NoSo 2023

Nordisk tidskrift för socioonomastik
Nordic Journal of Socio-Onomastics

Utgiven av Kungl. Gustav Adolfs Akademien för svensk folkkultur

3·2023

http://gustavadolfsakademien.se
noso@gustavadolfsakademien.se

NoSo 2023

Nordisk tidskrift för socioonomastik
Nordic Journal of Socio-Onomastics

Utgiven av Kungl. Gustav Adolfs Akademien
för svensk folkkultur

redigerad av / edited by
Emilia Aldrin, Linnea Gustafsson

Insända artiklar granskas i anonymiserad form av två oberoende och anonyma vetenskapliga experter.

För översättning resp. granskning av engelskspråkiga texter svarar Rhian Coggin (aukt. translator), Ystad.

NoSo har sedan årgång 1 (2021) utgivits av Kungl. Gustav Adolfs Akademien för svensk folkkultur i samarbete med Högskolan i Halmstad.

——

Submissions are reviewed by two independent and anonymous reviewers using a double-blind process.

English translation and language editing by Rhian Coggin, BA (Cantab.), Ystad.

NoSo has since volume 1 (2021) been published by the Royal Gustavus Adolphus Academy for Swedish Folk Culture in association with Halmstad University.

Distribution
eddy.se ab
e-post: order@bokorder.se
Box 1310, 621 24 Visby
Telefon 0498-25 39 00
http://kgaa.bokorder.se

Tidskriftens webbplats/The Journal's website
http://gustavadolfsakademien.se/tidskrifter/tidskrift/nordisk-tidskrift-for-socioonomastik-nordic-journal-of-socio-onomastics

Klostergatan 2
SE-75321 UPPSALA

ISSN 2004-0881

Produktion/Production: eddy.se ab, Visby 2023
Tryck/Printed by: eddy.se ab, Visby 2023

Innehåll / Content

Nordisk tidskrift för socioonomastik / Nordic Journal of Socio-Onomastics (5–12)

Introduction: Social representation and naming in sociocultural discourses

Terhi Ainiala & Minna Nevala

1. Introduction to the special issue

This issue is a collection of articles that share two main characteristics: they approach language use from the perspective of naming practice and group identity, and they focus on the study of naming and social representation that is constructed or represented in various sociocultural discourses. The joint theme for the five studies in this special issue was originally designed for a workshop at the XLVII Annual Conference on Linguistics in Tampere, Finland held in March 2021. The workshops addressed three partly overlapping areas: (a) naming patterns in group identity construction, (b) names as markers of social representation, and (c) naming as a tool for sociocultural inclusion and exclusion. These three themes are also addressed in the articles of the current issue.

The main focus of this issue is to study how group identity and public sociocultural representation are constructed through naming, and how naming separates and unites different groups. We address

Terhi Ainiala (University of Helsinki) & Minna Nevala (University of Helsinki).

NoSo 2023 | https://doi.org/10.59589/noso.32023.14398

certain questions through separate studies: What names and terms are used when we want to show that we belong to the in-group, and what means do we have to employ to show we belong to the out-group? Are the rules of naming always the same, or can negative become positive in some contexts? What happens when someone refuses to belong to a group and, at the same time, refuses to use the group's naming practices, as has happened with e.g. the colloquial names used of people living in Helsinki (Ainiala & Lappalainen 2017)? In summary, the issue explores the variety of linguistic and cultural resources that are applied in naming in various communication cultures.

This issue thus approaches the importance of naming as an identity marker in both synchronic and diachronic data and investigates their contribution to the construction of social representation. The data studied draw on a variety of text types that implement various sociocultural discourses, including newspapers, parliamentary debates, online communication, and literature. The methodologies used include corpus-assisted discourse analysis, categorisation analysis, lexical semantic analysis, and statistical analysis.

2. Social representation

The concept of social representation derives from social psychology. It is based on (cultural) values, ideas, metaphors, beliefs, and norms shared by various groups (cf. Moscovici 1984). Representation of a particular individual or a group can manifest itself by the use of 'positive' adjectives such as *healthy*, *normal*, or *natural*. Or, by terms denoting more 'undesirable' qualities, like *unnatural*, *deviant*, or *fat* (Limatius, *this issue*). This is what van Leeuwen (2008:109) calls 'moral legitimization'; it consists of the processes of evaluation, abstraction, and comparison. People are categorised into different sociocultural groups on the basis of positive and negative values. Hence, defining who we are requires a point of comparison. When we construct our social representation, or that of others, we tend to use what is called the 'in-group' and 'out-group' discourse, which is built on the notions of exclusion and inclusion (cf. Wodak 2008). When we say that someone is *a lunatic*, we simultaneously mean that we, as 'sane' people, do not belong in the same group as them (Nevala & Tyrkkö, *this issue*). Or, we may place ourselves morally and intellectually higher than those we verbally abuse by calling them by the appellativised names *uuno* or *tauno*, both with derogatory meaning (Sarhemaa, *this issue*).

Van Dijk (2009:52) discusses the state of polarised discourses in terms of 'our own place'. These are places where we want to be socially, politically, and culturally autonomous – we do not want interference from above, or from outside of our own place, i.e. group, including invasion into our way of using language (i.e. sociolect, idiolect, dialect, language). As Van Dijk (2009:141) states, giving attributes to the self and others concerns the interactional and societal context. This means that defining is not only governed by macro-level norms or shared knowledge, but it is also produced in micro-level interactions and situations, i.e. in sociocultural discourses, for example, when setting urban sociospatial categories (Ainiala et al., *this issue*). Social identities and representations can be seen to evolve and vary in social interaction, not only in response to the acts and stances of other interlocutors, but also according to the speaker's own attitude towards each interactional situation (Ochs 1993:298).

In the social identity process, we tend to exaggerate the differences between groups, as well as the similarities within the same group. Consequently, it is not unusual for groups to define their identity by their common opposition to some enemy or 'out-group'. While this process can be very effective in strengthening the in-group, it does so by significantly intensifying the intergroup conflict. Intragroup consensus can be reached by conforming to group norms. This process is called 'referent informational influence' (Hogg & Abrams 1988:172), which occurs in three stages, self-categorisation (a person defines themselves with a social category or identity), norm formation (a person creates or learns the stereotypical norms for the social category), and norm representation (a person assigns the norms to themself and starts behaving accordingly). In other words, we are influenced by others to the extent that they are in a position to be knowledgeable about group beliefs, norms, and values. That is particularly true of individuals who are most typical (prototypical) of the in-group – they guide and lead discussions about 'who we are' and consequently 'what we should do'. We might also develop an idealised picture of certain, publicly well-known, members of the (desired) in-group, in so far as being influenced by them enough to make us change our naming practices (Kanner & Raunamaa, *this issue*).

In this special issue, social representation manifests itself, for one, in actual labels and attributes, such as proper names and terms of reference. In addition, the studies show a variety of other discursive ways in which the interlocutors' social representations are expressed. The juxtaposition of different groups and their members can appear, for example, through the act of criticism, or by defining otherness. Another important factor taken up by the studies concerns sociocultural context. Naming proves to be one of the central tools for creating and maintaining social representation, whether within modern blog writing or historical media, and literature.

3. Naming

What is a name? In brief, a name is a word or combination of words, referring to one identified person, being, subject, or object, in which case the term proper name or proper noun can be used. Onomastics has seen extensive discussions on the definition and meaning of a proper name. Besides identification, names have many functions in society and culture. A person gives a name only to the referents which they feel are worth naming (see e.g. Ainiala et al. 2016:13–16; Nyström 2016). Furthermore, by examining which referents are given names and which are left nameless, we often obtain information about the objects a society deems important and valuable.

Additionally, names often convey sociocultural information, such as the social, linguistic, and ethnic identity of people and places. Proper names are not isolated elements in any language but represent a system with different sub-systems, such as the first name system and family name systems among personal names. These systems are culture-specific. In all cultures, giving a child a name means that they have been accepted as a member of the community. Richard Alford (1988) demonstrated that personal names express the identity of a person in two ways: for one, they tell the other members of the community who the individual in question is, and secondly, who they are expected to be. Thus, personal names have a significant role in building a person's individual and social identity, and constitute links between generations and families. Accordingly, when choosing a name for a child, name-givers may wish to connect the properties of a notable person to the child (Kanner & Raunamaa, *this issue*). Likewise, place names describe the cultural characteristics of objects and especially compared to other objects in a society and landscape.

A sociocultural perspective is fundamental in many other aspects. As a name is a word in a language that only has one referent, many different images, emotions, stances, and perceptions associated with this one special referent are attached to it. This becomes apparent in the use of slang names *Hesa* and *Stadi* for Helsinki and the demonyms derived from them (Ainiala et al., *this issue*). Geographically *Hesa* and *Stadi* identify the same city, but the images and affects connected to these names differ significantly.

Even though proper names are monoreferential, additional linguistic resources are available that can identify a referent, such as pronouns and classifying expressions (see e.g. Ainiala & Olsson 2020; Nyström 2016; Limatius, *this issue*; Nevala & Tyrkkö, *this issue*). Furthermore, the line between names and appellatives is not always clear-cut. Names associated with common images can also be used as appellatives (Sarhemaa, *this issue*).

In this issue, questions related to emotional and affective meanings of names and other identifying resources are discussed in multiple ways. By differentiating between various groups of people these meanings are utilised and even strengthened. All articles in this issue focus on the various ways to identify groups of people or individual people. Simultaneously, this issue highlights the vague and changeable line between proper names and other referential linguistic resources.

4. Issue contributions

The five contributions to this special issue offer a wide perspective on the topic of naming and representation, spanning the historical to present-day and both Finnish- and English-speaking contexts. The issue opens with Minna Nevala and Jukka Tyrkkö's study on the lexis of mental health in the British parliament and contemporary printed books from the early nineteenth century to the present day. The use of words referring to people with mental illness often reflects societal attitudes either in favour of or against particular group memberships. The semantic change that happened around the Second World War meant that terms like *lunatic* and *idiot* were not used to refer to people with diagnosed medical conditions, but rather as intensifiers or as distancing devices in intergroup relations.

In their article, Antti Kanner and Jaakko Raunamaa discuss the influence of media and literature on naming conventions in Finland between 1900 and 1939. Their statistical analysis of newspaper data shows that, contrary to common belief, people did not necessarily name their children after notable people. There are, however, some exceptions to how socially representative some first names were: the previously almost non-existent name *Ansa* only gained popularity in 1935–36 when Ansa Ikonen became one of the most famous actors in Finland.

Exploring present-day data from online discussion forums of Suomi24, Terhi Ainiala, Jarmo H. Jantunen, Salla Jokela and Jenny Tarvainen focus on the expressions of otherness. They perform a corpus-assisted discourse analysis by comparing the discursive naming practices for native and non-native Helsinkians, and those living in the capital region in contrast to those living elsewhere in Finland. Their results show that the capital city and its inhabitants have a distinctive role in people's socio-spatial thinking: for example, terms like *Hesa* and *Stadi* are used to distinguish 'us' (people living outside of the capital region) from 'them' (people living in Helsinki).

The practices of naming and describing bodies in plus-size women's fashion blogs are the subject in Hanna Limatius' article. Her analysis focuses on how bloggers use language to construct their identities

 NoSo 2023 | https://doi.org/10.59589/noso.32023.14398

through references to bodily characteristics and how they use terms to describe bodies that are not plus-size. The results demonstrate that, while the bloggers build counter-discourse to the mainstream media discourses presenting fatness as a negative characteristic, they also maintain particular hegemonic discourses on beauty, sexuality, and gender.

In the last article of this special issue, Maria Sarhemaa studies how proper names are used as terms of abuse in Suomi24 and internet survey data. Her specific focus is on the terms *uuno*, *tauno* and *urpo*, originally Finnish male names, which are commonly used pejoratively to express negative emotions. Her results show that these three appellativised names are not synonymous, rather they have different discursive functions: *uuno* is used to ease tensions and for self-disclosure, *tauno* to refer to 'ordinary' stupidity, and *urpo* to describe intentional negative behaviour.

