituary (Fowler 2007: 119). *The New York Times* editor interviewed by Fowler (2007: 119) noted that this is expected to change as more women enter the workforce and leadership positions.

The *Overlooked* project of obituaries can already possibly be considered as a part of this change. The series has been investigated by Boyce and Dove (2022), who show that it contrasts the regular obituaries section of *The New York Times.* Boyce and Dove (2022) pointed to the media's considerable influence through what is omitted and what remains in the obituaries. The retrospective obituaries feature marginalized groups, focusing on people from North America and occasionally on other historically significant people from other parts of the world who were previously not given an obituary in the newspaper (Boyce & Dove 2022: 511).

Within the study by Boyce and Dove (2022), two obituaries were used as case studies: those of Charlotte Brontë and Ida B. Wells-Barnett. The two *Overlooked* obituaries analyzed were found to carry an uplifting or inspiring overarching story that seeks to contribute to the current debates, using language that is also adjusted to present-day concerns (Boyce & Dove 2022: 519-520). The changes employed to contribute to present-day debates might include diminishing the coverage of views dated by today's standards or ignoring actions that could be viewed negatively, as the project has an overarching narrative of positive representation (Boyce & Dove 2022: 520). The opposite is also done: emphasizing aspects of the covered people's story that would be aligned with progressive ideas while portraying the subjects as pioneers of their respective fields (Boyce & Dove 2022: 520). Although highlighting people who have been historically ignored can be considered a noble cause, Boyce and Dove (2022: 520) claim that the *Overlooked* series instead fails to take the people they cover as people on their own merit but rather as models for some ongoing debate on the background of the male-dominated regular obituaries section. Despite this, the general public often praises *The New York Times* for creating the *Overlooked* series within the comments sections of the published obituaries, thanking the newspaper for the introduction of these people to the previously unaware reader (Boyce & Dove 2022: 519).

This cycle of obituaries can be viewed as an example of reversed memory. Zandberg et al. (2012) define reversed memory as helping to further the understanding of past events as continuous, extending to the present by commemoration of the traumatic past by narration of the triumphal present. They indicate that within the context of the construction of journalistic values and norms (such as neutrality, objectivity, and newsworthiness) and the values of collective memory construction (such as ethnocentrism and national solidarity), media research acknowledges that there are similarities between the processes of shaping collective memory and producing journalistic content, as both involve decisions about what is essential, giving meaning to facts by inserting them into cultural interpretive frames (Zandberg et al. 2012: 66). This important for both present-day and retrospective obituaries.

Fowler (2007: 34) also speaks of a 'counter-memory', in which there is a struggle over the meaning of different events and their interpretations, originating from those who are subordinate to contest the interpretation of different historical events. Those who have been historically marginalized possess a counter-memory of the common narrative; however, it only exists as an inversion of the dominant discourses, with an emphasis on a want to be recognized (Fowler 2007: 34-35). The *Overlooked* series can be viewed as a site of such struggle in which people who are alive in counter-memory are inserted into the mainstream narrative to shift the latter's focus at least somewhat.

Getting an obituary in *The New York Times* can be considered a type of recognition. The editors are aware of this and thus engage in deliberation when there might be moral issues involved in the choice (Fowler 2007: 121). Some of these rules about who is excluded have changed over time. For example, it is now acceptable to mention when someone died through suicide (although occasionally less concrete wording such as 'he died by his own hand' is used) or had extramarital affairs (Fowler 2007: 122). When a figure is of intrinsic sociological interest,

they may be included even if they have committed murder on a larger scale, with every editor agreeing that they would have done an obituary for Hitler, Goebbels, or Stalin (Fowler 2007: 122). Fowler (2007: 122) claims that all editors interviewed alluded to the ongoing process of molding collective memory despite not explicitly using the concept.

Schwartz (1991: 302) defines collective memory as an image that uses the familiar concepts of remembering and forgetting to articulate how a society retains and loses information about the past. Social institutions, among which we can also count news media and their obituaries, are the means through which members of society learn about the past, distinguishing significant events from the mundane and adding moral meaning to the past (Schwartz 1991: 302).

Hume, in her essay on the evolution of research on collective memory and mass communication, says that collective memory is "as much about current needs as about the events and people of the past" and that “history and memory /…/ are not necessarily only at odds but have a more complex relationship” (Hume 2010: 182, 183). The interpretations of historical facts can change, and the new meanings attributed to them can provide insights into today’s values, as the past is not immutable and can be reinterpreted to accommodate changes in beliefs over time (Hume 2010: 183). This correlates with the ideas of Boyce and Dove (2022), who also show that the past is used to meet the current needs of the newspaper or the public, creating a complex temporal relationship between the subject outside their historical context and today’s perspective.

Edy (1999: 72) discusses the importance of media in shaping and maintaining national collective memory. She stresses that history, as taught in schools, for example, tries to present objective facts, but the media can instead encourage an emotional and personal connection with history (Edy 1999: 72). Commemorative events (among which we can also classify obituaries)

tend to noncontroversial, but they can also become a place where social consensus in being built (Edy 1999: 74). This is especially likely in situations where social values are changing (for example, with respect to the position of formerly marginalized people) and where emotional connections, therefore, become especially important in embracing a broader interpretation of history.

However, obituaries are not only tools for shaping the collective memory but also respond to different emotional needs. Fowler (2005) defines the obituary as a tool to help in the grieving process and to provide a verdict on a person's life. This verdict includes judgments about people, their lives, and their contributions (Fowler 2005). Although Fowler focuses on the needs of the individuals or communities who lose those critical to their lives, the same sense of grieving is also essential to social groups who have suffered some historical tragedy or loss (e.g., the survivors of the Holocaust or victims of racial discrimination). The social verdict and its emotional force thus go beyond individual grief.

Fowler argues that the subjects chosen for obituaries have traditionally belonged to the highly educated elite and that obituaries further empowered them (Fowler 2005: 63). It has been shown that to this day, common people are underrepresented in obituaries and, through that, the collective memory because of our narrow definition of what kinds of achievements are worth commemoration. However, obituaries have, with time, become more universal (Fowler 2005: 63). Bressers and Hume (2010), for example, find that the non-professionally written online obituaries help to eliminate the disproportional representation of just the elite and aid in creating a more inclusive collective memory. In the case of *The New York Times*, in addition to the lengthy obituaries that can be provided only for a small number of people every day, there are shorter death notices, the number of which is much higher. We meet a much wider range of people in these death notices, usually written by families. The opposite is also true: it is argued

that a group can be removed from power by refusing them access to different forms of memory by a more dominant group (Fowler 2007: 35), such as was done, for example, in East Germany post-1945 by excluding Jewish wartime suffering from the narrative of regional resistance in World War II (Misztal 2003: 59).

Fowler (2005: 66) also notes different genres within obituaries, with differences that shape how the readers interpret the factual content, which is essential for collective memory. As outlined by Fowler (2005: 64), the first of these obituary types is the traditional positive obituary, characterized by the celebratory representation of its subject. The second is the negative obituary, acting as a kind of commemorative retribution by criticizing the subject in a non-neutral manner (Fowler 2005: 65). The subjects of the negative obituary are usually people characterized by power abuse or general unanimous dislike by the community. Their obituaries remind people of the danger of forgetting and create a distance between the subject and the reader (Fowler 2005: 66). The third type of obituary is the rare tragic obituary, the subject of which is described as the victim of misfortune and an object of sympathy (Fowler 2005: 65-66). For this obituary type, Fowler (2005: 65-66) gives the example of the boxer Bubi Scholtz, whose obituary in *The Independent* on August 23, 2000, described his rise from the workers' district of Berlin as a great boxer, his later risky business endeavors, followed by the shooting of his wife, prison, and death in obscurity, cementing him in public memory as a fallen hero. The fourth type of obituary is the ironic obituary, which, instead of outright othering, uses comparatively small amounts of praise to critique the subject (Fowler 2005: 66). More often this type is used for politicians and requires an understanding of the position of the newspaper within the media field (Fowler 2005: 66). This obituary type could be illustrated by that of Viscount Whitelaw in *The Guardian* on July 2, 1999, which recounts his job as a Deputy PM to Margaret Thatcher, his charms, shrewdness and his prison-building program; and inserts jokes about him while also

outlining his privileged background (Fowler 2005: 66). The ironic aspect is revealed in his loyal subordination, considered shameful in hindsight (Fowler 2005: 66). The last type is the untraditional yet positive obituary, where the subject does not follow the traditional, ascending path of life (Fowler 2005: 66). Fowler's (2005: 66) example is the obituary of Nora Kahn Piore (*New York Times*, June 15, 2000) where her dedication to union research led to her becoming a professor and a presidential adviser very late in her life.

As the *Overlooked* project of *The New York Times* is relatively new, there is little research on it specifically. This project's obituaries could be categorized primarily, by the definitions of Fowler (2005), into either traditional positive obituaries with a celebrated protagonist or untraditional yet positive obituaries with a less linear and less traditional path of success.

Haskins (2007: 402) argues that when the task of remembering the past is delegated to formal institutions such as museums, archives, and memorials, it "weakens the need for a political community to actively remember its past". These institutions tend to further the ideologies of the ruling elite and take over the transmission of the shared past through participatory performance and ritual (Haskins 2007: 402). According to Haskins (2007: 402), the objects that were historically remembered mainly belonged to the intellectual and artistic elite, not illiterate artisans and performers. The objects displayed in museums were chosen for their ability to tell a narrative of historical progress and to serve a national identity, thus contributing to the loss of productive context as the original context is changed or positioned within or according to the narrative needed (Haskins 2007: 402). This removal of original context to suit a narrative aligns with the *Overlooked* series, as it seeks to avoid representing only the elite while still working within the guidelines and aims of today's newspaper and its readers.

The efforts of the *Overlooked* series could be described as a kind of affirmative action. This term refers to different programs designed to increase the representation of underrepresented groups in different institutions (Gilens et al. 1998: 167). These programs include additional outreach and preferential treatment of target-group members (Gilens et al. 1998: 167). These are also among the aims of the *Overlooked* series. Historically, affirmative action has been opposed in the USA because it conflicts with the central American values of fairness, equality, and democratic opportunity (Herring & Henderson 2012: 630). To reduce these tensions, neoliberal elites have used the notion of diversity and expanded the inclusion conversation beyond race and socio-economic inequality to ethnic minorities, women, and other disadvantaged groups (Herring & Henderson 2012: 630). Herring and Henderson call for a move towards “”critical diversity’ that not only celebrates and embraces cultural differences that exist between groups but also examines issues of parity, equity, and inequality” (Herring & Henderson 2012: 640). It is essential to keep this in mind in studying the Overlooked series to see to what extent the texts call attention to structural issues that led to the forgetting of the subjects of the obituaries at the time of their deaths. Just showing exceptional people would not create a critical shift in the collective memory towards a greater acknowledgment of past injustices.

Gubitz and Avant (2020: 685) argue that programs created by businesses as a result of economic incentives, as a response to public pressure for more extensive representational diversity, with no public policy feedback mechanism as a means of accountability should be categorized as diversity initiatives and not called affirmative action. However, it is not the terminology that matters here, but the existence of different affirmative action strategies or diversity initiatives employed by media, for example, to alleviate previous prejudices or to create an overall more progressive image. Gubitz and Avant (2020: 698-699), in their study of

how racial attitudes affect the perception of such affirmative action and diversity initiatives in media, find that diversity initiatives affect how White Americans behave in the marketplace. If people trust a given entertainment company, public attitudes towards diversity initiatives can be changed, which is not possible within affirmative action attitudes, as those have been the subject of political debate for decades and are more challenging to change because so many stereotypical opinions have already been formed. The companies involved must make meaningful changes in their commitment to diversity to gain support among White Americans (Gubitz & Avant 2020: 699). Thus, simply talking about trying to become more diverse is not going to change public opinion, and there needs to be real action towards it. The *Overlooked* series can be seen as such a step for one major media company. Therefore, it is essential to look at its discourses more closely to see whether we are dealing with a surface-level gesture or meaningful action.

The building of collective memory can also be observed within the daily obituary cycle. Taussig (2017) explores how collective identity is created within obituary life stories using the examples of obituaries of veterans of World War II and the Vietnam War. He claims that obituaries both carry and are influenced by collective memory as they reflect, partly through which stories they tell, what values the given society had at the time of publishing (Taussig 2017: 461). However, these obituaries also include links to the public past by including and referring to culturally essential events or people at the time (Taussig 2017: 461). Although the subjects of obituaries lack control over their narrative, collective memory influences journalistic institutions, which have to balance official needs with vernacular ones (Taussig 2017: 262). He, similarly to Boyce and Dove (2022) and Zandberg et al. (2012), speaks of the obituary as a space for potentially remembering, reinterpreting, reiterating, or rethinking cultural norms (Taussig 2017: 463). The findings of his study show that there are differences in the representation of

World War II veterans and Vietnam War veterans, with the latter having less focus on war and more on danger and victimization, reflecting the differences in society's dominant narratives about these wars (Taussig 2017: 471)

Taussig's (2017: 465) research into the obituaries of World War II and Vietnam War veterans was based on a sample of 100 professionally written obituaries – 50 of World War II veterans and 50 of Vietnam War veterans from 12 different US newspapers – and used a qualitative narrative approach. Taussig claims that this approach allows us to find how journalistic norms, individual biography, and collective memory combine into how we remember veterans of different wars (Taussig 2017: 466).