References

Ainiala, Terhi & Lappalainen, Hanna. 2017. Orienting to norms: Variability in the use of names for Helsinki. In Ainiala, Terhi & Östman, Jan-Ola (eds.), *Socio-onomastics. The pragmatics of names*, 129–153. Amsterdam/Philadelphia: John Benjamins.

Ainiala, Terhi & Olsson, Pia. 2020. Places of power: Naming of affective places. *Nordisk tidskrift för socioonomastik / Nordic Journal of Socio-Onomastics* 1. 9–32.

Ainiala, Terhi, Saarelma, Minna & Sjöblom, Paula. 2016. *Names in focus: An introduction to Finnish onomastics*. Helsinki: Finnish Literature Society.

Alford, Richard D. 1988. *Naming and identity. A cross-cultural study of personal naming practices*. New Haven: HRAF Press.

Hogg, Michael A. & Abrams, Dominic. 1988. *Social identifications*. London: Routledge.

Moscovici, Serge. 1984. The phenomenon of social representations. In Farr, Robert M. & Moscovici, Serge (eds.), *Social representations*, 3–69. Cambridge: Cambridge University Press.

Nyström, Staffan. 2016. Names and meaning. In Hough, Carole (ed.), *The Oxford handbook of names and naming*, 39–51. Oxford: Oxford University Press.

Ochs, Elinor. 1993. Constructing social identity: A language socialization perspective. *Research on Language and Social Interaction* 26. 287–306.

van Dijk, Teun A. 2009. *Society and discourse: How social contexts influence text and talk*. Cambridge: Cambridge University Press.

van Leeuwen, Theo. 2008. *Discourse and practice: New tools for critical discourse analysis*. Oxford: Oxford University Press.

Wodak, Ruth. 2008. 'Us' and 'them': Inclusion/exclusion – discrimination via discourse. In Delanty, Gerard, Wodak, Ruth & Jones, Paul (eds.), *Migration, identity, and belonging*, 54–78. Liverpool: University of Liverpool Press.

Nordisk tidskrift för socioonomastik / Nordic Journal of Socio-Onomastics (13–44)

Geographical imagination and urban–rural binary in online discourses related to the capital region of Finland: A corpus onomastic study of Helsinki, Vantaa and Espoo

Terhi Ainiala, Salla Jokela,
Jenny Tarvainen & Jarmo H. Jantunen

Abstract: The article focuses on digital discourses related to Helsinki, Espoo and Vantaa, the three biggest municipalities in Finland's capital region. The data consist of texts from the discussion forum of Suomi24 that was analysed to find out how forum users produce socio-spatial distinctions by categorizing some groups as 'others' thus differentiating in-groups and out-groups. The analysis used methods of comprised corpus assisted discourse studies (CADS), including collocation analysis. The results show that discourses related both to native and non-native Helsinkians and to those living in the capital region in contrast to those living elsewhere in Finland are common and the juxtapositions between various groups are repeatedly constructed.

Terhi Ainiala (University of Helsinki), Salla Jokela (Tampere University), Jenny Tarvainen (University of Jyväskylä) & Jarmo H. Jantunen (University of Jyväskylä). Geographical imagination and urban–rural binary in online discourses related to the capital region of Finland: A corpus onomastic study of Helsinki, Vantaa and Espoo.

NoSo 2023 | https://doi.org/10.59589/noso.32023.14401

This points to the distinctive role of the capital region and its inhabitants in people's geographical imagination. The qualitative analysis reveals the contexts in which different groups of people are discussed. The findings shed light on some of the discursive practices through which the special role of Helsinki and the capital region in Finland are negotiated and perpetuated.

Keywords: corpus onomastics, digital discourse, capital region, spatial category, corpus linguistics, keyword, collocation, geographical imagination

1. Introduction

Intersubjective understandings of space and place are important to the construction of a shared social reality and coordinated action in the world. In research, this idea has been captured through the concept of 'geographical imagination', which refers to people's perceptions of space and their ways of acting in the world based on their spatial knowledge. Since the publication of Edward Said's (1979) seminal book *Orientalism*, many researchers have studied the ways in which geographical imaginations are produced through discourse on several interrelated sites, including news media, popular culture, online media, and political discussions (Eriksson 2010; Jansson 2003; Johnson & Coleman 2012). These studies show that emerging understandings of space frequently draw on conceptualizations of 'us' and 'them', and other related dualisms, reproducing 'moral geographies' where the virtues and vices of a community are assigned to specific places and regions (Johnson & Coleman 2012).

In this paper, we focus on one of the most dominant and persistent binaries associated with geographical imagination, namely that of the 'city' (urban) versus 'countryside' (rural) (cf. Eriksson 2010; Short 2005). Our interest in this topic stems from data-driven research on digital discourses surrounding the three biggest municipalities in Finland's capital region: Helsinki, Espoo and Vantaa. When studying how these cities have been discussed on Finland's biggest discussion forum Suomi24, we found that the urban–rural binary is prevalent in discussions related to all three cities – especially discussions related to the capital city Helsinki. This confirms earlier findings about the special role of the capital city (and capital city region) in the spatial order of Finland (Jokela & Linkola 2013). At the same time, it also calls for closer investigation of the discursive practices through which ideas of the city and the countryside are negotiated in the context of the recent growth and restructuring of the capital region, which has unfolded as part of a wider transformation of state space in Finland (Ahlqvist & Moisio 2014).

Recently, the populations of many small towns and rural municipalities in Finland have been aging and shrinking. While migration from small Finnish towns and rural areas to major cities has slowed down, many political actors view urbanization as unavoidable and even necessary, associating it with modernization, positive development, and competitiveness in contrast with more negative views of sparsely populated rural areas (Soininvaara 2022). In this context, Finland's capital region has gained renewed political and economic importance as a unit of competition of international investments and talent (Ahlqvist & Moisio 2014) and gone through major changes in terms of housing production, population size, and expanding infrastructure (City of Helsinki 2021). These processes are part of wider trends in Europe, which include the increase in regional inequalities, as well as the growing importance of suburban areas as concentrations of population, innovation, and economic activities (Keil 2017; Moore-Cherry & Tomaney 2019).

With this context in mind, we ask: How is the urban–rural divide negotiated and spatialized in a time when urbanization is viewed as a 'megatrend', leading to the intensification and expansion of urban infrastructure and movement of people to city regions that allegedly offer more opportunities than rural areas (see Rodríquez-Pose 2017; Soininvaara 2022)? What new knowledge can we obtain about the ideas of the 'city' and the 'countryside', if we direct our attention beyond the city centre (Helsinki) to more suburban areas (Espoo and Vantaa)?

As Barbara Johnstone's (2010) study on the discursive production of imagined dialects in the UK shows, new discursive practices and understandings of locality arise in the wake of social mobility and spatial changes. We follow this idea to study how the recent process of urbanization is reflected in digital discourses, where the relationship between Finland's capital region and more rural areas, as well as imagined boundaries between the city and the countryside are being (re)defined. Our data come from the Suomi24 corpus, which is a comprehensive collection of texts from threads on the Suomi24 discussion forum. Our four-billion-word corpus includes discussions from 2001–2017. We analyse this data to find out how the users of Suomi24 produce socio-spatial distinctions by using categories associated with the urban–rural division, and how these categories draw on and contribute to a particular geographical imagination. In doing so, we shed light on relationships and roles of the three cities in negotiations, through which people make sense of and gain ownership of the recent developments.

Theoretically and methodologically, we bring together socio-onomastics, human geography, corpus linguistics and discourse analysis. We adopt quantitative methods of corpus linguistics and combine them with qualitative methods of discourse analysis. This combination of methods leads to methodological synergy which is known as corpus-assisted discourse studies (CADS, see e.g. Partington et al. 2013). Statistical collocation analysis is the exact corpus methodological tool used in this article (cf. Chapter 3.2). Together, the quantitative collocation analysis and qualitative close reading analyses reveal which groups of people are discussed and in which contexts. This approach directs our attention to what Määttä et al. (2021:774) refer to as the 'micro-level linguistic workings of everyday discourse'. In other words, we view discourse as a site through which the boundaries and content of socio-spatial categories, such as the cities of the capital region or the capital region and countryside, are constantly produced and redefined. In doing so, we acknowledge that discourses are embedded in and constitutive of power relations and social identities. They may naturalize some ideas of space and social groups and create divisions between 'us' and 'them'. In the era of rapid urbanization, such knowledge is important, because it enables understanding

 NoSo 2023 | https://doi.org/10.59589/noso.32023.14401

the spatial dynamics of meaning-making in cities, as well as the ways in which urban discourses are entangled with the social processes that are transforming cities. Spatial categories also play an important role in constructing people's identities and life stories.

We begin by introducing our theoretical background before presenting the data and methodology used in the article. We then present and discuss the results of both the quantitative and qualitative analysis, and finally, we present the conclusions.

2. Theoretical background

The fascination of language is based on its ability to create order in a world that is inherently messy and complex. Language enables people to think in terms of categories and binary oppositions, to organize their impressions into coherent systems of knowledge, and to orient themselves in their daily lives. In social settings, language is organized into discourses, which are shared and historically contingent ways of making sense of the world. As the work of Michel Foucault (1980) shows, discourses are imbued with power, because they create taken-for-granted assumptions and position certain knowledge as being more relevant than other. Thus, not only do discourses reflect people's understanding of the world, but they also have ramifications for how the world is perceived and organized.

Discourses are integral to the emergence of geographical imaginations – that is, shared ways of thinking of the world as socially and spatially distinctive entities, which exist in people's mental maps and affect the way they act in the world (Said 1979). These conceptualizations have tangible consequences for how different geographical areas and groups of people are viewed and treated in practice.

The idea of geographical imagination has been applied to other contexts, including discursive processes and spatial differentiation within countries and across regions (e.g. Eriksson 2010; Jansson 2003; Johnson & Coleman 2012). Regardless of the scale of observation, geographical imagination often draws on and contributes to cultural dualities, which are mapped onto space in and through popular representations. These representations are entangled in complex ways and connected to similar representations from other geographical contexts (Eriksson 2010). One of the dualisms that manifests itself in different geographical contexts is that of city versus countryside, which is the focus of this paper. As Madeleine Eriksson (2010) shows in her study of Sweden, this distinction is often associated with other attributes (such as feminine/masculine or modern/traditional) and spatialized in different ways depending on the scale of observation. Such essentialized notions of the city and the countryside are not innocent, as they may conceal structural reasons for uneven development and justify policies that aim to move people to urban areas with more opportunities (see Eriksson 2010:101; Rodríguez-Pose 2017).

The distinction between the urban and rural has also been integral to geographical representations and discourses of Finland. Historically, the south-western part of Finland has been represented and conceptualized as the urbanized ‘core’ of Finland as opposed to the more rural and peripheral areas (Jokela & Linkola 2013). These representations and discourses have been primarily produced and controlled by influential individuals and networks organized around the national elite residing in the capital city. Because of this, Helsinki has come to symbolize a ‘superior’ place as opposed to the ‘other’ that has been associated with the provinces.

These kinds of distinctions often become intertwined with the emergence of spatially embedded social identities. For instance, people who live in the capital region may view themselves and their place in the world differently from those who live in the provinces, and geographical categories like the ‘city’ and the ‘countryside’ may come to play important roles in people’s life stories (Debarbieux & Petite 2014). Furthermore, there are often experienced distinctions even within a city, and non-native people may be characterized as ‘peasants’ instead of ‘authentic’ and native city dwellers (e.g. Ainiala & Lappalainen 2017).

Books and newspapers have traditionally been important to the construction of socio-spatial identities, not least because they have enabled the emergence of what Benedict Anderson (1991) terms

 NoSo 2023 | https://doi.org/10.59589/noso.32023.14401

'imagined communities'. These are communities of people, who may not know each other personally but who still share a feeling of togetherness. More recently, the role of online discussions and social media in the creation of socio-spatial identities has increased (Kavoura 2014). There has been many discussions on the formation of new social 'bubbles' through algorithms used by social media platforms, but online discussions also shed light on how traditional socio-spatial categories are conceptualized, challenged and (re)produced in everyday discourses. Furthermore, due to the abundance of entries around different topics, online discussions offer rich data for a fine-grained analysis of the spatialization of difference at various geographical scales.

In this article, we take the city names as our starting point. Generally, place names can have a central role in the construction of socio-spatial identities. For example, the recurrence of specific names in media and everyday representations reinforces the idea of the existence of these places and regions and consequently, their roles as part of a larger spatial order. Often, maps or pictures are attached to these names, linking them to a certain socio-spatial reality. Since place names are capable of indexing locality, names often shape people's senses of place and the social identities associated with place (see e.g. Ainiala 2020; Johnstone 2010).

In onomastics, our study is closely connected both to socio-onomastics and to corpus onomastics. In the field of socio-onomastics, questions related to names' roles in the construction of social identities as well as stances and attitudes towards certain names and their referents (see e.g. Ainiala & Östman 2017) are relevant to us. Self-evidently, we take the cultural and social context into account in our close analysis of names and their use. Corpus onomastics, in turn, is a relatively new approach in onomastics. It can be defined as a field in onomastic research, where electronic databases, i.e. corpora, are utilized as data, where analysis is based on corpus research methods (such as wordlist analysis, concordance and collocation analysis) and where the subject of the research is the prevalence of names, their usage in textual contexts (e.g. in collocations and other phraseological relations, and genres), as well as regional and local variation (Jantunen et al. 2022).

In sum, this multidisciplinary approach in our onomastic study will shed light on the meanings attached to the toponyms and their referents. As our special focus lies on the ways in which different groups in the capital region are viewed, compared, and discussed with each other and in relation to the rest of the country, various layers of identities and group-makings will be examined.

3. Data and methodology

3.1 Data: Suomi24 Corpus

Our data come from the 4-billion-word Suomi24 Corpus (City Digital Group 2021), which consists of posts published on the Suomi24 ('Finland24') discussion forum. The discussion forum is part of the social media website Suomi24, which comprises services such as chat rooms, blogs and dating service. Suomi24 is the most visited Finnish discussion forum with more than 3 million monthly users (in 2021), and it is one of the most visited Finnish social networking services (Kohvakka & Saarenmaa 2021). The discussion forum is divided into several sub-forums, such as health, hobbies, traffic, travelling and relationships. While the Suomi24 includes material from a vast number of users and is an all-round forum, we must bear in mind that the corpus data is not wholly representative of Finland's population in general. According to Ruckenstein (2017), for example, two-thirds of forum users are men and there is a possible over-representation of certain demographics. The Suomi24 corpus is available to researchers through The Language Bank of Finland. Our research corpus dates from the period 2001–2017. It consists of a total of approximately 10 million words divided into three sub-corpora as follows: 5.1 million words for the Helsinki corpus, 2.8 million words for the Vantaa corpus and 1.9 million words for the Espoo corpus. For the keyword analysis, the reference corpus was compiled using systematic sampling in which postings in the discussion forum were extracted at one minute past each hour around the clock. This method enabled us to avoid thematic and temporal bias in the reference data.

In this article, the aim and research questions have been formulated so that no personal information of the Suomi24 users is needed to conduct the analysis. On the Suomi24 discussion forum, writers use pseudonyms instead of personal information, such as real names. Thus, it is practically impossible to determine which user has produced a certain comment, although with some effort it is technically possible to find individual comments using search engines. In this article, the forum extracts are provided without the pseudonyms of the writers. On the user guide page for Suomi24, users are informed that their posts may be used for research purposes. However, it may of course be the case that users do not always understand that the posts may be used as research data. Thus, researchers must remember to respect the individual poster's culture and rights when analysing such data (Lagus et al. 2016).

3.2 Methods: Keywords and corpus-assisted discourse study

The data were analysed using a corpus-assisted discourse study (CADS), which combines qualitative discourse analysis with quantitative methods from corpus linguistics (cf. Partington et al. 2013:10–14). In the first phase, the data were analysed using a statistical keyword analysis. Keywords are words that are statistically more frequent in the research data in comparison with the reference data (Scott & Tribble 2006:58–59). For the calculation of keywords in Jantunen et al. (2022), we used the Keyword List program in the AntConc corpus toolkit (Anthony 2017), the Log-Likelihood test as the statistical measure (Rayson & Garside 2000) and a threshold 50 for occurrences to eliminate noise which may result from the repetition of identical postings. After this, the data were explored using qualitative close reading (cf. Mautner 2009); the results of this phase were reported in Jantunen et al. (2022). In their study, the 300 most significant keywords were grouped into discourse prosodies, i.e. associations of 'word, phrase or lemma and a set of related words that suggest a discourse' (Baker

2006:87). The discourse prosodies help us to shed light on the prevalent discourses that surround each city in the discussion. 'People' was one of the discourse prosody classes in Jantunen et al. (2022), and for the purpose of the present study, the keywords belonging to this discourse prosody were analysed more thoroughly and divided into sub-categories according to the co-textual meaning. The following analysis will concentrate on one of these sub-categories, namely demonyms, i.e. words that identify groups of people (natives, inhabitants) related to a particular place (cf. Roberts 2017).

Although the demonym keywords are clearly associated with Helsinki data (Jantunen et al. 2022), this does not mean that they do not exist in the contexts of Espoo and Vantaa. To the present study, four demonyms were taken into account. A statistical collocation analysis was carried out to find what discursive patterns the demonyms are related to. In this phase, collocations for each demonym were counted using the AntConc Collocate program. For the collocation analysis, a span of 4 words left and right and MI-test (with a cut-off point of 1.58) as a statistical measure were chosen. The MI-test was chosen since it does not correlate with collocation frequencies and thus also emphasises collocates that are not frequent in general but may reveal relevant information on discourses. See Appendix 1 for the frequencies and MI-scores of the collocates. In the final phase of this analysis, the collocates were grouped together in order to reveal the most common ways of discussion related to the demonyms.

The demonyms under study are words *stadilainen* [Stadi dweller], *hesalainen* [Hesa dweller], *maalainen* [peasant] and *juntti* [bumpkin]. The choice for these particular words arises both from their extensive occurrence in the data and their representative characteristics. Previous studies (e.g. Ainiala & Lappalainen 2017; Paunonen 2006) have revealed the meanings and functions of these words in the construction of in and out-groups and in negotiating urban–rural division. The words *stadilainen* and *hesalainen* are of special interest in our study, even for the reason they both include a common slang name for Helsinki, i.e. *Stadi* and *Hesa*. Usually, *Hesa* is known as a variant used by non-native Helsinkians and *Stadi*, in turn, as a variant used by native Helsinkians (Ainiala & Lappalainen 2017). The slang names *Stadi* and *Hesa* have their origin in Helsinki slang, which was a unique variety, a kind of pidgin language in Finland. Helsinki slang developed at the end of the 19th century and the beginning of the 20th century among the working class with both Finnish and Swedish language backgrounds in the densely populated areas of Helsinki. Consequently, these words can create discrepancies between the imagined and experienced urban and rural within the capital city Helsinki.

According to previous studies (e.g. Paunonen 2006), the words meaning 'peasant', both *maalainen*, also used in standard Finnish, and *juntti*, used especially in Helsinki slang and in colloquial Finnish are common in contexts where Helsinkians, particularly native Helsinkians are discussed in opposition to other Finns or non-native Helsinkians. Words meaning 'peasant' or 'country people' are extremely common in Helsinki slang and the concept of 'a peasant' belongs to the centres of semantic attraction in Helsinki slang. There are more than 150 words in Helsinki slang for 'peasants' (Paunonen 2006:353). The number of words given to 'a peasant' illustrates the significance of dividing the experienced rural and urban in the everyday speech among Finnish people, and likewise the words *hesalainen* and *stadilainen* the words *maalainen* and *juntti* are used also in discussions making contrasts between people living in Helsinki.

The collocation analysis was followed by a qualitative analysis interpreting the cultural meanings associated with the selected demonyms *maalainen* [peasant], *juntti* [bumpkin], *hesalainen* [Hesa dweller], and *stadilainen* [Stadi dweller]. We looked at the context of these words to analyse the mechanisms through which they were connected to wider cultural understandings of the socio-spatial order of the Finnish capital region, as well as its role in Finland. In doing so, we used the 'saturation principle' (Jokela & Raento 2012), which meant that we went through discussion entries entailing the selected words until new entries did not significantly add new insights into the contexts in which these words were used. In this phase, we also used our own cultural knowledge as Finns to interpret the nuances of the discussion entries, acknowledging that our social position inevitably affects our ways of making meaning out of the social phenomena we observe (Evans 1988).

4. Results

4.1 Discourses related to Hesa and Hesa dwellers vs. Stadi and Stadi dwellers

Hesa is a common slang name for Helsinki, and very often it is labelled as a variant used by non-native Helsinkians, in contrast to other well-known slang name for Helsinki, namely *Stadi* (Ainiala & Lappalainen 2017; see also Section 3.2). Followingly, *hesalainen* [Hesa dweller, Helsinkian] is frequently used to refer to someone who lives in 'Hesa'. The slang demonym *stadilainen* [Stadi dweller] also refers to the people living in Helsinki, but as well as *Stadi* is used by native Helsinkians, *stadilainen* also underlines that those people are born and raised in Helsinki, not those who have moved there. Table 1 lists the most statistically significant collocates of these keywords.

Genuineness and nativity versus non-nativeness are often discussed by Suomi24 users, as seen in Table 1. The table interestingly shows that this theme is represented using different collocations: collocates such as *aito* [genuine], *syntyperäinen* [native] and *paljasjalkainen* [born and bred] are common and statistically significant collocates of both Hesa and Stadi dwellers, as well as are *polvi* [generation] and *alkuperäinen* [original]. The words denoting nativity are noticeably associated with Helsinki, which tend to indicate that native Helsinkiness and social norms governing who can be counted as Helsinkian or especially as Stadi dweller are repeatedly negotiated. Extract 1 clearly points out that a first-generation resident of Helsinki has not lived there long enough to be considered *stadilainen* – the correct demonym would be *hesalainen*, according to the writer.

1) jos olet vain yhdenpolven helsinkiläisiä, olet hesalainen
[if you are just a first-generation Helsinkian, you are a Hesa dweller].

Table 1. Statistical collocates of *hesalainen* [Hesa dweller] and *stadilainen* [Stadi dweller] in data from Helsinki, Espoo and Vantaa.

	hesalainen	stadilainen
Native vs. non-native Helsinkians	Helsinki: *paljasjalkainen* [born and bred], *alkuperäinen* [original], *syntyperäinen* [native], *aito* [genuine], *polvi* [generation] Espoo: *aito* [genuine]	Helsinki: *paljasjalkainen* [born and bred], *alkuperäinen* [original], *syntyperäinen* [native], *aito* [genuine], *polvi* [generation] *wannabe* [wannabe], *sukupolvi* [generation] Espoo: *paljasjalkainen* [born and bred], *syntyperäinen* [native], *syntyä* [to be born] Vantaa: *paljasjalkainen* [born and bred], *oikea* [real], *syntyä* [to be born]
Capital area vs. countryside	Helsinki: *maalainen* [peasant], *savolainen* [Savonian], *hesalainen* [Hesa dweller], *stadilainen* [Stadi dweller], *turkulainen* [Turku dweller], *sakki* [gang; people living in Helsinki]; *pönde* [countryside], *Hesa* Espoo: *hesalainen* [Hesa dweller], *stadilainen* [Stadi dweller] Vantaa: *kaupunki* [town; city], *Vantaa*, *Helsinki*	Helsinki: *stadilainen* [Stadi dweller], *hesalainen* [Hesa dweller], *lande* [countryside], *landelainen* [lande dweller], *juntti* [bumpkin], *heinähattu* [bumpkin]; Espoo: *hesalainen* [Hesa dweller], *Stadi*, *lande* [countryside] Vantaa: *Stadi*, *Espoo*, *Helsinki*, *Vantaa*

There is a lively debate about who is eligible to use the demonym *stadilainen* and what kind of group of people it can or should refer to on the discussion forum. Extract 2, in which *stadilainen* occurs alongside with *paljasjalkainen*, illustrates that the discussion is also often affective. This extract also mentions *ylpeä*, which is one of the statistically significant collocates of *stadilainen* (cf. Appendix 1), and which signifies that people born in Helsinki are proud of their roots (cf. Ainiala & Lappalainen 2017).