In contrast, Fowler and Bielsa (2007) used quantitative methods when analyzing whom newspapers choose to remember within their obituaries section. They argued that although obituaries are becoming more inclusive with the increase of figures important for popular memory and counter-memory, they are still dominated by Western elites (Fowler & Bielsa 2007: 130). For this research, they used samples from newspapers from the UK (*The Times, The Daily Telegraph, The Independent,* and *The Guardian*), France (*Le Monde*), and the USA (*The New York Times*) from different periods, looking at the subjects' occupational and educational backgrounds, gender division, social mobility, nationality, ethnicity, and migration status. Fowler and Bielsa (2007: 221-222) found that obituaries are overwhelmingly of men, predominantly British, American, and European subjects. The predominant occupational backgrounds were found to lean towards higher-status professions – artists, the military, and political elites – while people with a background in farming, manual work, or even middle-class professions such as teaching and nursing were under-represented (Fowler & Bielsa 2007: 222). In their research into the *Overlooked* series, Boyce and Dove (2022) investigate only two

obituaries as case studies explicitly chosen because of the authors' research specialization in nineteenth-century celebrity culture.

Corpus-based research into obituaries is rare and often small-scale in terms of the number of texts used. Corpora have been used for obituary research previously by Moore (2002), who analyzed 100 obituaries in *The Economist* using corpus research to create a sociological profile of a typical person featured. Moore (2002: 524) found that ideology plays a role in both who is chosen for an obituary and how they are represented through text and image, with a typical profile of an obituary subject in *The Economist* being male, white, English-speaking, influential, from a western industrialized country and dying at an old age. The findings also showed that the subjects that deviated from this profile were treated as 'Them', opposing the 'Us' consisting of *The Economist,* its journalists, and its readership (Moore 2002: 524).

In the context of using computer-based methods to investigate news values in large datasets, Potts et al. (2015: 168) discuss the challenges of analyzing the construction of news values with corpus tools, as there is no defined set of grammatical or lexical devices that a corpus could be searched for to find news values. Instead, they discuss that, as news discourse is conventionalized, different devices used for news value construction could be investigated (Potts et al. 2015: 168). These devices include temporal markers (i.e., *yesterday, last*), typical emotion nous (i.e., *fear, concern*) or semantically related adjectives (i.e., *unexpected, astonishing)*, and also collocations of topic-associated content words (Potts et al. 2015: 168). A similar idea could also apply to investigating the creation of collective memory: finding standard devices that could be used to construct collective memory within obituaries. The lack of large-scale corpus research into obituaries and collective memory presents a possible gap that will be further investigated in this thesis, even if only in an explorative manner.

This literature review has shown that obituary writers and editors must make many choices about the content and language of the obituaries, which impact collective memory. As the regular obituaries column faces issues with poor representation of marginalized groups, the *Overlooked* series of obituaries published in *The New York Times* can be considered a move towards rectification of past gaps and the creation of a more inclusive vision of the past. While highlighting people previously not celebrated for various reasons, the retrospective series also aims to create uplifting or inspirational stories. This process, however, can create issues: for example, aspects of the subjects' lives that are today considered problematic are diminished. In contrast, aspects aligning with progressive views are emphasized to create a positive overarching narrative, thus changing the subjects' image in the collective memory and, perhaps, the broader collective memory by, for example, downplaying the presence of racism, sexism, or homophobia. Media has a vital role in shaping collective memory: the past is often reinterpreted within news media to accommodate the audience's needs or to adjust to the changes in beliefs, and the media helps to create a personal connection with the past. However, this process has to remain sensitive to the need to tell the story of past mistakes so we do not repeat them.

Obituaries can influence collective memory through the language they use to convey information, and this is why the following section will focus on the vocabulary used in obituaries. Most previous research into obituaries and collective memory at the time of writing has been either qualitative or not based on large-scale data. Therefore, the present thesis complements the existing research by adding a corpus-based research component.

2 COMPARATIVE CORPUS-ASSISTED DISCOURSE ANALYSIS OF THE OBITUARIES *THE NEW YORK TIMES* AND THE *OVERLOOKED* SERIES

2.1 Method

Discourse, as described by Fairclough (2015: 55-56), is language as a social practice, as it is intrinsically a part of society and is socially conditioned by other non-linguistic parts of society. According to Fairclough, linguistic phenomena are social as "whenever people speak or listen or write or read, they do so in ways which are determined socially and have social effects" (Fairclough 2015: 56). Social phenomena are also linguistic – language activity in social contexts is a part of different social processes and practices (Fairclough 2015: 56). In this view, language does not just reflect social realities but also contributes to shaping it. What differentiates discourse from text is that discourse is a process of social interaction including text, process of production, and process of interpretation with the text as its source (Fairclough 2015: 57). Around interpretation and production exist other, non-linguistic social conditions which influence these processes (Fairclough 2015: 57-58). My study will, above all, focus on the textual aspect of discourse, but like in the literature review above, I will place the textual analysis in a social context.

Corpus-assisted discourse studies approaches discourse through corpora, allowing for work with a larger body of texts than is habitual in traditional qualitative discourse analyses. While qualitative analysis allows the analyst to get a detailed look at a small number of texts, this approach does not allow the results to be generalized, and there is always the danger of choosing texts that match the research aims (Gillings et al. 2023: 1). Subjectivity is always to an extent present in qualitative analysis. Projects that use corpus-assisted discourse studies often work with a social question in combination with a linguistic one or are driven by what might link different social practices with their associated linguistic choices (Gillings et al. 2023: 1).

The data for this type of analysis can be anything, as long as it contains language and there is understanding that language and society influence each other (Gillings et al. 2023: 2).

Corpus-assisted discourse studies combines corpus linguistics and discourse analysis and thus allows us to look at a much larger number of systematically sampled texts, helping to avoid selection bias, issues with representativeness and giving the following analysis stronger empirical foundations (Gillings et al. 2023: 6). Corpus linguistics also help by creating the possibility for entirely new lines of inquiry through playing around with the data available, such as frequency lists (Gillings et al. 2023: 6). While corpus analysis does add quantitative potential it also has qualitative qualities, both equally as crucial for corpus-assisted discourse studies (Gillings et al. 2023: 6-7).

Corpus-assisted discourse studies, thus, demand commitment to both the computer-assisted quantitative and the human-led qualitative investigation of discourse, ultimately trying to figure out how discourse works and is created (Gillings et al. 2023: 7). Thus multiple research designs are possible in which corpus linguistics and discourse studies are brought together (Gillings et al 2023: 7). One possibility is to begin the research with discourse studies by first identifying discursive struggles in a small number of texts and then using corpus linguistics to find out their typicality. The other option is to start with corpus linguistics, the results of which will be interpreted using discourse studies on the basis of the textual analyses (Gillings et al. 2023: 7). The thesis at hand will primarily employ the latter approach.

This thesis will also use two corpora built explicitly for the purpose of this thesis, representing the language of two different types of obituaries present within *The New York Times*. This is necessary as simply using a reference corpus (massive corpora created to create hypotheses about general language) is not enough when the questions asked are about a more specific genre of language, in this case, obituaries. The type of data and its volume are essential

here (Gillings et al. 2023: 10). The creation of a new corpus has the benefit of including only pertinent texts from a time period of concern and thus creating a better focus for the whole study (Gillings et al. 2023: 10). Meanwhile, care needs to be exercised when choosing the texts: bias must be avoided not to cherry-pick texts that prove the hypothesis of the project (Gillings et al. 2023: 10). The corpora used within this thesis are made with the assumption that obituaries carry many cultural values of the given time through their inclusions and omissions. *The New York Times* is a widely read and well-regarded newspaper in the United States; hence, it is safe to hypothesize that it contributes to shaping American collective memory in its obituary genre. However, no generalizations will be made within the thesis about American collective memory as it is possible that the liberal newspaper does not necessarily correspond to the values of all members of the American public. All claims here will thus be limited to the shaping of collective memory within *The New York Times*.

Corpus-assisted discourse analysis will be used for this thesis in combination with Critical Discourse Analysis to ensure the soundness of qualitative interpretations while handling more significant amounts of text (Marchi & Taylor 2018: 4). The first stage of data analysis is performed with corpus tools to ensure greater objectivity but the data is then analyzed qualitatively to assess the context that is usually absent in large-scale data analysis (Marchi & Taylor 2018: 4). Despite this, complete objectivity cannot be ensured, but the patterns that can become visible within corpus data allow for findings to be organic, rather than evidence-fitting predetermined categories (Marchi & Taylor 2018: 5)

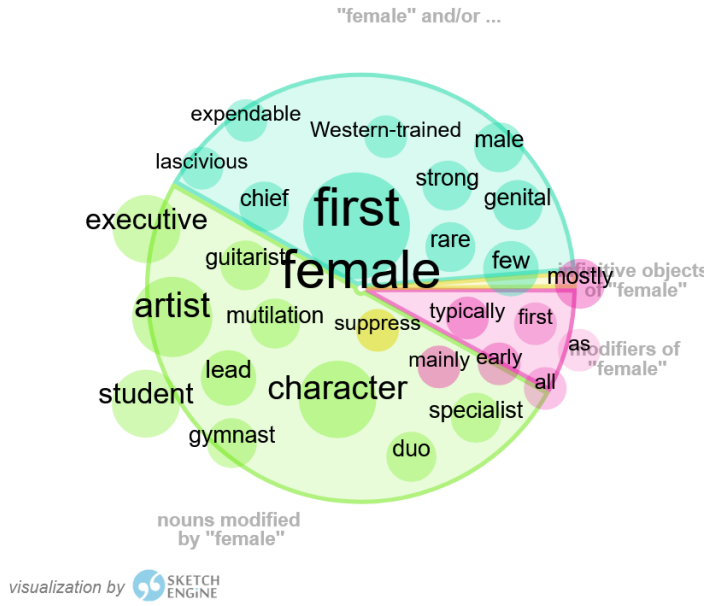
Two separate corpora were created for the purpose of this thesis. The first corpus consists of the regular obituaries published in *The New York Times* and was gathered from the online version of *The New York Times'* 'Today's Paper' section from January 1, 2021, to December 31, 2021, from the section titled 'Obituaries'. The second corpus was collected from the entire *The*

New York Times' *Overlooked* series up to and including December 2023. For this, the phrase "overlooked no more" was inserted into the website's search function, as this phrase is always present in the titles of the series' obituaries. Texts that contain these words elsewhere in the newspaper were omitted. Matches were manually searched, and only articles in the series were selected. The choice to include only a single year's texts in one corpus and the entirety of available texts of the series in the other was made as the *Overlooked* series has produced much less content than the regular obituary cycle. Adding the entirety of the series up to a point makes the text base large enough in size even to be worth analyzing with corpus methods. The choice of the year 2021 for the first corpus was random. The final size of the first corpus based on the regular obituaries cycle is 1,388,794 words (1,294 individual texts, average word count ~1,073 words). The final size of the *Overlooked* corpus is 241,153 words (195 individual texts, average word count ~1,237 words).

Texts for both corpora were copied into the *WordPad* program and thus converted into a .txt file, with obituary authors, image captions, and publishing dates omitted. The obituaries belonging to the *Overlooked* series were also omitted from the corpus of traditional obituaries. The *.txt* files were categorized into folders by month and year of publication for a more accessible organization.

Corpus-assisted discourse analysis will be used to research the two collected corpora with the help of the corpus management and text analysis tool Sketch Engine. Quantitative analysis will be performed using the tools 'word sketch' and 'word list' functions, and qualitative analysis will be performed using the 'concordance' tool. 'Word sketch' is a tool showing word behavior grammatically and within collocations, and the 'word list' feature arranges words by frequency (Kilgarriff et al. 2014: 9-10, 16). When comparing frequency between the two corpora relative frequency will be used, as due to the significant differences in corpus size, only looking

at raw frequency would not say much (Gillings et al. 2023: 18). Sketch Engine allows the creation of different word sketch visualizations, which show frequency and typicality. These will be employed throughout the thesis to illustrate trends within the corpora. The size of the circles within the graph shows how frequent something is (the bigger the circle, the more frequent it is), and the distance from the center shows typicality (the stronger the collocation, the closer to the center it is). Different colors indicate which segment something belongs to, and segment size indicates how many collocations belong to the grammatical relation segment.



An example of a word sketch visualization based on the adjective *female*.