2) paljasjalkainen stadilainen, varsinkaan ylpeä stadilainen ei ikinä käyttäisi itsestään nimitystä "hesalainen"!! EI IKINÄ! NEVER! Bonjaatsä? Se olisi majesteettirikos! [a Stadi dweller born and bred, especially a proud Stadi dweller would never call themself a 'Hesa dweller'!! NEVER! NEVER! Do you get it? That would be a crime against humanity!]

However, despite how clearly defined the use of *hesalainen* and *stadilainen* seems to be in some postings, the difference in the denotations is not at all clear (see also Ainiala & Lappalainen 2017). An interesting example of this is Extract 3, in which the poster refers to himself as Hesa dweller, even though he was born in Helsinki and has family roots there.

3) Vaikka olen syntyperäinen toisen polven Hesalainen en ollut käynyt Myyrmannissa joten en tunnistanut ko. paikkaa. [Although I am a native, second generation Hesa dweller I had not been to Myyrmanni so I'm unfamiliar with the place.]

Moreover, the non-nativeness of Hesa dwellers is often discussed in connection to collocates of 'countryside' and thus, the juxtaposition between rural Finland and the capital region is explicit. In Extract 4, people moving from the countryside (collocate *lande*: a slang word for 'countryside'; see also Table 2) are considered as people who instantly regard themselves as Hesa dwellers when they have managed to rent a flat in the outskirts of Helsinki, in a neighbourhood of Vuosaari (cf. Jantunen et al. 2022:31–32). The writer expresses a somewhat condescending view towards these kinds of people through the lexical choices made (e.g. *lande*, *ängetä* [force one's way]) (see also Ainiala & Lappalainen 2017:137–141).

4) Landelta änkee väkeä Hesaan ja rupeaa pitämään itseään hesalaisina, kun ovat saaneet vuokrakämpän Vuosaaresta. [People from the countryside force their way into Hesa

and regard themselves as Hesa dwellers as soon as they have managed to sign a contract for a flat in Vuosaari].

Likewise, in Extract 5 the writer makes a confrontation between Hesa dwellers and the people from the countryside. The writer somewhat ironically, and using a laughing smiley in his comment, regards himself as *maalaisjuntti* ([country bumpkin], cf. Section 4.2) and even explicitly mentions he is from Tornio, a city in Northern Finland, Lapland.

5) Aattelin että te hesalaiset tiiätte paljon paremmin ku tällanen maalaisjuntti Torniosta :D [I thought you Hesa dwellers know much better than this kind of country bumpkin from Tornio :D]

Otherwise, as Extract 5 above shows, the demonym *hesalainen* may refer to Helsinkians only or all the people in the capital area. Often, it remains unclear which ones are included as *hesalainen*. As a common idea, other Finns may categorize all the people in the capital region as Hesa dwellers. This becomes apparent in Extract 11 (see Section 4.2), where the writer expresses how Hesa dwellers is a category which sometimes may even include people living in Espoo, however, this is a conception possible only for peasants. Simultaneously, the writer makes a strong categorization between Stadi dwellers and Hesa dwellers.

Thus, as the previous extracts illustrate, *hesalainen* is sometimes a categorization given to those who have moved to the capital region, not simply to the city of Helsinki. In Extract 6 below, the writer expresses how people who have moved to Espoo deliberately label themselves as 'genuine Hesa dwellers' (*muuttaa* [to move] is also among the significant collocates, cf. Appendix 1). Thus, even here, the collocate *aito* [genuine] is represented (see Table 1). Otherwise, the approach is very similar as in Extract 4, the capital region and the rest of Finland are confronted. It is noteworthy that the city mentioned here is Oulu, a city situated in North Ostrobothnia and presumably regarded by many living in southern Finland as a very northern part of Finland (see Pietilä et al. 2019).

6) Se on ihka aito hesalainen eli oulusta espooseen muuttanut uuskaupunkilainen. [He is absolutely a genuine Hesa dweller, that is, a new city dweller who moved from Oulu to Espoo.]

As the previous examples have shown, Hesa dwellers, the non-native Helsinkians and people who have moved to the capital region, are regarded as an out-group, as 'country people', and very often are the subject of pejorative comments. This can be seen also in Table 1, which lists several disparaging labels for non-Helsinkians and areas outside Helsinki (e.g. *landelainen, heinähattu, pönde*). Additionally, in comparison to Stadi dwellers, Hesa dwellers are the more underestimated group (cf. Extract 7).

7) ne pahimmat suunsoitajat ovatkin niitä hesalaisia ei stadilaisia. [Anyway, the worst boasters are Hesa dwellers not Stadi dwellers.]

Occasionally, Hesa dwellers are framed explicitly as people who live in Helsinki and thus differ from people living in the neighbouring cities of Espoo and Vantaa. In these discussions, *hesalainen* is usually a categorization used by a non-native Helsinkian. Excerpt 8 also shows another theme closely related to the nativeness, that is, moving to Helsinki or to the capital area (collocates *muuttaa* [to move], *asua* [to live]; cf. Appendix 1)

8) Muutettiin vantaalle reilu kuukausi sitten, Onko näin että hesalaiset on yhtä tyhmii ku vantaalaiset? [We moved to Vantaa more than a month ago. Are Hesa dwellers as stupid as Vantaa dwellers?]

Even though the collocations in every sub-corpora, Helsinki, Espoo and Vantaa, are in many respects similar, there are some differences.

 NoSo 2023 | https://doi.org/10.59589/noso.32023.14401

In discussions involving Helsinki, the content words for country people are much more common than in Espoo and Vantaa. Additionally, demonyms referring to people in some other parts of the country (*savolainen*, *turkulainen*) appear only in discussions related to Helsinki. In general, the variety of different content words is remarkably bigger in the Helsinki material. Common to all sub-corpora is, nevertheless, the striking confrontation between urban (native and 'real' Helsinkians or all the people living in the capital area) and rural areas (non-native and 'unreal' Helsinkians or all the people living outside the capital area).

4.2 Discourses related to peasants and bumpkins

We examined discussions which involved the keywords *maalainen* [peasant] and *juntti* [bumpkin]. Table 2 illustrates statistically significant collocates that denote either people living in the countryside (e.g. *landepaukku*, *savolainen*) or in city (e.g. *stadilainen*, *kaupunkilainen*), and collocates referring to urban (*Stadi*, *Espoo*) and rural (*lande*) area. In Finland, *maalainen* refers to a person living in the countryside, but also more generally to a stereotypical idea of an unsophisticated country person. *Juntti* is more explicitly a derogatory term for a simple-minded, awkward, uneducated, and often bad-mannered person (Ahjopalo 2015). English equivalents of *juntti* would be *yokel* and *bumpkin*. *Maalainen* and *juntti* are often used interchangeably, which is evident also in our data.

Table 2. Statistical collocates of *juntti* [bumpkin] and *maalainen* [peasant] in the Helsinki, Espoo, and Vantaa data.

	juntti	maalainen
Capital area vs. countryside	Helsinki: *landepaukku* [bumpkin], *maalainen* ‘peasant], *juntti* [bumpkin], *savolainen* [Savonian], *stadilainen* [Stadi dweller]; *lande* [countryside] Espoo: *Stadi, Helsinki, Espoo* Vantaa: *Helsinki*	Helsinki: *maalainen* [peasant], *kaupunkilainen* [city dweller], *juntti* [bumpkin], *stadilainen* [Stadi dweller], *hesalainen* [Hesa dweller], *helsinkiläinen* [Helsinki dweller]; StadiEspoo: *Helsinki* Vantaa: *kaupunki* [town; city]

The keywords *maalainen* and *juntti* are regularly mentioned in discussions concerning differences and similarities between the municipalities of the capital region on the one hand, and between the capital region and the rest of Finland on the other hand. This is a way of negotiating boundaries between urban sophistication and small-town backwardness, as well as constructing and reworking a multi-layered socio-spatial order associated with these categories.

In this spatial order, Helsinki clearly appears superior to the rest of Finland. This superiority is constructed and justified by statements, which combine ideas of the city and countryside with common binary oppositions, such as large–small, and good–bad. For example, the geographical features of Helsinki are contrasted with those of the provinces, as evidenced by an entry where a discussant refers to Finland’s biggest lake, Saimaa, as ‘savolaisten junttien meri’ [the sea of the Savo bumpkins] as opposed to the ‘real’ sea by Helsinki. Similarly, discussions on people’s behaviour in traffic demonstrate the alleged superiority of the Helsinki people over those from the countryside (Extracts 9 and 10) (cf. also Appendix 1 for collocates denoting traffic).

9) Helsingissä osataan ajaa, eikä autoa vie kun kuski! [In Helsinki people know how to drive, and nobody except the driver steers the car!]

10) Maalaiset tukkivat liikenteen, koska he pitävät liian pitkiä turvavälejä. [Country people cause traffic jams, because their distance between cars is too long.]

While Helsinki is clearly portrayed as the core of urban sophistication, the boundary between what counts as Helsinki or 'urban' on the one hand and *juntti* or *maalainen* on the other hand is not clear-cut. Rather, it is continuously remade and context dependent. Many discussions hint that the degree of backwardness increases as a function of distance from Helsinki. From the perspective of the provinces, the residents of Espoo and Vantaa may appear as 'Helsinki people' (cf. Extract 6), but when the geographical scale of observation changes, differences between these municipalities become more evident (Extract 11).

11) Maalaisille espoolaiset ovat hesalaisia mutta stadilaisten kanssa niillä ei ole mitään tekemistä! [For country people, the residents of Espoo are Hesa people, but they have nothing to do with residents of Stadi.]

When observed in more detail, Espoo and Vantaa varyingly appear as part of a transitional zone between the capital city and countryside or as intimately connected Helsinki. This is evident, when the content of *maalainen* or *juntti* are negotiated in relation to the physical characteristics, urban structure, and mentality of Espoo and Vantaa (Extract 12) or their surrounding municipalities, such as Tuusula (Extract 13).

12) Espoo on maalaisille, jotka haluavat asua Helsingissä. Ei yhtäkään kunnon keskustaa, pelkkiä autoteitä ja omakotitaloja. [Espoo is for country people, who want to live in Helsinki. No real centre, only roads and detached houses.]

13) Tuusulassa olis halvempaa, mutta on vähän liian maalaistyylinen ja kaukana. Haluttaisiin olla espoolaisia, vantaalaisia tai hesalaisia. [It would be cheaper in Tuusula, but that's a little too country-like and far away. We'd like to be residents of Espoo, Vantaa or Hesa].

As the former discussion entry shows, Espoo and Vantaa are often deemed suitable for country people who want to enjoy the pros of city life. In contrast, the people of Helsinki are more often portrayed as self-important people who look down on country people, pretending to be more urban than they are. These discussions emphasize how Helsinki is not as urban as it appears to be, because 'along with migration waves, Helsinki has become a city dominated by country people' and 'there is no other place in Finland that has so many country people in a small area as Helsinki.'