Bednarek and Caple (2014: 151) have shown that corpus linguistic analyses of frequency can prove helpful in examining the discursive construction of newsworthiness and thus could also possibly be employed here to examine public memory construction. Generally, when comparing multiple corpora, the frequency can bring out differences between those corpora, with relative frequency (occurrences per million words) being especially pertinent due to the difference in corpus sizes (Hunston 2010: 3-5). 'Concordance' will be used to check the words

examined qualitatively in the context they appear to ensure no misinterpretation of quantitative data and to further examine the possible reasons for their frequency through closer investigation.

The thesis will analyze how public memory is constructed through the language used within obituaries and whether and how it differs in the regular cycle of obituaries and those of the *Overlooked* series. In their work, Bednarek and Caple (2014) introduce a framework for analyzing news discourse and news values. Bednarek and Caple (2014: 141) mention that traditional corpus linguistic techniques, such as analysis of frequency, analysis of keywords/clusters, or grammatical/semantic tags, can also be used for news values analysis. They use a corpus of 100 news stories of about 70,000 words and, as a first step, examine word frequency to understand how frequently repeated things could indicate how newsworthiness is being communicated to the reader (Bednarek & Caple 2014: 143). In my thesis, I will employ methods used by Bednarek and Caple (2014: 143-146) to identify different news values and discuss how these contribute to the production of collective memory. The different categories of news values and the word categories that could indicate their presence, as established by Bednarek and Caple (2014), will be used. I will exclude the category of personalization due to its irrelevance to the present topic (further expanded on below). Extra linguistic resources for identifying news values proposed by Potts et al. (2015: 151) will complement the method of Bednarek and Caple (2014). These modifications will be made as the text corpora analyzed differ in subject/genre; thus, different linguistic devices will be relevant to my research. The following news values and the linguistic devices used to construct them will be looked at:

Eliteness

This value is expressed by words and phrases that convey the high status or prominence of the subject (in this case, the subject of the obituary) and include words such as role titles (*boss*, *prime minister*, *doctor*/*Dr.*, *chief,* etc.) (Bednarek & Caple 2014: 143-144).

Superlativeness

This category includes words that convey the scale of the subject and could include the symbol # standing for a number, different numbers themselves, and word forms and bi-grams such as *percent*, *more/more than/than*, *the most*, *less than.* Concordance is essential in this instance as the context in which these words occur shows whether or not they are related to superlativeness (Bednarek & Caple 2014: 144). Intensified lexis (i.e., *stun, epidemic, wreck*) will also be investigated here (Potts et al. 2015: 151).

Proximity

Place names help to bring out what gets geographically and culturally highlighted within obituaries and, thus, what is projected into the collective memory regarding locations of importance. This category also includes the first-person plural pronouns *we* or *our*, but again, concordance must be used to see if the usage concerns proximity (Bednarek & Caple 2014: 144).

Negativity

This category will include words that are not directly negative but, within a particular context, could show negativity (i.e., *the police/police* could indicate crime) or groups that are generally considered negatively evaluated (i.e., *the IRA, The Islamic State*) (Bednarek & Caple 2014: 144-145). Negative evaluative language (i.e., *terrible, horrible, slaughter*) refers to negative emotions or attitudes (i.e., *distraught, condemn, criticize*), and general negative lexis (i.e., *conflict, death, damage*) could also be included here (Potts et al. 2015: 151).

Timeliness

This category includes potential temporal references (i.e., *yesterday, last night/month/weekend/November*) and modal verbs such as *will/would/could* (Bednarek & Caple

2014: 145). These timeliness indicators are essential when looking at the general *The New York Times* obituaries next to the *Overlooked* series, as temporality is a key factor in the texts.

Personalization

Personalization means that the subject of the news (event or issue) is given a personal face, conveys personal experiences and eyewitnesses, and references these people and experiences (Potts et al. 2015: 151). Obituaries are primarily written as a memorial; they intrinsically include personal experiences and references to the family of the deceased, together with their comments and experiences regarding the subject. Thus, as this is already a genre feature (compared to general news writing, where personalization might be used to create news value), looking at personalization as a part of collective memory production within obituaries might not be productive.

Novelty

One important bigram to consider here is *the first,* as it clearly conveys Novelty (Bednarek & Caple 2014: 146). Other indicators of newness (i.e., *fresh, new, historic*) or unexpectedness (i.e., *strange, weird*) could also be investigated (Potts et al. 2015: 151).

2.2 Analysis

2.2.1. Adjectives

Analyzing the frequency of adjectives and their context gives insight into the portrayals of the obituary subjects. Table 1 presents the top 15 most frequent adjectives from both corpora side by side. There are visible similarities and differences. The first two most frequent adjectives are the same in both – *first* and *other* – and most of the top 15 are the same adjectives in both corpora, but with a different placement, with a few exceptions. The *Overlooked* corpus' top 15 most frequent adjectives include adjectives that could possibly be attributed to race (e.g., the

words *black* and *white*). The top 15 most frequent adjectives of the corpus of regular obituaries includes words to do with Timeliness (e.g., *former* and *late*), which do not appear in the top 15 most frequent adjectives of the *Overlooked* corpus. These similarities and differences will be explored further in the following analysis section.

No.	Item	Frequency	Freq. Per Million
1.	first	2356	1409.22408
2.	other	2028	1213.03329
3.	more	1403	839.19413
4.	many	1308	782.37058
5.	early	1210	723.7526
6.	good	1039	621.47021
7.	own	999	597.5445
8.	young	908	543.11352
9.	new	889	531.74881
10.	American	812	485.69183
11.	last	803	480.30855
12.	high	710	424.68128
13.	former	691	413.31657
14.	several	682	407.93329
15.	late	629	376.23173

No.	Item	Frequency	Freq. Per Million
1.	first	526	1821.86585
2.	other	318	1101.43221
3.	many	268	928.25104
4.	more	261	904.00568
5.	own	218	755.06988
6.	young	194	671.94292
7.	black	178	616.52495
8.	early	163	564.5706
9.	new	161	557.64335
10.	good	158	547.25248
11.	American	147	509.15262
12.	white	137	474.51639
13.	few	134	464.12552
14.	female	130	450.27103
15.	several	117	405.24393

Table 1. Top 15 adjectives by frequency in *The New York Times* obituaries corpus (left) and the *Overlooked* corpus (right)

When comparing the top 50 adjectives in terms of frequency in both *The New York Times* obituaries corpus and the *Overlooked* corpus (Appendix 1 and Appendix 2), the word *first* is the most frequent, appearing 1409.2 times per million and 1821.9 times per million, respectively, being slightly more frequent per million in the *Overlooked* corpus. This word, as found through concordance analysis, most often conveys Novelty (e.g., *"*Marcia Freedman, who was the first American-born woman to serve in the Israeli Parliament /…/"). This is not in itself surprising as obituaries, in general, cover subjects who have made a significant contribution to a field in their lifetime.

The adjective *black* appears in the regular *The New York Times* obituaries' corpus as the 49th most common adjective (138.8 times per million). In comparison, in the *Overlooked* corpus, it is the 7th most frequent adjective (616.5 times per million). In the *Overlooked* corpus, it is found to be primarily used in the context of race. In contrast, in the corpus of regular obituaries, it more often occurs as a general descriptor of the color of something. When looking at which nouns are most often modified by *black*, the importance of race within the *Overlooked* corpus becomes apparent, as most often *black* modifies the nouns *ant, hole, and stain* in the corpus of *The New York Times* obituaries, but the nouns *community, woman,* and *people* in the corpus of *Overlooked* obituaries. Again, this could be construed as Novelty, as the subjects' race is highlighted as a point of difference and a means of correcting biases in previous representations of the past. The subjects within the *Overlooked* series are used as representatives of a larger whole, similar to Boyce and Dove's (2022: 520) claims about the *Overlooked* series. The *Overlooked* series, in this case, is making space in the collective memory for Black people who still do not get enough recognition in the regular obituaries section.

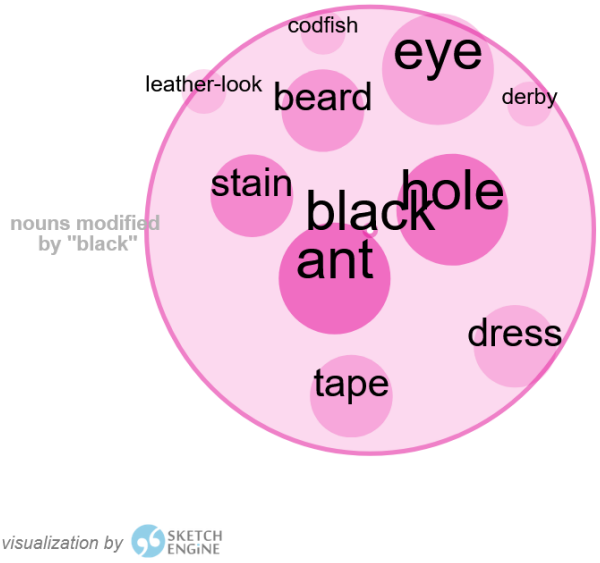


Figure 1. Nouns modified by *black* in the corpus of *The New York Times* obituaries.

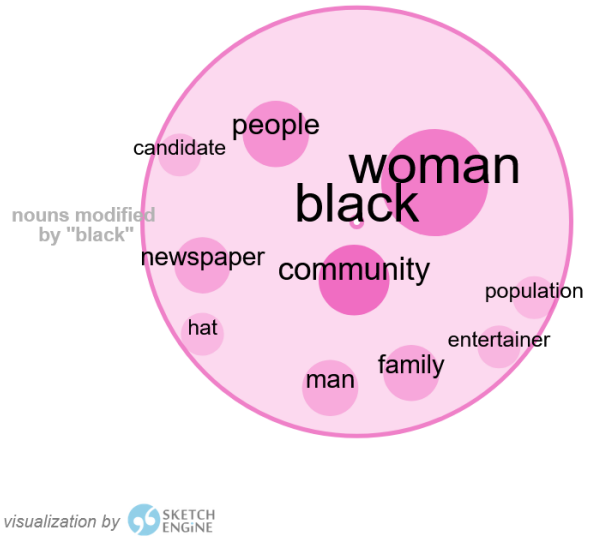


Figure 2. The nouns modified by *black* in the corpus of *The New York Times Overlooked* series obituaries.

W*hite* is the 12th most frequent adjective (474.5 times per million) in the *Overlooked* corpus but only the 42nd most frequent in the corpus of regular *The New York Times* obituaries (241.05 times per million). Both in the corpus of *The New York Times* obituaries and the *Overlooked* corpus, the word mainly occurs either as a descriptor of color or, more often, of race. Within the context of the latter, it usually conveys negativity (e.g., "Three months earlier, a white man shot and killed his father"). The fewer occurrences of the adjective *white* in *The New York Times* obituaries corpus in a racial context can be attributed to whiteness still being considered the unspoken norm and thus not deserving special mention. As the *Overlooked* corpus portrays people of color relatively more often, the white race is brought up as the contrast. When looking at which nouns are usually modified by *white,* the words *supremacy* and *supremacist* are more present within the regular obituaries (only a few occurrences in the *Overlooked* corpus). The regular obituaries corpus, as found through concordance, includes obituaries of actual white supremacists (e.g., "Historians say, however, that Mr. Patterson will

most likely be remembered as one of the most intractable white supremacists of his day /…/"), something that would likely never appear in the *Overlooked* series. There are also more generic allusions to the white race with a negative connotation as something to be opposed to ("/…/ knocked loose the stranglehold of white supremacy on the state /…/"). The *Overlooked* series refrains from using the words *white supremacy* but is more likely to use the adjective *white* in general to refer to race. This can be interpreted as trying to create accountability in the white reader – by using the adjective *white* to modify more words such as *man* or *people* and putting them in an opposing, negative context in relationship with the subject. This can create the effect of guilt through association, while using *white supremacy*/*supremacist* can alienate the average reader.

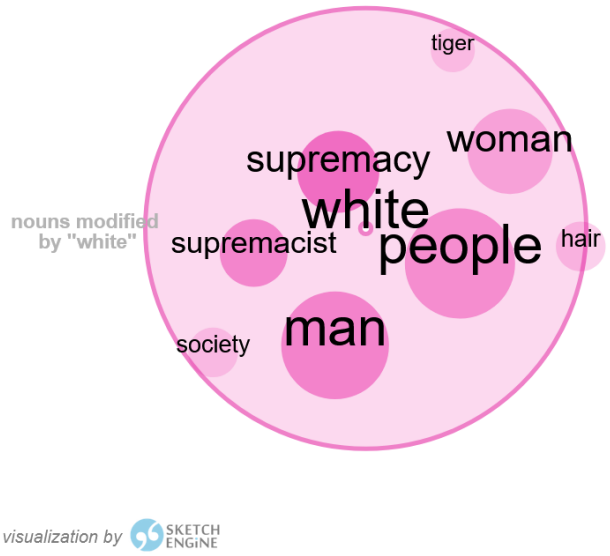


Figure 3. Nouns modified by *white* in the corpus of *The New York Times* obituaries.