These discussions also show that it is easy to become classified as a *maalainen* or *juntti* and, because of this, several discussants claim that people make a conscious effort to speak in the 'Helsinki way'. For instance, calling Helsinki *Hesa* instead of *Stadi* or simply *Helsinki* is a clear giveaway that the speaker is from the countryside (cf. Ainiala & Lappalainen 2017). It is also claimed by one discussant that people living in Helsinki try to hide their dialects to avoid being classified as country people.

In short, it is evident from our data that living in Helsinki does not automatically make one a 'real' Helsinki resident. This also enables debates about who is to blame for the perceived ills of Helsinki, such as self-importance or arrogance (Extracts 14 and 15).

14) Kyllä junttikin voi asua stadissa, mutta ei se siitä stadilaista tee. [A juntti can live in Stadi, but that doesn't make them a 'Stadi'.]

15) Jos helsingissä joku on tyly tai muuten juntti, niin hän on jostain muualta. [If someone in Helsinki is arrogant or otherwise juntti, they come from somewhere else.]

In these discussions, the relationship between Helsinki and the countryside resembles that of an adult and a child, who is trying to fit in with a group of adults but still lacks some of the knowledge, manners or autonomy that are attained with age. This contrast echoes a geographical imagination embedded in popular visual representations of the spatial order of Finnish nation-state (Jokela & Linkola 2013). It is also evident in discussions about the economic relationship between the capital region and rural Finland. For example, several discussants claim that 'country people' earn money as income transfers from the growing municipalities in the capital region (Extract 16). On the other hand, this view is also challenged by some, who point to the dependency of the urban areas on the countryside, which appears as a sphere of people who are not alienated from nature or physical work (Extract 17).

16) maalaisille kärrätään rahaa kärry tolkulla kasvukuntien kirstusta [piles of money are being carted to the country people from the coffers of the growing municipalities]

17) Ettei vaan olis niinpäin, että ahkerat maalaiset elättää stadin muovikassi-miehet [Isn't it rather so that the hard-working country people support the plastic bag men ('drunkards') of stadi]

These divergent interpretations echo wider political sentiments that emphasize the value of urban areas as engines of national economies and portray rural centres in the context of crisis-speech related to problems of declining regions (Soininvaara 2022:12–43). As Eriksson (2010:101) points out, representations of peripheral areas as 'weak' and subordinate to the urban areas are inherently political, not least because they help to justify regional policies that favour urban areas by highlighting their capacity to provide economic growth for the entire nation. Eriksson (2010) examines the situation in Sweden, but similar developments are evident in Finland, where the idea of investing uniformly throughout the country has been largely replaced by political viewpoints that emphasize the development of Helsinki into an attractive and innovative ‘world city’ that connects Finland into global economic flows (Ahlqvist & Moisio 2014). In this context, the efforts of some Suomi24 discussants to show the economic value and authenticity of the countryside (Extract 17) can be interpreted as acts of self-empowerment and ways of justifying living in areas associated with the prevailing crisis-speech.

5. Discussion and conclusions

In this article, we have discussed digital discourses of geographical imagination produced in the discussions related to Helsinki, Espoo, and Vantaa, the three biggest municipalities in Finland’s capital region. Using texts from the Suomi24 discussion forums as our data, we analysed the ways in which the users produce socio-spatial distinctions by categorizing some groups as ‘others’ and thus differentiating in-groups and out-groups. In the analysis, we started with statistical collocation analysis which was complemented with close reading and an interpretation of the cultural meanings of the utterances extracted from the data. Our aim has been to combine methods in a manner that pays attention to the strengths and limitations of them. While the analysis of a large body of corpus data may lose detailed information, qualitative analysis without statistical information may strengthen the role of researcher’s intuition. By using mixed method approaches, such as keyword and collocation analyses and close-reading in the present study, the methodologies can reinforce each other (cf. Baker 2006:15–16) and produce perhaps more accurate information of the phenomenon under study.

The results show that discourses related both to native and non-native Helsinkians and to those living in the capital region in contrast to those living elsewhere in Finland recur, and the juxtapositions between the experienced ‘urban’ and ‘rural’ are repeatedly constructed. This points to the distinctive role of the capital city and its inhabitants in people’s geographical imagination (Eriksson 2010; Jokela & Linkola 2013). The qualitative analysis reveals the contexts in which the division between the urban and rural are discussed.

All the four keywords analysed qualitatively in our article, namely *hesalainen* [Hesa dweller], *stadilainen* [Stadi dweller], *maalainen* [peasant] and *juntti* [bumpkin], are used in making comparisons between people living in Helsinki or in the capital region and people living outside Helsinki region. Whereas Stadi dwellers are categorized as native Helsinkians and often presented as 'genuine' Helsinkians, Hesa dwellers are framed as non-natives and simultaneously often as country people or at least equal to them in comparisons between different Helsinkians (cf. Ainiala & Lappalainen 2017). Respectively, in discussions between the capital region and the rest of the county, Hesa and Stadi dwellers are confronted with all the other Finns, who, in turn, are often presented as country people.

Our analysis supports earlier studies that have shown the importance of binary oppositions in meaning-making and discourses, which work towards particular geographical imaginations (Eriksson 2010; Said 1979; Silva 2017). The attributes associated with people living in and outside of Helsinki confirm the special role of the capital city in the spatial order of the Finnish nation-state and contrast it with the 'rest of Finland' (Jokela & Linkola 2013). What is new in our study is that by comparing discourses associated with Helsinki, Espoo, and Vantaa, we have demonstrated the important role of suburban areas around the capital city in the negotiation of what counts as 'urban' as opposed to 'countryside' or sophisticated as opposed to 'juntti'. This points toward the need for new research that focuses on the special urban citizenship and forms that are emerging in these areas that have less clearly defined and stable positions in the geographical imagination of Finnish people (cf. Keil 2017).

Usually, the ones who label people in Helsinki (region) as Hesa dwellers are from outside of the capital region, and thus Helsinki and people living there are observed from outside. This shows that, as a capital city, Helsinki is a special geographical entity impregnated with symbolic power and charged with meaning. It is within mundane discourses, like the ones apparent on the Suomi24 discussion forum, that this meaning is mobilized for the construction of spatially embedded social identities (cf. Määttä et al. 2021). Comments about

Helsinki and Hesa dwellers serve to distinguish ‘us’ (people living outside of the capital region) from ‘them’ (people living in Helsinki) and, simultaneously, to gain ownership of some of the meanings associated with the capital city.

This is interesting in the context of the recent processes of urbanization and related political discussions, where shrinking and aging towns and rural municipalities are frequently framed as problem areas (Soininvaara 2022). The ‘othering’ of Helsinki is a discursive strategy that enables people living in small towns and rural areas to gain agency by challenging some of the prevailing ideas of the alleged virtues of the capital region. In other words, the formation of social identities around an adversarial relationship between Helsinki and the rest of Finland may serve to foster or return the self-regard of people who do not identify themselves with the values and meanings associated with the urbanized ‘core’ of Finland.

With this in mind, we suggest that studying digital discourses may help researchers and urban professionals become more aware of discourses through which people’s geographical imagination is continuously remade as a response to various ideas that underlie urban policies and regional development agendas. Identifying possible tensions and anticipating conflicts enables these professionals to direct their actions in a socially sustainable way, ensuring that no one is left behind.

References

Ahjopalo, Janne. 2015. Junttius on kiehtovan monitasoinen ilmiö – ”Se voi jopa buustautua lottovoitosta”. Yle, October 20, 2015. https://yle.fi/uutiset/3-8393303 (accessed 5 June 2022)

Ahlqvist, Toni & Moisio, Sami. 2014. Neoliberalisation in a Nordic state: From cartel polity towards a corporate polity in Finland. *New Political Economy* 19(1). 21–55.

Ainiala, Terhi. 2020. Aspects of localness: A pilot study of kiosk and grill food names in Finland. *Onomastica* 64. 167–177.

Ainiala, Terhi & Lappalainen, Hanna. 2017. Orienting to norms. In Ainiala, Terhi & Östman, Jan-Ola (eds.), *Socio-onomastics. The pragmatics of names*. (Pragmatics & Beyond New Series 275), 204–226. Amsterdam: John Benjamins.

Ainiala, Terhi & Östman, Jan-Ola. 2017. Introduction: Socio-onomastics and pragmatics. In Ainiala, Terhi & Östman, Jan-Ola (eds.), *Socio-onomastics. The pragmatics of names*. (Pragmatics & Beyond New Series 275), 1–18. Amsterdam: John Benjamins.

Anderson, Benedict. 1991. *Imagined communities* (2nd edition). London: Verso.

Anthony, Laurence. 2017. *AntConc* (Version 3.5.8) [Computer Software]. Tokyo, Japan: Waseda University. Available at: https://www.laurenceanthony.net/software2017

Baker, Paul. 2006. *Using corpora in discourse analysis*. London: Continuum.

City Digital Group. 2021. Suomi24 virkkeet -korpus 2001–2020, Korp-version [text corpus]. The Language Bank of Finland. Retrieved from http://urn.fi/urn:nbn:fi:lb-2021101525.

City of Helsinki. 2021. *Statistical Yearbook of Helsinki* 19. The Urban Research and Statistics Unit at Helsinki City Executive Office. https://www.hel.fi/hel2/tietokeskus/julkaisut/pdf/22_03_31_statistical_yearbook_of_helsinki_2021.pdf (accessed 10 September 2022)

Debarbieux, Bernard & Petite, Mathieu. 2014. Do people live in geographic categories? Cities, countryside, and mountains in biographical narratives. *Annales de géographie* 693(5). 483–501.

Eriksson, Madeleine. 2010. "People in Stockholm are smarter than countryside folks"–Reproducing urban and rural imaginaries in film and life. *Journal of Rural Studies* 26(2). 95–104.

Evans, Mel. 1988. Participant observation: The researcher as research tool. In Eyles, John & David M. Smith (eds.), *Qualitative methods in human geography*, 197–218. New York: Barnes & Noble.

Foucault, Michael 1980. *Power/knowledge: Selected interviews and other writings*. New York: Pantheon Books.

Jansson, David R. 2003. Internal orientalism in America: WJ Cash's *The Mind of the South* and the spatial construction of American national identity. *Political Geography* 22(3). 293–316.

Jantunen, Jarmo, Ainiala, Terhi, Jokela, Salla & Tarvainen, Jenny. 2022. Mapping digital discourses of the Capital Region of Finland: Combining onomastics, CADS, and GIS. *Names* 70(1). 19–39.

Johnson, Corey & Coleman, Amanda. 2012. The internal other: Exploring the dialectical relationship between regional exclusion and the construction of national identity. *Annals of the Association of American geographers* 102(4). 863–880.

Johnstone, Barbara. 2010. Indexing the local. In Coupland, Nikolas (ed.), *Handbook of language and globalization*, 386–405. Oxford: Oxford University Press.

Jokela, Salla & Linkola, Hannu. 2013. 'State idea' in the photographs of geography and tourism in Finland in the 1920s. *National Identities* 15(3). 257–275.

Jokela, Salla & Raento, Pauliina. 2012. Collecting visual materials from secondary sources. In Rakić, Tijana & Chambers, Donna (eds.), *An introduction to visual research methods in tourism*, 53–69. London: Routledge.

Kavoura, Niki. 2014. Social media, online imagined communities and communication research. *Library Review* 63(6). 490–504.

Keil, Roger. 2017. *Suburban planet: Making the world urban from the outside in*. Cambridge: Polity Press.

Kohvakka, Rauli & Saarenmaa, Kaisa. 2021. Median merkitys on kasvanut pandemian aikana – monet ikäihmiset ovat ottaneet melkoisen digiloikan. *Tieto & Trendit*. Helsinki:Tilastokeskus. https://www.stat.fi/tietotrendit/artikkelit/2021/median-merkitys-on-kasvanut-pandemian-aikana-monet-ikaihmiset-ovat-ottaneet-melkoisen-digiloikan/ (accessed 10 September 2022)

Lagus, Krista, Ruckenstein, Minna, Pantzar, Mika & Ylisiurua, Marjoriikka. 2016. *Suomi24: muodonantoa aineistolle*. Helsinki: Helsingin yliopisto.

Mautner, Gerlinde. 2009. Corpora and critical discourse analysis. In Baker, Paul (ed.), *Contemporary corpus linguistics*, 32–46. London: Continuum.

Moore-Cherry, Niamh, & Tomaney, John. 2019. Spatial planning, metropolitan governance and territorial politics in Europe: Dublin as a case of metro-phobia? *European Urban and Regional Studies* 26(4). 365–381.

Määttä, Simo K., Suomalainen, Karita & Tuomarla, Ulla. 2021. Everyday discourse as a space of citizenship: The linguistic construction of in-groups and out-groups in online discussion boards. *Citizenship Studies* 25(6). 773–790.

Partington, Alan, Duguid, Alison & Taylor, Charlotte. 2013. *Patterns and meanings in discourse: Theory and practice in corpus-assisted discourse studies (CADS*). Amsterdam: John Benjamins.

Paunonen, Heikki. 2006. Synonymia Helsingin slangissa. *Virittäjä* 110(3). 336–364.

Pietilä, Juulia, Tillotson, Jack S. & Askegaard, Søren. 2019. Mythologies of Finnishness in advertising. In Askegaard, Søren & Östberg, Jacob (eds.), *Nordic consumer culture*, 239–265. Cham: Palgrave Macmillan.

Rayson, Paul & Garside, Roger. 2000. Comparing corpora using frequency profiling. In *The workshop on comparing corpora*, 1–6. Hong Kong: Association for Computational Linguistics.

Roberts, Michael. 2017. The semantics of demonyms in English: Germans, Queenslanders, and Londoners. In Ye, Zhengdao (ed.), *The semantics of nouns*, 205–220. Oxford: Oxford University Press.

Rodríguez-Pose, Andrés. 2017. The revenge of the places that don't matter (and what to do about it). *Cambridge Journal of Regions, Economy and Society* 11(1). 189–209.

Ruckenstein, Minna. 2017. Suomi24-käyttäjäkyselyn tuloksia. Citizen Mindscapes -project. http://blogs.helsinki.fi/citizenmindsca-pes/2017/08/21/suomi24-kayttajakyselyn-tuloksia/ (accessed 5 June 2022)

Said, Edward. 1979. *Orientalism*. New York: Vintage Books.

Scott, Mike & Tribble, Chris. 2006. *Textual patterns: Key words and corpus analysis in language education*. Amsterdam: John Benjamins.

Short, John Rennie. 2005. *Imagined country: Environment, culture, and society*. Syracuse, New York: Syracuse University Press.

Silva, Derek. M. D. 2017. The othering of Muslims: Discourses of radicalization in the New York Times, 1969–2014. *Sociological Forum* 32(1). 138–161.

Soininvaara, Ilppo. 2022. *The politics of urbanization: From a global imperative to national struggles*. Department of Geosciences and Geography A98. Helsinki: University of Helsinki.

Appendix 1. The most significant content word collocates (denoting to authenticity, opposition between Helsinki and countryside, characteristics, traffic and moving and living) of the four studied node words in three data. The frequencies and statistical MI-values are presented in those data categories in which the collocates exist.

Semantic group	Node word	Collocates with frequencies and MI-values in three data		
genuine Stadi dweller	*hesalainen*	Helsinki-data	f	MI
		paljasjalkainen [born and bred]	14	7.91
		syntyperäinen [native]	8	7.55
		alkuperäinen [original]	12	6.89
		aito [original]	14	6.44
		polvi [generation]	8	6.61
		Espoo-data	f	MI
		aito [original]	3	8.91
	stadilainen	Helsinki-data	f	MI
		paljasjalkainen [born and bred]	66	9.75
		syntyperäinen [native]	42	9.55
		wannabe [wannabe]	8	8.88
		aito [original]	91	8.55
		polvi [generation]	33	8.26
		alkuperäinen [original]	37	8.13
		sukupolvi [generation]	7	6.71
		Espoo-data	f	MI
		paljasjalkainen [born and bred]	6	12.35
		syntyperäinen [native]	4	11.05
		syntyä [to be born]	4	6.46
		Vantaa-data	f	MI
		paljasjalkainen [born and bred]	9	9.94
		syntyä [to be born]	6	4.28
		oikea [real]	6	3.81

Helsinki vs countryside	*hesalainen*	Helsinki-data	f	MI
		pönde [countryside]	5	9.36
		turkulainen [Turku dweller]	11	7.36
		hesalainen [Hesa dweller]	42	6.92
		stadilainen [Stadi dweller]	41	6.49
		savolainen [Savonian]	9	6.24
		maalainen [peasant]	16	6.18
		sakki [gang]	5	6.12
		Hesa	221	4.04
		Espoo-data	f	MI
		hesalainen [Hesa dweller]	4	9.92
		stadilainen [Stadi dweller]	3	9.19
		Vantaa-data	f	MI
		kaupunki [town; city]	5	5.07
		Vantaa	8	4.03
		Helsinki	7	3.71
	stadilainen	Helsinki-data	f	MI
		heinähattu [bumpkin]	6	7.87
		landelainen [land dweller]	9	7.46
		lande [countryside]	51	6.69
		stadilainen [Stadi dweller]	60	6.65
		hesalainen [Hesa dweller]	41	6.49
		juntti [bumpkin]	35	6.44
		Espoo-data	f	MI
		hesalainen [Hesa dweller]	3	9.19
		lande [countryside]	11	8.70
		Stadi	3	7.77
		Vantaa-data	f	MI
		Stadi	10	7.41
		Espoo	14	4.62
		Helsinki	10	3.56
		Vantaa	6	2.95

	maalainen	Helsinki-data	f	MI
		maalainen [peasant]	18	5.00
		kaupunkilainen [city dweller]	9	5.57
		juntti [bumpkin]	10	3.68
		stadilainen [Stadi dweller]	13	3.18
		hesalainen [Hesa dweller]	7	3.17
		helsinkiläinen [Helsinki dweller]	8	2.82
		Stadi	92	1.68
		Espoo-data	f	MI
		Helsinki	13	12.22
		Vantaa-data	f	MI
		kaupunki [town; city]	17	2.80
	juntti	Helsinki-data	f	MI
		juntti [bumpkin]	66	7.93
		landepaukku [bumpkin]	6	7.11
		maalainen [peasant]	20	6.68
		stadilainen [Stadi dweller]	35	6.44
		lande [countryside]	26	6.29
		savolainen [Savonian]	7	6.06
		Espoo-data	f	MI
		Stadi	9	7.65
		Espoo	53	3.84
		Helsinki	9	3.78
		Vantaa-data	f	MI
		Helsinki	7	9.87
characteristic	*hesalainen*	Helsinki-data	f	MI
		leuhka [boastful]	5	9.21
		ylimielinen [arrogant]	9	7.86
		kusi [piss]	6	6.72
	stadilainen	Helsinki-data	f	MI
		dille [stupid]	5	9.92
		ylpeä [proud]	15	7.52
	maalainen	Helsinki-data	f	MI
		typerä [fool]	21	19.80

	juntti	Helsinki-data	f	MI
		idiootti [idiot]	5	5.81
traffic	*hesalainen*	Helsinki-data	f	MI
		kolaroida [to collide]	5	8.33
		Espoo-data	f	MI
		metro [underground; metro]	4	6.34
		ajaa [to drive]	4	5.15
moving and living	*hesalainen*	Helsinki-data	f	MI
		muuttaa [to move]	41	4.60
		muualta [from elsewhere]	13	4.39
	stadilainen	Espoo-data	f	MI
		muuttaa [to move]	6	5.60
		asua [to live]	16	5.19
		Vantaa-data	f	MI
		asua [to live]	10	12.68
	juntti	Helsinki-data	f	MI
		muuttaa [to move]	52	5.13
		asua [to live]	60	4.20

Nordisk tidskrift för socioonomastik / Nordic Journal of Socio-Onomastics (45–70)

Large-scale statistical analysis on representation of public figures in newspapers and naming conventions in Finland in 1900–1939

Antti Kanner & Jaakko Raunamaa

Abstract: This article examines the influence public figures had on the name-giving in Finland from 1900 to 1939 via their presence in newspapers. The study is based on digital materials and the results were obtained using statistical research methods. The main finding of the study is that generally, the prevalence of notable peoples' names does not reflect the person in the way people choose first names for their children. On the other hand, more in-depth analyses reveal consistent interaction: short-term peaks in prominence of public figures in the newspaper data coincide with peaks in the popularity of their first names 50% more often than could be expected by random. Also, high-correlation outliers reveal personalities with a definite impact on first given name trends, interpretable as rule-confirming exceptions of the general pattern of very modest correlations.

Keywords: name-giving, statistical analysis, early 20th century, history of Finland, digital humanities.

Antti Kanner (University of Helsinki) & Jaakko Raunamaa (University of Helsinki & University of Tallinn). Large-scale statistical analysis on representation of public figures in newspapers and naming conventions in Finland in 1900–1939.

NoSo 2023 | https://doi.org/10.59589/noso.32023.14404

1. Introduction

This article focuses how newspapers and public figures influenced trends for first names in Finland during the first half of the 20th century (1900–1939). It seeks to answer the following research question: Is there a statistically observable interaction between the prominence of public figures in the newspaper data and trends with first names attributable to an influence of the former to the latter?

The role of external factors (such as popular culture and literature) in name giving is a familiar and discussed topic in onomastics (e.g. Ainiala et al. 2012:178; Leibring 2016; Nyström 2016:12). There are a few examples in Finnish onomastic literature that are often used to demonstrate the influence of public figures and literature characters on name-giving. For example, it is claimed that the popularity of the female name *Armi* increased greatly after 1952 when Armi Kuusela won the Miss Universe pageant (e.g. Kiviniemi 2006:217). In similar fashion, the case of *Olavi* is often used to illustrate the impact of literature on Finnish naming conventions (e.g. Vilkuna et al. 1988:125; Kiviniemi 2006:183).

The most relevant study for this research was conducted by Eero Kiviniemi in 1982 and consisted of surveys organized in the late 1970s that focused on the principles of name giving in Finland (1982:129). The results include close to 600 answers, which, in turn, contain details of approximately 4,500 first names. Based on the outcomes, Kiviniemi created a nine-part classification model to demonstrate the different categories of naming principles and their relative proportions. According to Kiviniemi, three percent of the first names were given based on famous individuals (173). Furthermore, five percent of the cases fell into the category of ‘historical and literature characters’ (ibid).

Hence, it seems plausible to assume that external cultural and social factors such as celebrity influence name-giving. On the other hand, one must remember that this assumption is based on very limited data. The examples discussed in onomastic literature and Kiviniemi’s statistical analysis (1982) only cover a very small portion of all the names given annually in Finland. According to the records of Official

Statistics of Finland (OSF), the average annual number of new-borns at the beginning of the 20th century (1900–1939) was approximately 80,000. The conclusion is that no scholar in Finland (or to the best of our knowledge, nowhere in the world) has comprehensively analysed the influence of external factors in name giving through statistical methods.

Limited access to suitable materials and methods has made it difficult to conduct statistical research on external factors in name-giving from previous decades. Since the 2010s, various archives have been digitized and made available for digital analyses (Ogilvie 2016). It is also clear that the field of digital humanities and the methods used within it have evolved and expanded during the past decades (Brennan 2018).