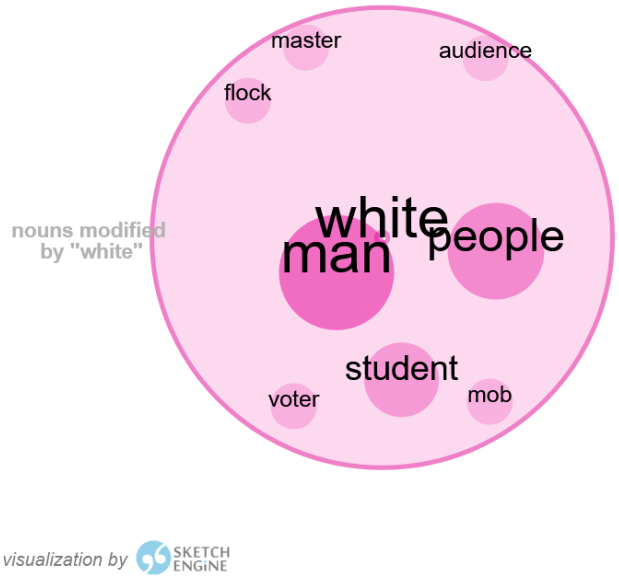


Figure 4. The nouns modified by *white* in the corpus of *The New York Times Overlooked* series obituaries.

The adjective *American* is present in both corpora's top 15 most frequent adjectives. It occurs in the corpus of regular *The New York Times* obituaries 485.7 times per million and in the *Overlooked* corpus 509.2 times per million, being 10th most frequent and 11th most frequent, respectively. *The New York Times* is an American publication, so references to American people are to be expected. The adjective *American* in the corpus of regular obituaries modifies nouns such as *history, life, flag, woman, citizen, troops,* and *culture*. The modification of the noun *woman* by *American* is interesting, as the concordance shows that often the usage cases have to do with different achievements, and *American* is also preceded by another adjective like *Asian* or *African* that creates a racial meaning, not the meaning of Proximity through national belonging (e.g., "/…/ the first **African American woman** to win Wimbledon"). *Woman* is the most common noun modified by *American* in the corpus of *Overlooked* obituaries with similar context patterns as in the corpus of regular obituaries, with the difference that in the corpus of regular obituaries, the most common noun modified by *American* is "history". This shows the regular obituaries cycle's need to justify the subjects as a part of history, specifically American

history, as Proximity is likely to be important to the readership. The *Overlooked* corpus also features, as most frequent nouns modified by *American*, the noun *history* and *citizen*, but also *audience* and *soldier*. Nouns that deal with war are present in both corpora in connection with the word *American*. This and the presence of the phrase *American citizen* shows the presence of patriotic values and, indirectly, also an interest in shaping the collective memory.

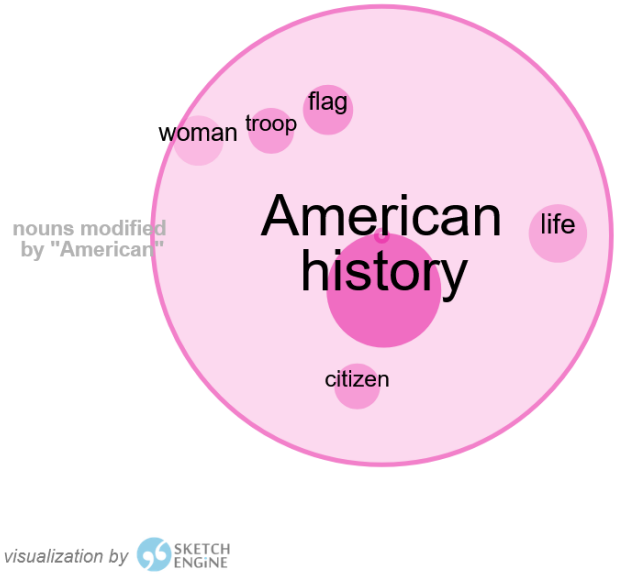


Figure 5. The nouns modified by *American i*n the corpus of *The New York Times* obituaries.

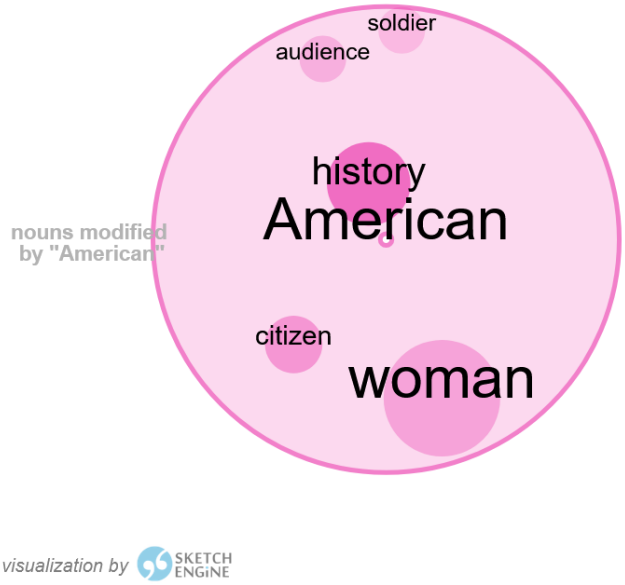


Figure 6. The nouns modified by *American* in the corpus of *The New York Times Overlooked* series obituaries.

Another adjective of interest is *former*, being the 13th most frequent adjective (413.3 times per million) in the corpus of ordinary *The New York Times* obituaries but only 39th (214.7 times per million) in the *Overlooked* corpus, occurring only half as often. The word occurs in the main corpus of obituaries mainly in the context of interviews with people close to the subject, naming their former occupations ("In recent years, fans would leave flowers and notes on a small table in the hallway outside Ms. Grossman's door, said Mr. Lambert, her **former** neighbor"), most probably to legitimize their opinion on the subject by showing their position towards them. The same pattern can be observed in the *Overlooked corpus.* Differences in usage become visible when looking at the nouns the adjective *former* modifies. In the main *New York Times* corpus, some of the most frequently modified nouns deal with Eliteness, such as *president* (21.5 times per million), *editor* (13.6 times per million), and *director* (10.8 per million). Within the *Overlooked* corpus, the most commonly modified noun by *former* is *slave* (17.32 per million). This can be interpreted as further proof that the *Overlooked* obituaries often fit into

Fowler's (2005) category of untraditional yet positive obituaries with a less linear and less traditional path of success, especially when accounting for the context of occurrence. Within the *Overlooked* corpus, the phrase *former slave* occurs only in plural and is always in the context of the subject helping the formerly enslaved. Other nouns modified by *former* in the *Overlooked* corpus are not present in significant enough numbers or are only present within one obituary.

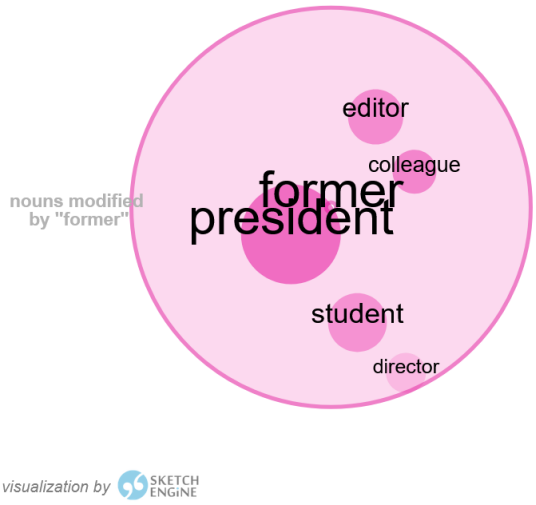


Figure 7. Nouns modified by *former* in the corpus of *The New York Times* obituaries.

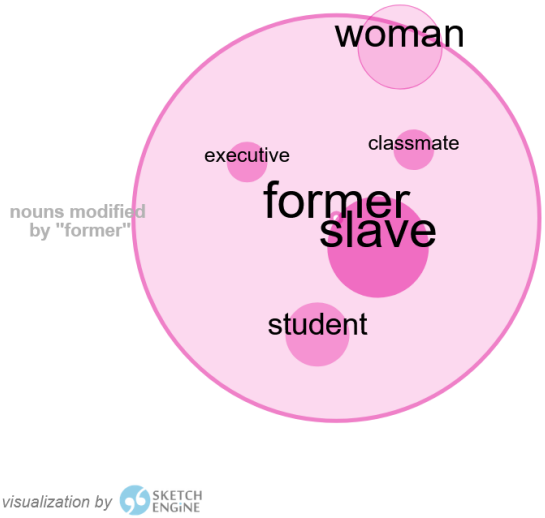


Figure 8. Nouns modified by *former* in the corpus of *The New York Times Overlooked* series obituaries.

The adjective *female* is present in the top 50 most frequent adjectives within the *Overlooked* corpus (14th; 450.3 times per million) but not within the corpus of regular *The New York Times* obituaries (99.3 times per million). Based on the concordance, it can be said that the adjective *female* in the *Overlooked* corpus mainly appears in the contexts that explicitly highlight the gender of the noun it describes (e.g., "Khan was the first **female** radio operator to be sent by Britain into occupied France /…/"). Again, this could be explained by the aim of the series to highlight people previously missing from obituaries, including women. The regular obituaries cycle does not need to highlight this difference or use this aspect as Novelty, and thus, the adjective is not as frequently used. By using language that creates Novelty with the help of the obituary subjects, the collective memory of the person is inherently tied to the aspect that is being portrayed as novel, such as the subject's gender. The word within the corpus of regular obituaries occurs in similar contexts but with much lesser frequency, indicating the lesser need to point out the Novelty of the subject.

2.2.2. Nouns

Analyzing the frequency of the nouns within the two corpora gives an overview of who the subjects are and what is essential as a subject within the obituaries, and thus, what is portrayed as worthy for the collective memory.

The top 15 most frequent nouns in both the corpus of *regular New York Times* obituaries and the corpus of *Overlooked* obituaries (see Table 2), differently from the similar tables about adjective frequency, have more differences than similarities. Only a few words occur in both top 15's – *year, people, school, work, book*. Others are different for the two corpora. Both corpora's top 15 include *new, York,* and *time,* as the newspaper the data is derived from is often mentioned within the obituaries and thus will not be further examined as the usage of the words in the contexts outside of the newspaper's title would be difficult to distinguish with the analytical tools used in this thesis.

No.	Item	Frequency	Freq. Per Million
1.	mr.	13343	7981.01734
2.	ms.	4664	2789.73731
3.	year	4250	2542.10625
4.	new	3024	1808.78337
5.	time	2593	1550.98388
6.	york	2345	1402.64451
7.	dr.	2061	1232.772
8.	people	2012	1203.46301
9.	university	1925	1151.4246
10.	school	1906	1140.05989
11.	work	1851	1107.16204
12.	father	1794	1073.06791
13.	book	1684	1007.27222
14.	world	1611	963.60781
15.	son	1608	961.81338

No.	Item	Frequency	Freq. Per Million
1.	woman	1102	3816.91287
2.	year	669	2317.16399
3.	time	556	1925.77455
4.	life	415	1437.40367
5.	new	407	1409.69468
6.	people	350	1212.26815
7.	world	344	1191.48641
8.	child	339	1174.1683
9.	work	328	1136.06844
10.	school	326	1129.14119
11.	man	325	1125.67757
12.	family	323	1118.75032
13.	book	320	1108.35945
14.	mother	287	994.05989
15.	york	282	976.74177

Table 2. Top 15 nouns by frequency in *The New York Times* obituaries corpus (left) and the *Overlooked* corpus (right)

Comparing the top 50 most frequent nouns extracted with the 'word list' tool in Sketch Engine from both corpora provides interesting results. The most frequently used noun within the *Overlooked* corpus is *woman*, present 1102 times (3816.9 times per million), while this noun does not occur within the top 50 most frequent nouns of the corpus of regular obituaries at all. This aligns with the series' original goal, which is to highlight important women as part of the International Women's Day celebration. While pointing out the female gender can be seen as a positive move towards more inclusive representation, it can also create an opposite effect. Women can, through the use of this strategy, appear as a Novelty, stressing that men are, in this instance, still considered the norm. For example, in "/…/ Rice became one of the first **women** to graduate from the architecture program /…/" the subject's gender is brought up to show that historically, this field had been mostly male-dominated. The subject changed this *status quo*, but this can be only highlighted by implying the norm was men. The series' goal to remedy one-sided representation within collective memory can only be achieved by this language of Novelty, which the *Overlooked* obituaries employ.

The modifiers of the noun *woman* further show the use of Novelty. The five modifiers most commonly occurring with *woman* are *first*, *young*, *Black*/*black*, *other*, and *only*. Here all but *other* can immediately indicate Novelty, explicitly showing why the subject is worthy of being honored with a retrospective obituary in addition to already showing Novelty by referring to gender (e.g., "She is also the only **woman** to have done it alone."). The modifier *other* together with the noun *women* in the *Overlooked* corpus is mainly used when the subject of the obituary achieved something together with other women who are not mentioned by name but were present (e.g., She and other working-class **women** in her Brooklyn neighborhood formed tenants' unions, /…/").

No.	Collocate	Frequency	Freq. Per Million
1.	first	51	176.64
2.	Black/black	45	155.86
3.	young	30	103.91
4.	other	17	58.88
5.	only	12	41.56

Table 3. Top 5 modifiers of the noun *woman* by frequency in *The New York Times Overlooked* corpus.

Nouns *Mr. and Ms.* are the first and the second in terms of their frequency in the top 50 most frequent nouns in the regular *New York Times* obituaries corpus (7981.0 times per million and 2789.7 times per million) but are not found within the *Overlooked* corpus. This could indicate the creation of Eliteness within the normal obituary cycle, and the use of polite titles could possibly be due to the short time between the death of the subject and the written obituary demanding more respect. Contrastingly, *Mr.* is used in the *Overlooked* corpus only 22 times (76.2 times per million) and Ms. 38 times (131.6 times per million), showing a vast difference in usage. The concordance shows that within the regular obituaries' corpus, both *Mr*. and *Ms*. are used to refer to the subject consistently throughout the obituary. However, within the *Overlooked* corpus, when it comes to *Ms.,* the usage is mainly restricted to references to one subject – Cuban printmaker Belkis Ayón. It is not used for any other subjects. When it comes to *Mr.,* the concordance shows that it is mainly used as a polite title for men who have written about the subject of the obituary (e.g., "She was survived," **Mr.** Miller wrote in his biography, "by her husband, Earle, her sister Lavada, her brother Arvada and her stepbrother Arthur (Artemus) Bush /…/").

Dr. is another title that appears among the top 15 most frequent nouns within the corpus of regular obituaries but not within the corpus of *Overlooked* obituaries. *Dr*. is the seventh most frequent noun within the corpus, occurring 2061 times (1232.8 times per million). This title has a significantly smaller presence within the *Overlooked* corpus, only occurring 65.8 times per

million. In the *Overlooked* series, the concordance tools show the title can be seen chiefly in connection with male doctors who appear within the subject's story (e.g., "*In 1862, she married* ***Dr.*** *Charles H. Gordon, a Scot /…/*"). Another occurrence within the concordance hints at why the frequencies are so different. The *Overlooked* obituary of doctor Elizabeth Hayes, where the title *Dr*. appears, speaks of her Novelty as "not just brave and principled but also young (she was 33) and a rare female doctor". This explains why there is such a difference in frequency. As the people covered in the *Overlooked* series often represent those pioneering in their different fields, there will be a distinct lack of other representations of the same groups within *Overlooked*. When trying to represent everyone previously unrepresented, there is bound to be a certain lack of repetition to keep things new and always representative of new groups or people. Another explanation is that marginalized groups who are usually covered in the section did not often have the opportunity to acquire advanced degrees.

People is a noun with a similar frequency in both corpora – occurring 1203.5 times per million in the regular *New York Times* obituaries corpus and 1212.3 times per million in the *Overlooked* corpus. Notably, the noun occurs most frequently in both corpora in the plural form instead of the singular *person.* This indicates that it displays Proximity by being an indication of community and its impact on it or of it. This sense of community can also be seen when looking at the modifiers of the noun *people*. Within the corpus of regular obituaries, the most common modifiers are *many, Black,* and *other;* within the *Overlooked* corpus *many*, *white,* and *black* – all invoking a certain public or community.

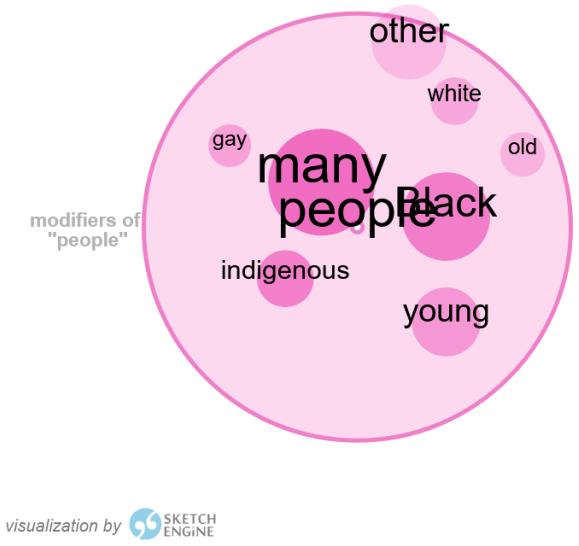


Figure 9. Modifiers of the noun *people* in the corpus of *The New York Times* obituaries.

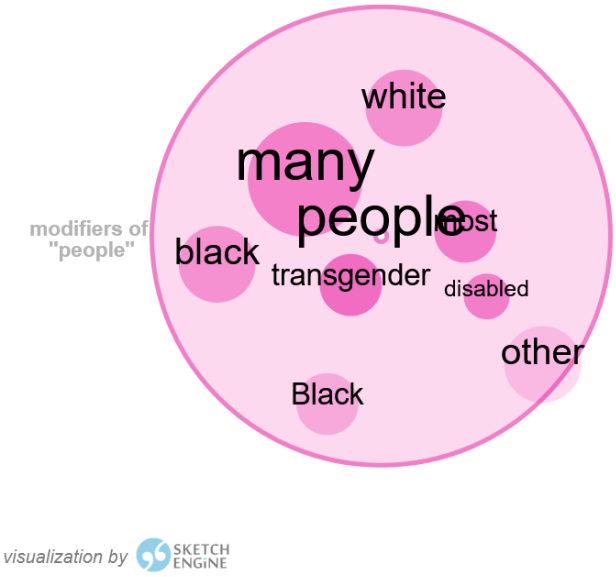


Figure 10. Modifiers of the noun *people* in the corpus of *The New York Times Overlooked* series obituaries.

Within both corpora, the use of the noun *people* can be most often seen in one context: the impression or impact the subject has left on the community (the people), either positive (e.g., " /…/ he was acclaimed for changing the way Australia's Indigenous people were portrayed and viewed") or negative (e.g., "After being interrogated for four months, she was deemed an enemy

of the people and executed"). Within the *Overlooked* corpus, the community is often brought up regarding the perceived injustices they were subjected to, while the corpus of regular obituaries mostly speaks of the positive impacts and impressions. Implicating an entire community in the injustices committed against a subject can create a sense of guilt in the reader, possibly helping to further the ideas carried by the *Overlooked* series. The slight difference in usage between the two corpora can also be attributed to the temporal differences in the time of publishing, as when the person has recently passed, the writing demands more neutral to positive language use.

Work as a noun is present among both corpora's top 15 most frequent nouns – 1851 times (1107.2 times per million) in the corpus of regular obituaries and 328 times (1136.1 times per million) in the corpus of *Overlooked* obituaries. The concordance analysis shows that the subjects' work and achievements are important when it comes to having an obituary published, regardless of whether it is published closely after death or retroactively. Within the regular obituary cycle, the noun *work* is often even present within the first line of the obituary, right under the title (e.g., "George Ferencz, Innovative Theater Director, Dies at 74. He directed **works** by Sam Shepard, Amiri Baraka and others at La MaMa and similar theaters known for experimentation"), showing the importance of achievement in a specific field within the subjects' life story. This again justifies the publication of the given subject's obituary. The most common modifiers of *work* within the corpus of regular obituaries are *new, own, early, other,* and *hard,* showing that obituaries prefer novel work that spans a longer timeline and has taken effort. The most frequent modifiers of the noun *work* in the *Overlooked* corpus are *good* (e.g., *her best work*), *social, early,* and *published,* showing a similar demand for linearity but also a more significant emphasis on social impact, as the section is interested not just in individual achievements but also their broader impact. The emphasis on being published can be explained by the fact that the people written about were relatively unknown at their death.

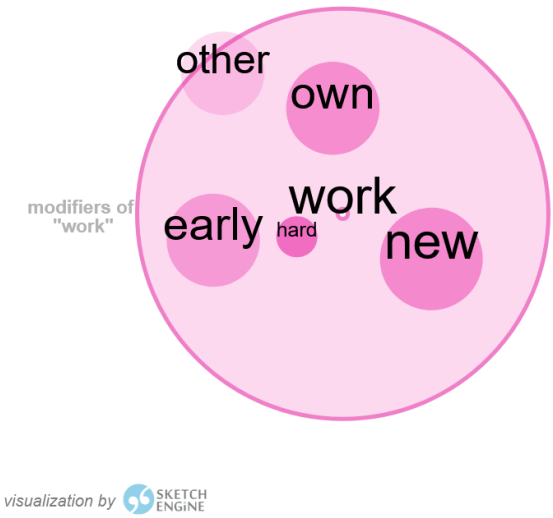


Figure 11. Modifiers of the noun *work* in the corpus of *The New York Times* obituaries.

Figure 12. Modifiers of the noun *work* in the corpus of *The New York Times Overlooked* series obituaries.

The noun *life* is the 4th most frequent noun in the *Overlooked* corpus, occurring 1437.4 times per million, and the 16th most frequent in the corpus of regular obituaries, occurring 955.2 times per million. The verbs most frequently occurring with the noun *life* as object in the corpus of regular *The New York Times* obituaries are mostly either positive or neutral – *live* (*lived the life*), *be* (*was her life*), *save* (*saved my life*), *have* (*had a glamorous life*), *change* (*changed her*

life) etc. The most frequent verbs occurring with the noun *life* as object in the *Overlooked* corpus, however, also include more verbs expressing Negativity – *take* (*took her own life*), *lost* (*lost their lives*) – as well as positive ones. *Take* is the second most frequent verb with *life* as object in the *Overlooked* corpus, indicating that suicide is no longer considered a taboo, and this also changes the kinds of deaths that can be openly added to the collective memory.

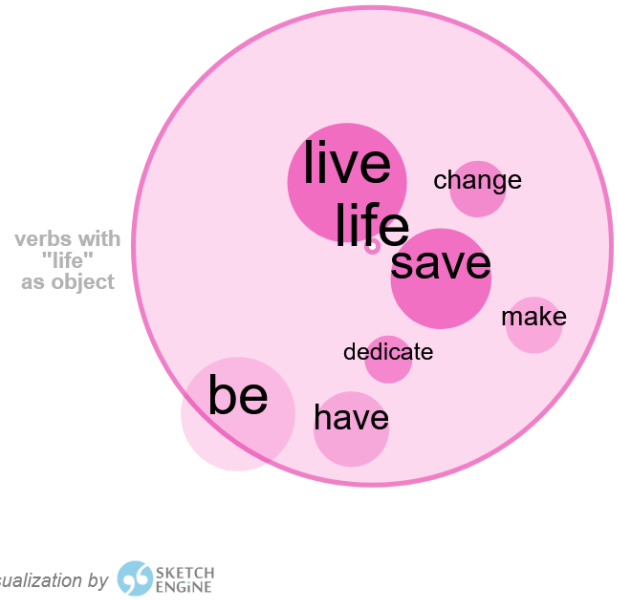


Figure 13. Verbs with *life* as object in the corpus of *The New York Times* obituaries.

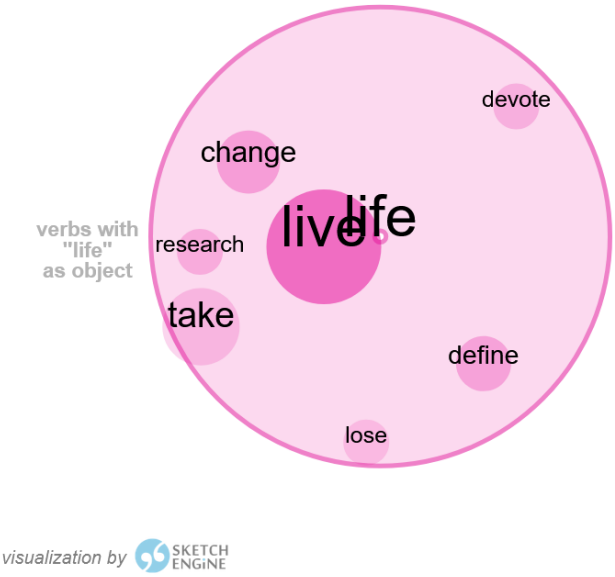


Figure 14. Verbs with *life* as object in the corpus of *The New York Times Overlooked* series obituaries.