In the present article, we bring digitized archive materials and robust statistical analysis methods together to investigate the interaction between the occurrences of public figures in Finnish newspapers and the naming conventions in Finland during the first part of the 20th century (1900–1939). More specifically, we have collected a list of names of famous Finns from the National Biography of Finland (NBF) and recorded how often they are mentioned in the Finnish newspapers digitized by the National Library of Finland. This is done in three parts: First, by using Pearson's correlation coefficient to investigate the overall correlation between the occurrences of public figures in the press and the popularity of their first names according to the first given name statistics of Finland.[1] Second, we analyse the predictability of the distribution across all given first names for each year, given the distribution across relative frequencies of public figures' mentions for the respective years. Third, a peak detection method has been used to detect those names and famous persons whose popularity and prominence have peaked at the same time. Based on the

[1] When discussing public figures, the expression *first name* refers to the name by which the person was widely known. It was not necessarily always the person's official first given name. It was common to have two first names in Finland at the beginning of the 20th century (Ainiala et al. 2012:174). See Section 2.2 for more details regarding the choice of names used for public figures.

results, some names are analysed at the more fine-grained level. We discuss the reasons and features that have either enabled or prevented the popularity of a first name being influenced by the prominence of a public figure.

The article comprises four sections. The first one provides an overview of Finnish history from the latter decades of the 19th century to the end of the 1930s. The second section focuses on the research materials and more specifically explains how the data is compiled and adjusted for the purposes of this work. Issues with the research material are also indicated and discussed. The analyses and their results are presented in the third section. We have selected three complementary statistical analyses for our study and each of these will be presented alongside their results. This choice of structure keeps the methodological description of each analysis in close proximity to its results. Thus, it is more reader-friendly for audiences who are not well-versed in statistical details.

Because large-scale statistical analyses have not previously been conducted on our material, we chose to apply the most intuitive and robust statistical methods for looking at the interaction between press prominence and the popularity of first names. The fourth and final part presents the conclusions together with two more detailed analyses of public figures and their influence on how names are chosen.

We also argue that while during the period of our study celebrity culture, newspaper media and naming conventions all were quite different to what they are today, the interaction of the three can be meaningfully studied as long as interpretations of the results do not rely on unfounded assumptions regarding their historical nature. For this reason, we will seek to present our results in robust statistical terms and only offer reserved interpretations of the changes in underlying cultural, social and linguistic structures involved. We see that our study can form a starting point for more specific research. Future contributions employing more elaborated and detailed statistical analyses may then benefit from the groundwork laid out here.

2. Historical background

There are many examples of major changes in naming conventions coinciding with the most important social and cultural developments in the surrounding society (Salway 1994; Raunamaa 2021:72). This line of thought is further supported by early 20th century Finland. To begin with, naming choices changed during the first part of the 1900s. From the Late Middle Ages onwards (c. 1400 CE, most of the inhabitants of Finland were named after Christian saints and their names recorded mainly in their Swedish forms (Ainiala et al. 2012:161–166; Raunamaa 2020). The system, which had remained fairly stagnant for centuries, began to break down from the end of the 19th century. In short, the popularity of Finnish names started to rise, whereas the number of names written in non-Finnish form declined rapidly. This phenomenon is depicted in Figure 1 below (based on Kiviniemi 2006:265). One can notice how the rapidly declining popularity of the first category "Names in foreign form" (e.g. *Anders*, *Jakob* and *Margareta*) coincides with rise of second and third sections that include Fennicised foreign names (= names adapted into Finnish) (e.g. *Anders* -> *Antti*, *Jakob* -> *Jaakko* and *Margareta* -> *Reetta*) and names based on pre-Christian Finnish nomenclature (e.g. *Kauko*), Finnish mythology (e.g. *Väinö*) or Finnish vocabulary (e.g. *Urho*). These changes in naming conventions concurred with turmoil in Finnish society. The events leading to the social upheavals and concurrently created a fertile ground for changes in the use of given names in Finland at the turn of the 20th century are described in more detail below.[2]

[2] This section is based on Saari (2012) and Ylikangas (2002:103–189) unless stated otherwise.

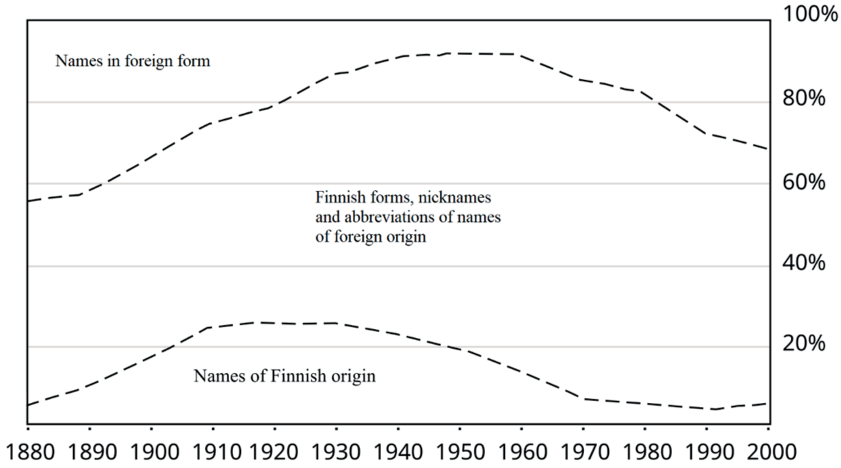

Figure 1. Proportions of different groups within the names given in Finland between 1880 and 2000 based on Kiviniemi 2006: 265. The upper section "Names in foreign form" refers to names that were used in their original foreign form and were not Fennicised. The second section "Finnish forms, nicknames and abbreviations of names of foreign origin" includes all the foreign originated names that were adapted into Finnish. The third category, "Names of Finnish origin", covers all the names that derive from pre-Christian Finnish nomenclature, Finnish mythology or were created on the basis of Finnish language from the mid-19th century onwards.

During the latter part of the 19th century, the Grand Duchy of Finland was linguistically a diglossic society: Swedish was predominantly used as the prestige language of administration, culture, politics, public sphere and education while Finnish was the household language of rural population (Marjanen et al. 2019:57). However, towards the turn of the century, Finnish gradually became more prominent in public discourse and administrative use and especially as the symbol of national identity. Crucial steps for the wide adoption of Finnish language were, for example, the establishment of the first Finnish teacher-training college in 1863 and Finnish achieving equivalent legal status to Swedish in 1902 (Saari 2012:180). Finnish education institutions also developed rapidly during the 19th century. At the beginning of the 1800s many Finns had basic literacy as the Lutheran Church required that adults could read religious texts (Kuikka 2010:2). From

 NoSo 2023 | https://doi.org/10.59589/noso.32023.14404

1850 onwards, the national government became interested in the education of ordinary people (ibid. 5–9). The result was that by the turn of the 20th century, most of the Finns were able to read.

The improving prestige of the Finnish language was influenced by the European intellectual trends of nationalism and romanticism. When these ideological trends arrived in Finland, the local cultural elite, mostly Swedish speaking at the time, started to promote a distinct national identity of Finns, drawing clear separations from the neighbouring nations, especially Russians and Swedes. Improving the status of Finnish language was seen as one of the key expressions of the national identity and formed the core of the political agenda of the so-called Fennoman political movement.

The Russian Empire showed mixed reactions to Finnish nationalistic developments. On the one hand, Finnish national identity was seen as instrumental in severing the ties Finns still had to the Swedish Realm under which Finland had belonged before the Finnish War (1808–1809). On the other hand, a distinct Finnish national identity also hindered the assimilation of Finns as loyal subjects of the Russian Empire. During the 19th century, the position of the Russian government shifted between these two emphases, the shifts resulting in legislation sometimes promoting and sometimes demoting the status of Finnish as an official language. From the perspective of this article, one of the most impactful periods started at the turn of the 20th century when Russification started with the aim to linguistically, culturally and economically absorb Finland into Russia. This period lasted until independence. One of the consequences of Russification was that Finns (both Finnish and Swedish speaking) united against Russification and the trend of nationalism was strengthened.

Besides linguistic and cultural turmoil, it is also safe to claim that Finnish society also underwent radical economic and social changes. At the beginning of the Russian era, Finland was still an agrarian society, but from the 1830s onwards industrialization took hold and the economy grew. Paper was one of the key goods produced in Finland, which, of course, contributed to the development of the local newspaper publishing.

In 1917, after many years of unrest and upheaval within the Russian Empire, the Parliament of Finland declared independence. The declaration was soon followed by political turmoil, which quickly transformed into an open conflict. The result was a civil war between the Conservative 'White' and Revolutionary Communist 'Red' armies, the former representing the forces of Finnish Socialist Workers' Republic and the latter being a coalition of various groups united by the opposition of the communist and left-wing Reds.

The Whites won the short but fierce war, after which society rapidly began to normalize, as did economic growth. This was supported in particular by the forest industry as well as the economic openness that Finland promoted after the civil war. The new nation also craved fresh national heroes which were obtained, for example, among the long-distance runners and Nordic skiers who brought many Olympic gold medals to Finland.

3. Research material

This section gives a description of the materials used in this study. We start with the newspapers, continue with public figures and finally, describe the material containing the Finnish first name statistics. We also discuss the issues related to the materials in the subsections.

3.1 Newspaper material

Newspapers were the most widespread form of media in the early 20th century Finland, making them the most viable source for investigating the influence of public figures on name-giving in Finland. Since most of the Finns were introduced to newspapers by the end of 19th century (Marjanen et al. 2019:56), the degree of literacy even in rural regions of Finland was relatively high (Suomen tilastollinen vuosikirja, 1903) and as radio did not become widespread until the 1930s (Endén et al. 1996:67), the printed press was likely the main source of information at the time. However, it must be taken into consideration that, although newspapers were widely published during the early 20th century, regional differences were still significant. For example, in 1920 many rural areas were hundreds of kilometres away from the closest newspaper publisher and therefore, it is likely that the inhabitants of these remote regions did not encounter newspapers regularly (Marjanen et al. 2019:71).

Second, the historical newspapers in Finland are readily available for research. Many of the early 20th century Finnish newspapers have been digitized by the National Library of Finland and their accessibility varies across several platforms. The data we have used come from the annotated corpus version maintained by the Language Bank of Finland (www.kielipankki.fi). It was extracted using the API interface of the Korp corpus tool by which Kielipankki provides access to their resources (Borin et al. 2012). We only included Finnish language newspapers published between 1900–1939. From these newspapers, we collected yearly relative frequency counts for all sequences matching the full names in the Finnish National Biography.

Due to licensing issues, the Kielipankki version of the historical newspaper collection only has all papers until 1910. After that, the selection of newspapers is much smaller: in 1910 the number of tokens is close to 261 million, whereas the 1920 data contain approximately 20 million tokens. Although using the relative frequencies mitigates the immediate issues related to uneven sizes of the yearly corpora, the problems related to unpredictable sampling of the later years remain. The fact of the matter is that we cannot know whether the latter two thirds of our data are systematically biased compared to the first third. This can only be validated in the future once the rest of the material becomes available for research. For now, we simply assume that the variation in the inclusion of newspapers in the corpus (i.e. exclusion of most of the newspapers from the later yearly corpora) does not systematically affect the prominence of certain public figures and present our results as conditional to that assumption.

A summary inspection of some public figures and the relative frequencies in the newspaper corpus does not seem to depend on the differences in the composition of the corpus, or at least to a great extent.

Below (Figure 2) are relative frequency graphs for two famous Finns, long distance runner Paavo Nurmi and composer Jean Sibelius, both widely known public Finnish figures of their time. Both graphs show variation in frequency congruent with their respective biographies, but not at least obviously attributable to the major segmental variation in the data. More importantly, the scale of the relative frequencies also seems to be well comparable across the years, with these two figures reaching peak prominence in around one to two occurrences per million tokens of running text across the period.