Words dealing with family are also of interest in this comparison of frequencies. The noun *mother* does not occur within the top 15 nouns of the corpus of regular obituaries and instead is the 17th most frequent noun. However, it is the 14th most frequent noun within the *Overlooked* corpus. The relative frequencies are similar, despite this: 938.5 times per million in the corpus of regular obituaries and 994.1 times per million in the corpus of *Overlooked* obituaries. The word *father* contrasts the word *mother* as it is present in the top 15 most frequent nouns of the corpus of regular obituaries (1073.1 times per million) but is only 19th within the *Overlooked* obituaries (834.7 times per million) but with a more considerable difference in relative frequency.

The most frequent modifier of the word *mother* in both corpora is *single* (*single mother*). The other modifiers in the corpus of regular obituaries are *birth* (*birth mother*) and, in the *Overlooked* corpus, *own (her own mother*), both emphasizing the importance of the relationships between mothers and their children. This personal aspect is also seen in the most frequent

prepositional phrases including *mother*. In the corpus of regular obituaries, the most frequent prepositional phrases with mother are *by 'mother'* and *with 'mother'*. Prepositional phrase *by 'mother'* occurs most often within the common obituary genre feature *survived by* (*is survived by his mother*) and with *raised* (*raised by his mother*). The latter especially emphasizes the importance of early family upbringing in which the mother plays a decisive role. The prepositional phrase *with 'mother'* occurs in multiple different contexts, usually referring to aspects of life in which the mother played a central role, e.g., *lived with his mother* and *moved with her mother*. The prepositional phrase *with 'mother'* is the most frequent within the *Overlooked* corpus, occurring in similar contexts. *By 'mother',* however, is nearly non-existent, as the genre feature phrase *survived by* is largely irrelevant in the *Overlooked* obituaries due to the often significant time lag between the recorded death and the writing making references to the surviving relatives irrelevant.

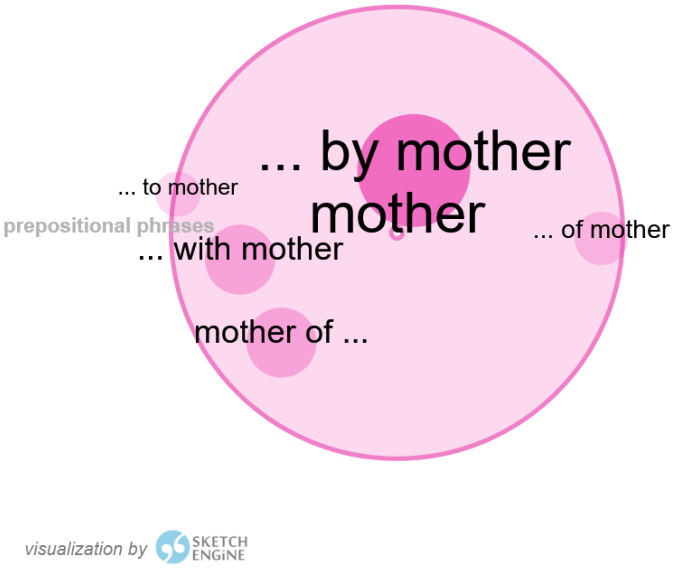


Figure 15. Prepositional phrases of the noun *mother* in the corpus of *The New York Times* obituaries.

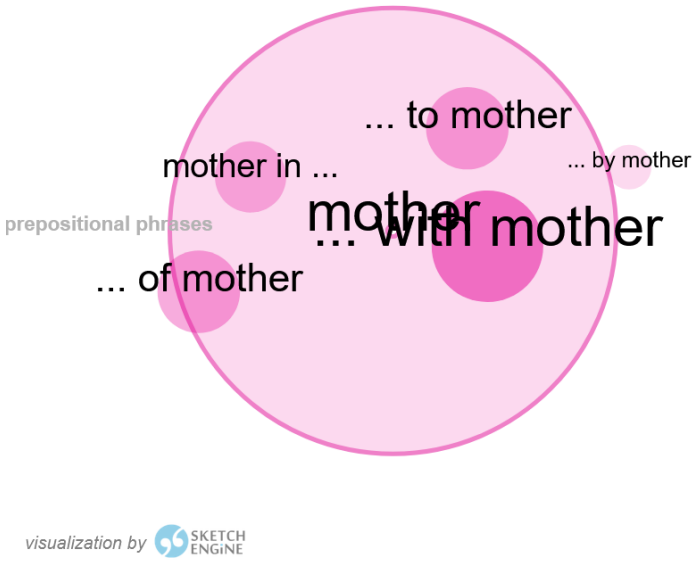


Figure 16. Prepositional phrases of the noun *mother* in the corpus of *The New York Times Overlooked* series obituaries.

The noun *father* does not appear in significant patterns with any modifiers. The most frequent prepositional phrase with *father* in the corpus of regular *New York Times* obituaries is *with 'father',* which appears in similar contexts as *with 'mother'*. The second most frequent is the construction *'father' of*, mostly used to indicate that the subject was the founder of something and thus acting as an indicator of Eliteness (e.g., *father of the atomic bomb),* while a similar *'mother' of* construction is less frequent overall, showing gendered bias. Another interesting occurrence is the 4th most frequent prepositional phrase with *father – from 'father'*. The things a subject can receive from their father include *beating* ("*receiving severe beatings from his father*"), *wisdom,* and *training.* Thus, it can be said that while mothers are mentioned in order to stress personal connections, fathers mostly appear as authority figures or mentors. This echoes the difference in how parental roles are viewed to this day. Most frequent prepositional phrases with *father* in the *Overlooked* corpus are *to 'father'* and again *'father' of* the former appearing

in the context of the obituary subjects' stories being conveyed through texts, for example, "*wrote to her father"* or "*letter to her father"*.

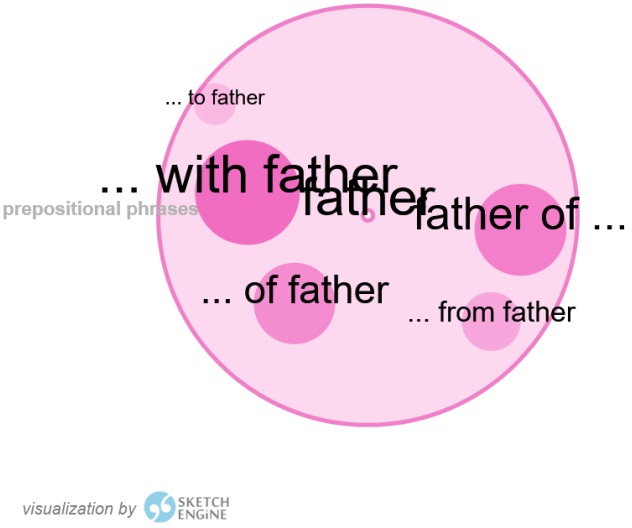


Figure 17. Prepositional phrases of the noun *father* in the corpus of *The New York Times* obituaries.

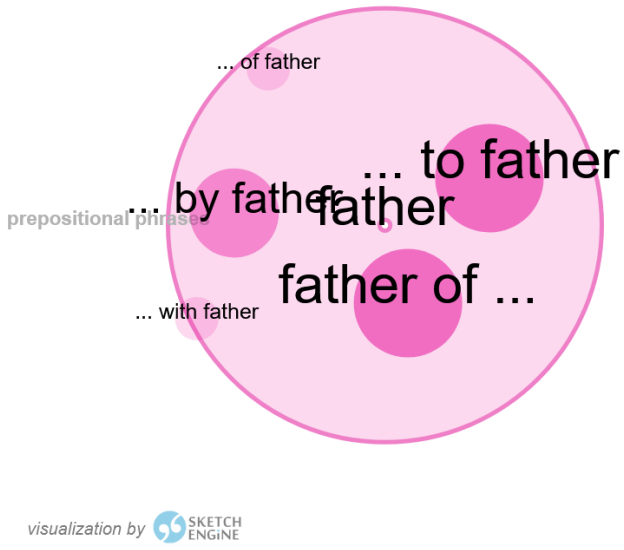


Figure 18. Prepositional phrases of the noun *father* in the corpus of *The New York Times Overlooked* series obituaries.

The noun *child* appears within the top 15 most frequent nouns of the *Overlooked* corpus (1174.2 times per million) and is 19th most frequent in the corpus of regular obituaries (875.1 times per million). The most frequent pronominal possessors of *child* in the corpus of regular obituaries is *their,* followed by *her* and only then by *his*. Thus, the predominant focus is on the familial unit as one, followed by the mother's relationship with the child, although the most common subject of the obituaries are men. The male obituary subjects do not always have their children mentioned in the obituaries. When child-parent relations are mentioned, they are always connected with the entire familial unit (*their children)*. In the corpus of the *Overlooked* obituaries, the most frequent pronominal possessor of *child* is *her*, followed by *their* and again, only then *his*. In this case, *his* is expected to be the least frequent as most subjects of the *Overlooked* obituaries are women. Here *her* is more frequent than *their*, suggesting that the traditional family unit of man-woman-child is backgrounded and the singular, female subjects of the obituary and the relationship they have with their child/children is foregrounded. The focus on the subject can be an attempt to shift the attention away from the traditional gender norms and to move towards more progressive ideals.

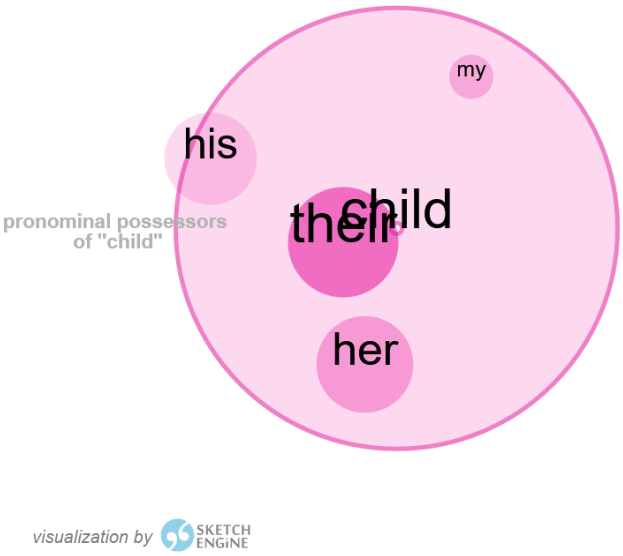


Figure 19. Pronominal possessors of the noun *child* in the corpus of *The New York Times* obituaries.

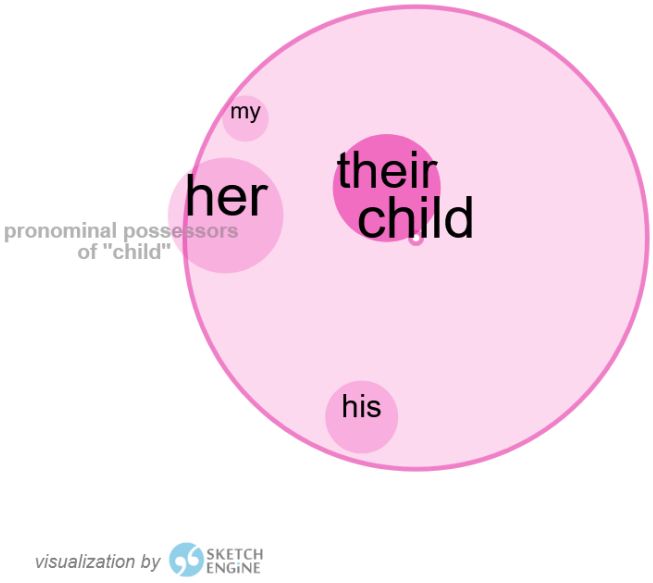


Figure 20. Pronominal possessors of the noun *child* in the corpus of *The New York Times Overlooked* series obituaries.

Nouns dealing with education are also present within the top 15 most frequent nouns of both corpora. *School* appears in both top 15 most frequent noun lists with a similar frequency. The noun *school* occurs in the corpus of regular obituaries 1140.1 times per million and in the

Overlooked corpus 1129.1 times per million. The modifiers of *school* can give insight into the types of institutions the subjects completed (other than university education). The most frequent modifier of *school* in both the corpus of regular obituaries and the corpus of *Overlooked* obituaries is *high* (*high school*). The context shows that *high school* appears within both corpora as a part of the subject's life story, often trying to create a sense of humble beginnings, especially if the subject did not complete high school (e.g., "Having dropped out of high school, Ms. Minner was determined to make something of herself /…/"). This helps create a link to the American Dream and the belief that one can succeed if one works hard. The corpus of *Overlooked* obituaries features the modifier *public* (*public school*) as the second most frequent modifier of *school,* and it is also frequently present in the corpus of regular obituaries, again a possible tool for bringing the subject closer to the reader as someone who did not have the privilege of attending private school. The corpus of regular obituaries also has the noun *school* frequently modified by *law, medical, private,* and *Catholic*, with the corpus of *Overlooked* obituaries having similar modifiers but also *segregated* and *boarding*, showing the difference between the average subjects portrayed by the different obituary types.

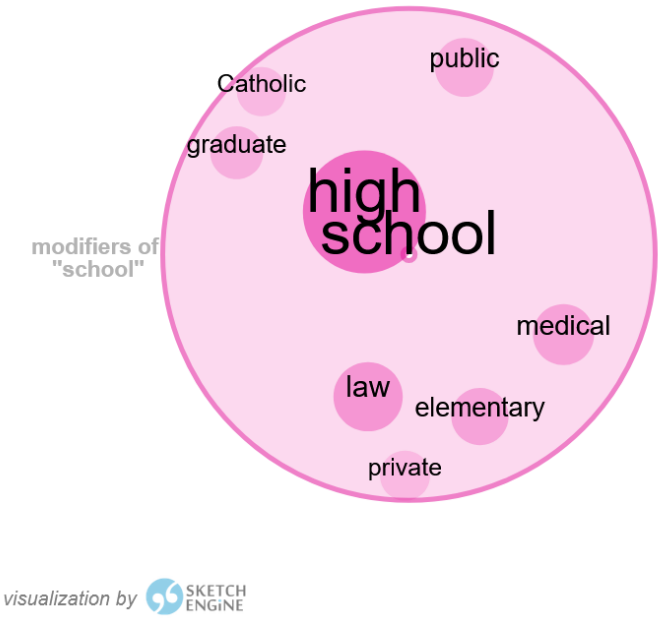


Figure 21. Modifiers of the noun *school* in the corpus of *The New York Times* obituaries.

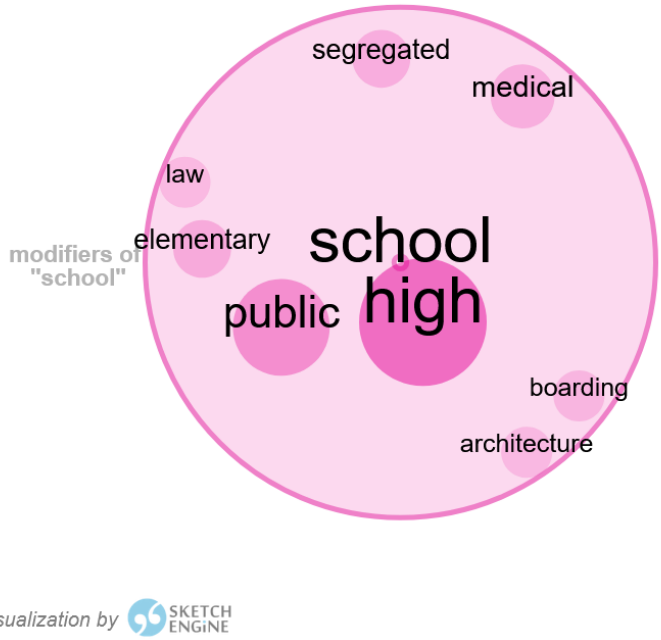


Figure 22. Modifiers of the noun *school* in the corpus of *The New York Times Overlooked* series obituaries.

The noun *university* appears among the top 15 most frequent nouns of the corpus of regular obituaries (as the 9th most frequent noun (1151.4 times per million)) but is only the 22nd most frequent noun of the *Overlooked* corpus (744.7 times per million). The concordance shows

that in both corpora, the noun is mostly used within the names of different universities from which the subjects have graduated. The higher frequency within the corpus of regular obituaries might indicate that the subjects of the common obituaries' cycle are expected to finish higher education. The *Overlooked* obituary subjects can follow an unconventional path to success, which often does not include traditional education.

This analysis of different word lists and frequencies within the *New York Times* corpus of obituaries and the *Overlooked* corpus has shown that the two types of obituaries use distinctly different strategies in representing their subjects. The language of the two types of obituaries differed regarding how collective memory was constructed. From the different choices in frequent vocabulary use, it was possible to discern that the traditional obituaries cycle is trying to become more inclusive, but it still relies, above all, on eliteness in representing its subjects, while the *Overlooked* series is more likely to present non-traditional paths to success. The frequency of nouns showed that race and gender are brought up primarily to create Novelty. Similar strategies were used with adjectives to create oppositions with the average readership to create guilt within the reader and enforce a particular collective memory. While this strategy is understandable, it also re-creates the sense that white men continue to be the primary category of people whose lives deserve commemoration.

CONCLUSION

The obituary is a genre used for recounting people's lives after their death, helping to link the individual lives of people with collective memory, satisfying a given society's need for a shared sense of the past, and reflecting shared cultural values in mass media form. First, as known today, obituaries were inherently linked with class. Over time, the genre has shifted towards commemoration based on personal achievement and newsworthiness, reflecting changes in social values. What is considered worth reporting is constantly changing, and who is considered worthy of an obituary and how they are written about can influence collective memory and people's perceptions of different events, people, stories, and history.

The professionally-written obituary is a genre based on a series of subjective choices – who gets represented mainly depends on the story their life provides for the publication. While our societies have become more diverse, obituary editors admit that they do not usually strive for equal representation of subjects covered. A part of these inequalities can be attributed to the differences in merit, as it has been traditionally perceived. For example, the number of women in the workforce in the generations that are currently passing away was relatively small compared to today's situation. Hence, women could not reach the public status usually recognized in obituaries. Similarly, historically, people of color were only granted obituaries when their narratives were particularly newsworthy; the poor, social outcasts and most working-class people were also often omitted. While this is changing, there will be no sudden changes as long as obituaries focus on people who have achieved success in public sphere professions of great social prominence.

The *Overlooked* series of obituaries in *The New York Times* represents a more active method of addressing past gaps in collective memory. The series retroactively gives obituaries to subjects who did not receive an obituary at the time of their passing, with stories that often

contribute to different present-day debates. Obituaries have the benefit of retrospection – past events and people can be reframed in the light of present-day events and knowledge. While the series' goals are overarchingly positive, it often omits parts of the subjects' stories to achieve this larger positive narrative, handling them as a model for different ongoing debates.

Previous research into media and collective memory production speaks of a counter-memory, where historically marginalized people offer alternatives to the standard narrative. The *Overlooked* series fits into this counter-memory and the broader process of narrative struggle over the past. Like counter-memory, it only exists as an inversion of the dominant discourses, a product of the cycle of main obituaries with its continuing underrepresentation of different minorities. The media has also been found to be a site where people can make a personal connection with history, and noncontroversial commemorative events can become a site for building social consensus and a broader interpretation of history.

How the content of the obituaries is brought to the reader can also affect collective memory. Obituaries have been shown to have different genres, where the differences in how the story is told can shape the readers' interpretation of the factual content presented. For example, when the subject is portrayed negatively, this affects the emotional focus of how they are remembered. *Overlooked* could also be seen as an example of affirmative action, as it aims to increase the representation of those previously underrepresented within the obituaries *The New York Times* publishes. Previous research has shown that simply speaking of initiatives for diversity is not enough for change to occur, and real, meaningful action is needed.

Previous studies into *Obituaries* using corpus research have mainly been small-scale and have often only used qualitative methods. Previous research into collective memory in obituaries using corpus research or other computer-based methods was not found at the time of writing.

This is why, for the thesis, corpus-assisted discourse analysis was chosen as the method, as it combines the qualitative with the quantitative by working with discourse through corpora. It also permits greater objectivity, as there is no human component in finding different linguistic patterns. Methods used by Bednarek and Caple (2014) and Potts et al. (2015) to identify different news values were also employed to find how a subject's insertion into the collective memory could have been justified.

In order to compare the regular and *Overlooked* obituaries, two corpora were created to find out what the vocabulary used can tell us about collective memory production within the two types of obituaries in the *New York Times*. The first corpus was created from the regular obituaries published in *The New York Times* from January 1, 2021, to December 31, 2021, and the second from the entire *The New York Times' Overlooked* series up to and including December 2023. The first corpus contained 1,388,794 words (1,294 individual texts, average word count ~1,073 words), and the second 241,153 words (195 individual texts, average word count ~1,237 words).

When looking at the adjectives by their frequency within both corpora, it was found that the words modified by adjectives often show much about how collective memory might be produced within obituaries. Adjectives with possible racial connotations were shown to often modify words more likely to deal with negative subjects in the *Overlooked* obituaries, using guilt as a tool to invite the readers to reflect on the past. Adjectives dealing with potential temporal references regarding life stages, such as employment, show that the regular obituaries feature more professions dealing with Eliteness while *Overlooked* does not, emphasizing instead the untraditional yet positive life paths. Different adjectives dealing with the aspects that differentiate the subject from the typical subject of the regular obituaries (for example, race, gender) add Novelty. While this is done with the best intentions, it unfortunately also recreates

stereotypes and places minorities in the counter-memory instead of challenging the traditional narrative of social worth.

The frequency of nouns showed what subjects are deemed normal within obituaries and how they are portrayed. Again, race and gender are brought up primarily to create Novelty. While different titles are the most common nouns within the corpus of regular obituaries, these are very uncommon among the *Overlooked* obituaries. Titles, for example, *Ms*., are nearly wholly absent among *Overlooked* obituaries.

The regular obituaries cycle and *Overlooked* obituaries evoke different communities. The *Overlooked* obituaries often bring up different communities when talking about injustices committed against them to raise the awareness of the perhaps privilege-blind white male reader. Familial relations are among the most frequent nouns in both corpora, with more sensitivity to more inclusive familial language appearing in the *Overlooked* obituaries. Vocabulary dealing with education reveals that the regular obituary cycle demands higher education from its subjects. In contrast, the *Overlooked* subjects can more often follow less conventional education paths.

The present study demonstrated that while the *Overlooked* obituaries show a conscious move towards more inclusive language and an effort to change public memory of its subjects, they are still often portrayed as just exceptions in the larger social narrative. The *Overlooked* obituaries often use language to create Novelty even more than the regular obituaries cycle. The *Overlooked* series is a valiant attempt to change collective memory on the background of a mainly white, male-dominated regular obituaries cycle. It has gained positive attention from its audience but leaves its subjects to only serve counter-memory rather than change the dominant narrative in the regular obituary cycle.

While the thesis at hand produced intriguing exploratory results about the language that shapes collective memory, more corpus-based research could be done into other media genres to focus on collective memory creation and other affirmative action media projects to see if the language of creating collective memory is similar to those within obituaries.

REFERENCES

Bednarek, Monika ja Helen Caple. 2014. Why do news values matter? Towards a new methodological framework for analysing news discourse in Critical Discourse Analysis and beyond. *Discourse & Society*, 25: 2, 135-15

Boyce, Charlotte and Danielle Mariann Dove. 2022. Obituary, gender, and posthumous fame: the New York Times Overlooked project. *Celebrity Studies*, 13: 4, 507-523.

Edy, Jill A. 1999. Journalistic Uses of Collective Memory. *Journal of Communication*, 49: 2, 71–85.

Fairclough, Norman. 2015. *Language and Power*. Third edition. London, New York: Routledge, Taylor & Francis Group

Fowler, Bridget, and Esperança Bielsa. 2007. The lives we choose to remember: a quantitative analysis. In Fowler, Bridget (ed). *The Obituary as Collective Memory*, 129-156. New York: Routledge.

Fowler, Bridget. 2005. Collective memory and forgetting: components for a study of obituaries. *Theory, Culture & Society*, 22: 6, 53-72.

Fullinwider, Robert. 2018. Affirmative Action. In Zalta, Edward N. (ed). *The Stanford Encyclopedia of Philosophy.* Available at https://plato.stanford.edu/archives/sum2018/entries/affirmative-action/, accessed 11 May 2024.

Gilens, Martin, Paul M. Sniderman and James H. Kuklinski. 1998. Affirmative Action and the politics of realignment. *British Journal of Political Science*, 28: 1, 159–183.

Gillings, Mathew, Gerlinde Mautner and Paul Baker. 2023. *Corpus-Assisted Discourse Studies*. Cambridge: Cambridge University Press.

Gubitz, S. R. and Denzel Avant. 2020. Racializing Captain America: How racial attitudes affect perceptions of Affirmative Action and diversity initiatives in media. *Journalism & Mass Communication Quarterly*, 97: 3, 683–703.

Haskins, Ekaterina. 2007. Between archive and participation: public memory in a digital age. *Rhetoric Society Quarterly*, 37: 4, 401–422.

Herring, Cedric and Loren Henderson. 2012. From Affirmative Action to diversity: toward a critical diversity perspective. *Critical Sociology*, 38:5, 629–643.

Hume, Janice and Bonnie Bressers. 2010. Obituaries online: new connections with the living—and the dead. *OMEGA – Journal of Death and Dying*, 60: 3, 255-271.

Hume, Janice. 2000. *Obituaries in American Culture*. Jackson: University Press of Mississippi.

Hume, Janice. 2010. Memory matters: the evolution of scholarship in collective memory and mass communication. *Review of Communication*, 10: 3, 181-196.

Hunston, Susan. 2010. *Corpora in Applied Linguistics.* Cambridge: Cambridge University Press.

Kliger-Vilenchik, Neta, Yarif Tsfati and Oren Meyers. 2014. Setting the collective memory agenda: examining mainstream media influence on individuals' perceptions of the past. *Memory Studies*, 7: 4, 484-499.

Lester, Marilyn. 1980. Generating newsworthiness: the interpretive construction of public events. *American Sociological Review*, 45: 6, 984–94.

Marchi, Anna and Charlotte Taylor. 2018. Introduction: partiality and reflexivity. In Marchi, Anna and Charlotte Taylor (ed). *Corpus Approaches to Discourse: A Critical Review*, 1-15. London/New York: Routledge.

McDonald, William. 2018. From the death desk: why most obituaries are still of white men. The New York Times. Available at www.nytimes.com/2018/03/08/obituaries/overlooked-from-the-death-desk-why-most-obits-are-still-of-white-men.html, accessed 21 May 2022.

McDonald, William. 2022. How the Times decides who gets an obituary. *The New York Times*. Available at https://www.nytimes.com/article/new-york-times-obituary-process.html, accessed 02 February 2024.

Misztal, Barbara. 2003. *Theories of Social Remembering*. Maidenhead: Open University Press.

Moore, Stephen H. 2002. Disinterring ideology from a corpus of obituaries: a critical post mortem. *Discourse & Society*, 13: 4, 495–536.

Potts, Amanda, Monika Bednarek and Helen Capleet al. 2015. How can computer-based methods help researchers to investigate news values in large datasets? A corpus linguistic study of the construction of newsworthiness in the reporting on Hurricane Katrina. *Discourse & Communication*, 9: 2, 149–172.

Schwartz, Barry. 1991. Iconography and collective memory: Lincoln's image in the American mind. *The Sociological Quarterly*, 32: 3, 301–319.

Taussig, Doron. 2017. Your story is our story: collective memory in obituaries of US military veterans. *Memory Studies*, 10:4, 459–473.

Zandberg, Eyal, Oren Meyers and Motti Neiger. 2012. Past continuous: newsworthiness and the shaping of collective memory. *Critical Studies in Media Communication*, 29:1, 65–79.

RESÜMEE

TARTU ÜLIKOOL
ANGLISTIKA OSAKOND

Karoliine Laaneoja

Production of Collective Memory in *The New York Times'* Obituaries and its *Overlooked* Series: A Corpus-Based Study (Kollektiivmälu tootmine The New York Timesis ja selle "Overlooked" seerias: korpusepõhine uurimus)

Magistritöö

2024

Lehekülgede arv: 67

Annotatsioon:

Järelhüüded on žanr, milles üksikud elud muutuvad kollektiivse mälu osaks selle kaudu, kuidas sündmusi ja inimesi kujutatakse. *The New York Times*'is ilmunud sarja „Overlooked" eesmärk on luua järelhüüded inimestele, kes oma surma hetkel neid erinevatel põhjustel ei saanud. Magistritöö eesmärk on võrrelda *The New York Times*'i tavapäraseid järelhüüdeid ja selle "Overlooked" sarja, et selgitada välja, millised on kahe erineva eesmärgiga järelhüüete sarnasused ja erinevused, mida võib kasutatav sõnavara meile öelda kollektiivse mälu tootmise kohta ja kuidas see kahe järelhüüete tüübi puhul erineb. Uurimuseks kasutatakse korpusepõhist diskursusanalüüsi. Analüüsiks kasutatakse korpustööriista Sketch Engine. Töö eesmärkide saavutamiseks on loodud kaks korpust. Üks korpus koosneb *The New York Times*'is 1. jaanuarist 2021 kuni 31. detsembrini 2021 avaldatud tavapärasest järelhüüetest (1 388 794 sõna) ja teine *The New York Times*'i "Overlooked" sarjast kuni detsembrini 2023 (kaasa arvatud) (241 153 sõna).

Sissejuhatuses käsitletakse järelhüüete žanri ajalugu ja „Overlooked" sarja ning seda, kes on järelhüüete subjektideks. Esimeses peatükis antakse ülevaade varasematest uuringutest, mis käsitlevad järelhüüete žanri, selle rolli kollektiivse mälu kujundamisel, varasemaid uuringuid järelhüüete kui diskursuse kohta ning seda, kuidas meedia on püüdnud kasutada positiivse diskrimineerimise strateegiaid, et luua kaasavamat kollektiivset mälu. Teises peatükis tutvustatakse kasutatud meetodeid, korpuspõhist diskursusanalüüsi, kasutatud korpuseid ning *New York Times*'i järelehüüetes ja selle "Overlooked" sarjas kasutatud sõnavara korpuspõhise analüüsi peamisi tulemusi.

Korpuspõhise analüüsi tulemustest joonistub välja, et "Overlooked" sarja järelhüüetes kasutatakse erinevaid meetodeid et juhtida tähelepanu muudes järelhüüetes puuduvatele subjektidele paigutada nad kollektiivsesse mällu. Samas aga kujutavad need järelhüüded vähemusi uudsusena. Kuigi seda tehakse õilsatel eesmärkidel, taastoodab see arusaama, et tavapärane kujutamisväärne inmene on valgenahaline mees. Seega ei suuda need järelhüüded laiemat ajaloolist mälu muuta, kuid siiski laiendavad seda.

Märksõnad: inglise keel ja keeleteadus, korpusanalüüs, järelhüüded, kollektiivmälu, meediadiskursus, *The New York Times*

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina, Karoliine Laaneoja,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose

   Production of Collective Memory in *The New York Times'* Obituaries and its *Overlooked* Series: A Corpus-Based Study**,**

   mille juhendaja on Raili Marling,

   reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.

2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commonsi litsentsiga CC BY NC ND 3.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.

3. Olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.

4. Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Karoliine Laaneoja

14.05.2024

Autorsuse kinnitus

Kinnitan, et olen koostanud käesoleva magistritöö ise ning toonud korrektselt välja teiste autorite panuse. Töö on koostatud lähtudes Tartu Ülikooli maailma keelte ja kultuuride kolledži anglistika osakonna magistritöö nõuetest ning on kooskõlas heade akadeemiliste tavadega.

Karoliine Laaneoja

14.05.2024

Appendix 1. Top 50 most frequent adjectives in the corpus of *The New York Times* obituaries

No.	Item	Frequency	Frequency Per Million
1.	first	2356	1409.22408
2.	other	2028	1213.03329
3.	more	1403	839.19413
4.	many	1308	782.37058
5.	early	1210	723.7526
6.	good	1039	621.47021
7.	own	999	597.5445
8.	young	908	543.11352
9.	new	889	531.74881
10.	American	812	485.69183
11.	last	803	480.30855
12.	high	710	424.68128
13.	former	691	413.31657
14.	several	682	407.93329
15.	late	629	376.23173
16.	old	598	357.6893
17.	few	585	349.91345
18.	next	575	343.93202
19.	same	552	330.17474
20.	great	546	326.58589
21.	public	507	303.25832
22.	political	502	300.26761
23.	most	473	282.92147
24.	big	473	282.92147
25.	small	443	264.97719
26.	second	432	258.39762
27.	white	403	241.05149
28.	long	398	238.06077
29.	social	392	234.47192
30.	only	388	232.07935
31.	major	368	220.11649
32.	large	352	210.54621
33.	local	333	199.1815
34.	national	320	191.40565
35.	different	319	190.8075
36.	chief	291	174.05951
37.	real	279	166.8818
38.	important	278	166.28366
39.	little	274	163.89109
40.	popular	273	163.29294
41.	French	271	162.09666
42.	British	269	160.90037
43.	much	253	151.33009
44.	musical	249	148.93752
45.	civil	247	147.74123
46.	medical	243	145.34866
47.	Jewish	242	144.75052
48.	military	238	142.35795
49.	black	232	138.76909
50.	international	232	138.76909

Appendix 2. Top 50 most frequent adjectives in the *Overlooked* corpus

No.	Item	Frequency	Frequency Per Million
1.	first	526	1821.86585
2.	other	318	1101.43221
3.	many	268	928.25104
4.	more	261	904.00568
5.	own	218	755.06988
6.	young	194	671.94292
7.	black	178	616.52495
8.	early	163	564.5706
9.	new	161	557.64335
10.	good	158	547.25248
11.	American	147	509.15262
12.	white	137	474.51639
13.	few	134	464.12552
14.	female	130	450.27103
15.	several	117	405.24393
16.	last	108	374.07132
17.	high	106	367.14407
18.	only	106	367.14407
19.	public	105	363.68045
20.	great	104	360.21682
21.	next	102	353.28958
22.	[number]	99	342.89871
23.	old	94	325.58059
24.	Chinese	93	322.11697
25.	small	92	318.65334
26.	social	89	308.26247
27.	british	88	304.79885
28.	most	82	284.01711
29.	German	81	280.55349
30.	national	75	259.77175
31.	same	75	259.77175
32.	late	73	252.8445
33.	political	71	245.91725
34.	long	68	235.52638
35.	local	67	232.06276
36.	little	67	232.06276
37.	large	66	228.59914
38.	second	65	225.13551
39.	former	62	214.74464
40.	important	61	211.28102
41.	medical	56	193.9629
42.	French	55	190.49928
43.	such	55	190.49928
44.	big	54	187.03566
45.	Japanese	53	183.57203
46.	popular	52	180.10841
47.	much	51	176.64479
48.	modern	51	176.64479
49.	short	51	176.64479
50.	sexual	50	173.18116

Appendix 3. Top 50 most frequent nouns in the corpus of *The New York Times* obituaries

No.	Item	Frequency	Frequency Per Million
1.	mr.	13343	7981.01734
2.	ms.	4664	2789.73731
3.	year	4250	2542.10625
4.	new	3024	1808.78337
5.	time	2593	1550.98388
6.	york	2345	1402.64451
7.	dr.	2061	1232.772
8.	people	2012	1203.46301
9.	university	1925	1151.4246
10.	school	1906	1140.05989
11.	work	1851	1107.16204
12.	father	1794	1073.06791
13.	book	1684	1007.27222
14.	world	1611	963.60781
15.	son	1608	961.81338
16.	life	1597	955.23381
17.	mother	1569	938.48581
18.	family	1551	927.71925
19.	child	1463	875.08269
20.	city	1417	847.56813
21.	interview	1411	843.97928
22.	music	1403	839.19413
23.	home	1383	827.23128
24.	times	1379	824.83871
25.	woman	1373	821.24985
26.	art	1335	798.52043
27.	daughter	1310	783.56687
28.	death	1271	760.2393
29.	company	1194	714.18232
30.	wife	1169	699.22875
31.	day	1147	686.06962
32.	president	1134	678.29376
33.	career	1127	674.10676
34.	show	1100	657.95691
35.	film	1092	653.17177
36.	man	1058	632.83492
37.	way	1055	631.04049
38.	part	1044	624.46092
39.	professor	992	593.35751
40.	group	989	591.56308
41.	war	972	581.39465
42.	cause	959	573.6188
43.	black	945	565.2448
44.	college	939	561.65595
45.	director	891	532.9451
46.	state	891	532.9451
47.	role	884	528.7581
48.	country	877	524.5711
49.	brother	828	495.26211
50.	friend	827	494.66397

Appendix 4. Top 50 most frequent nouns in the *Overlooked* corpus

No.	Item	Frequency	Frequency Per Million
1.	woman	1102	3816.91287
2.	year	669	2317.16399
3.	time	556	1925.77455
4.	life	415	1437.40367
5.	new	407	1409.69468
6.	people	350	1212.26815
7.	world	344	1191.48641
8.	child	339	1174.1683
9.	work	328	1136.06844
10.	school	326	1129.14119
11.	man	325	1125.67757
12.	family	323	1118.75032
13.	book	320	1108.35945
14.	mother	287	994.05989
15.	york	282	976.74177
16.	war	268	928.25104
17.	art	258	893.61481
18.	interview	245	848.58771
19.	father	241	834.73321
20.	day	239	827.80597
21.	home	219	758.5335
22.	university	215	744.67901
23.	united	214	741.21539
24.	death	204	706.57915
25.	black	195	675.40654
26.	city	191	661.55205
27.	way	191	661.55205
28.	film	190	658.08843
29.	story	190	658.08843
30.	states	188	651.16118
31.	husband	181	626.91582
32.	name	179	619.98857
33.	friend	173	599.20683
34.	history	172	595.74321
35.	right	172	595.74321
36.	part	171	592.27958
37.	artist	165	571.49784
38.	group	156	540.32523
39.	country	152	526.47074
40.	house	150	519.54349
41.	girl	150	519.54349
42.	newspaper	149	516.07987
43.	movement	144	498.76175
44.	college	135	467.58915
45.	daughter	131	453.73465
46.	student	130	450.27103
47.	state	128	443.34378
48.	decade	125	432.95291
49.	career	122	422.56204
50.	son	122	422.56204