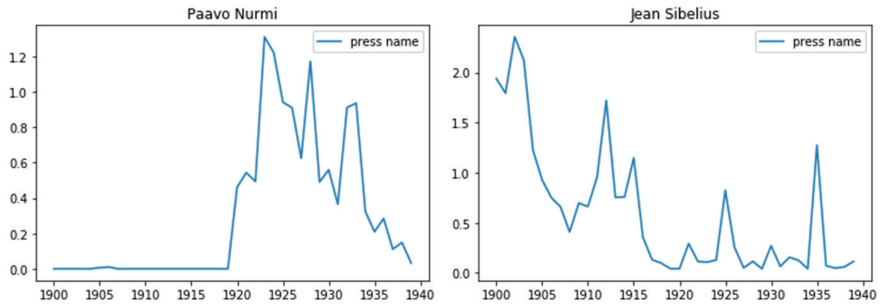

Figure 2. Relative frequency graphs of occurrences in the newspapers for two famous Finns, long distance runner Paavo Nurmi and composer Jean Sibelius.

3.2 List of public figures

A list of those public figures who were active in Finland during 1900–1939 is derived from the National Biography of Finland (NBF). In total, this material contains more than 6,500 Finnish biographies. The backgrounds of the selected individuals vary, but the most common categories of those who were active during the 20th century are (starting with the largest group) members of the government, visual arts & photography, literature, humanities and military & national defence. Our material is based on a dataset created by the Biografiasampo project (BS) (Hyvönen et al. 2018) and obtained through Yasgui SPARQL query interface. Our query was searching for public figures born between 1850 and 1939. The outcome contains 3,393 public figures. We removed all those figures that had three or less first name occurrences in our data but made no other frequency-based filtering. After this procedure, we comprised a list of 2,649 public figures, including information on their sex, date of birth and class categories. The majority of these notables (2,291) are men and the minority (358) are women.

The main issue with the above presented list of public figures is that it is not complete. There are probably many such notables that were often mentioned in the press but were not considered important enough to be added to the National Biography. Another problem is that the material mainly contains Finns and accordingly, many well-known foreign public figures are excluded. The names of the public figures are an issue as well. Especially at the beginning of the 20th century, there was variation in the spelling of names, and many notables were not presented under their official names. This is particularly true for those public figures who were Finnish speakers but were baptised with a Swedish name. However, it seems that at least the majority of public figures are in the BS/NBF dataset with the name they were known for.[3] Furthermore, there are some public figures whose prominence is exaggerated as their names occurred in company names as well (e.g. *Werner Söderström* and *Victor Ek*).

Despite the above-described issues, the number of public figures derived from NBF is extensive and the list consists of people from a wide range of backgrounds. Furthermore, there is extensive press coverage of the majority of the public figures. Thus, it can be argued that the public figures listed in NBF are a comprehensive sample of those who could have influenced Finnish naming conventions at the beginning of the 20th century.

[3] For example, famous Finnish sportsmen *Hannes Kolehmainen* and *Ville Ritola* are listed with the names they were known for although they were officially *Johannes* and *Viljo* (BS/NBF).

3.3 First name statistics of Finland

Statistics of first given names in Finland have been obtained from the records of the Finnish Population Register Centre.[4] The material contains annual details of first given names, their numbers and gender.[5] Our dataset is limited to the years 1900–1939. Furthermore, we obtained Finland's annual birth rate statistics from Statistics Finland and, using these figures, we calculated the relative annual frequency of each name. The data contain 87,406 different names.

Although the records of the Finnish Population Register Centre are comprehensive, it should still be emphasized that they do not always depict the real usage of the names. This is especially true at the beginning of the 20th century when many Finnish speaking new-borns were still baptised with Swedish names although in speech, they were addressed differently. On the other hand, this issue is also an interesting subject for research, as one could assume that the prominences of Finnish and Swedish names in the newspaper data would be in line with the Finnish and Swedish first given names in the official name records.

4. Methods and results

In this section, we present our three computational methods and the results they have yielded. We start with the Pearson's correlation coefficient, then move to Shannon-Jensen Divergence and finally present our analysis regarding coinciding peaks. The results of each are presented immediately following a brief discussion of what the method seeks to capture and what are the major reservations concerning their use.

[4] We obtained the name statistics through the GitHub page of the Finnish Broadcasting Company (YLE): https://github.com/Yleisradio/yle-uutiset/tree/master/names.

[5] For privacy reasons, the value 10 is used when the annual frequency of a name is between one and ten.

4.1 Pearson's correlation coefficient

We sought to test the hypothesis of the influence of public figures' degree of prevalence in the newspapers and popularity of their respective first names with as simple, robust and intuitive method as possible. Perhaps the most common approach for an analysis of this kind conceptualizes the interaction between two phenomena as that of mutual dependence and correlation. In other words, the number of occurrences of public figures' names per year is seen as the independent variable and the number of their first names in name records of the same year form the dependent variable. As we aim to provide the most robust statistical observations in the matter, we study the dependence of these two variables as that of linear correlation and select Pearson's correlation coefficient for measuring the degree of correlation between them.

We calculated the Pearson correlation coefficient between the yearly relative frequencies of each of our 2,649 public figures full name mentions in the newspaper data and the relative frequencies of their first name in the name records. This yielded a correlation score for each name, which is represented in the histogram (Figure 3) below:

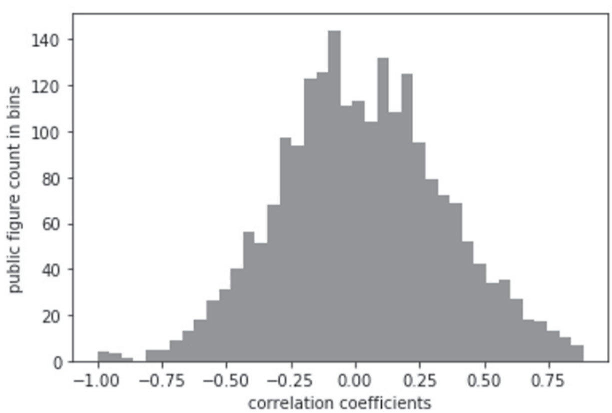

Figure 3. Distribution of correlation coefficients for each NBF public figure as a histogram.

The distribution in Figure 3 is reminiscent of a multimodal distribution with two modes: one near –0.3 and the other near 0.2. While this is surprising at first, this bimodality has a ready explanation. As is noted in Figure 1 in Section 1.1, Finnish first name conventions underwent a relatively remarkable tectonic shift during the period being studied, as names congruent with non-Finnish orthography steadily decreased in both type and token frequency. This means that public figures whose peak of fame gradually declined during the period automatically have some correlation with the frequency of their first name if they happened to have an orthographically non-Finnish name. Similar trends can be seen in name variants with explicitly Finnish orthography: almost all become more frequent over time and thus produce correlation with public figures with growing fame. Thus, a modest positive or negative correlation is to be expected here.

While correlation does not entail causality, it can be reasonably surmised that a very simple and blunt version of causation between public figure's names and first given name trends should result in linear correlation between mentions and given names. As no such correlation is detectable in Figure 3, apart from the modest bimodal peaks explainable by broader patterns of change in names, we can conclude that the frequency of first names is largely independent of frequencies of public figures mentioned in the newspapers. However, Figure 3 also shows possible outliers at both ends of the histogram, where positive or negative correlations are remarkably strong.[6] Some true cases of exceptional mutual co-dependence could indeed be found among these sections, but they do little to challenge the overall interpretation.

4.2 Jensen-Shannon Divergence

However, there is another, complementary, possibility for looking into this matter. While the correlation coefficient focused on the temporal variation of individual names and sought to detect patterns of co-dependency, the repertoires of mentions and given names can be studied

[6] The outliers are represented by the bump near –1.0 mark and comparatively heavier mass near 0.75.

as a whole. In this line of analysis, the question is instead formulated as, 'How well does the distribution across all public figure occurrences predict the distribution of first given names for every year?' A common method used for such analyses is the Küllbeck–Leibler Divergence (KLD) (cf. e.g. Degaetano-Ortlieb and Teich 2018). However, as not all the public figures appear in newspapers every year, nor do their first names appear in name records every year, we opted to use its symmetrical variant Jensen–Shannon Divergence (JSD), following an approach established in similar studies using language data (Pechenick et al. 2015). While mathematically alike in many regards, unlike KLD, JSD yields identical divergence scores regardless of which distribution is selected for base of comparison. The values of JSD range from 0 (meaning the distributions are the same or identical) to 1 (meaning they are maximally different).

The prevalence distribution was obtained by summing over relative frequencies of occurrences of each public figure with the same first given name each year. That distribution was then compared to the distribution of first given names for the same year by computing the JSD similarity between them. Because the data consist of potential confounding factors related to wider patterns of change in naming practices, as was noted above, the JSD similarity scores were compared to a confidence interval obtained by a bootstrapping process. In the bootstrapping, each yearly mention distribution was compared to given first name distributions for every year. The maximum, minimum and mean of these 28 scores were noted and plotted in Figure 4 below. If there was a systematic correspondence between distributions over mentions and given first names, it would be fair to assume that this correspondence was strongest between distributions obtained from the same year rather than random pairings of year.

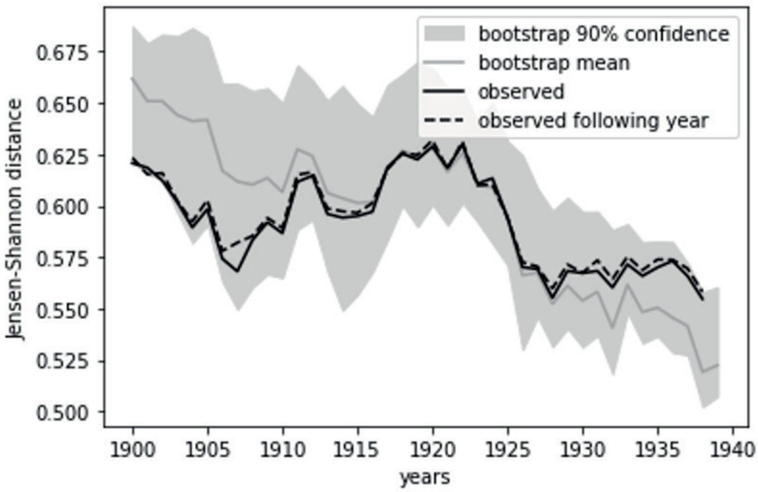

Figure 4. Jensen-Shannon Divergences of consecutive years. The grey area marks the lowest and highest scores when each year is compared to every other year. The grey line is the mean score across all years and black line is the score for the same year.

No such correspondence is visible in Figure 4 above, perhaps apart from the first five years. Rather, generally, the JSD score between the distributions of the same year are neither stronger nor weaker than could be expected when comparing distributions from two random years. This observation complements the one in Figure 3, by giving further grounds to reject an overarching hypotheses of interaction between the occurrences of public figures in the newspapers and their first names in name records. The exceptional pattern in the first five years does however warrant some further questions on whether that truly would be a sign of better predictive power of prevalence distributions. On the other hand, it might simply be to do with the fact that as some (mainly Swedish, see Section 1.1 for more details regarding this development) names present in the data at that time fall out of use. For that reason alone, the earliest years are more similar than the later years. This should result in both the early end having high similarity (low JSD score) and the latter end having low similarity (high JSD scores), which is congruent with Figure 3 above. For the most part, it can be noted that there is no clear evidence supporting a strong correspondence between mentions and given names.

4.3 Coinciding peaks

The analyses above sought to look at very broad trends and temporal variation and investigate whether changes in degrees of publicity resulted in names becoming more frequent or infrequent over time. These resulted in largely negative results. However, previous studies indicate another pattern of interaction between degree of public presence and naming conventions; Kiviniemi (1982:173) observed that questionnaire data showed how roughly 3% of parents were inspired by public figures when selecting first given names for their children. In addition to this survey study, Kiviniemi (e.g. 2006:212–226) has presented many examples of cases where publicity has influenced the popularity of a name.

Examining peaks in popularity of first given names was our third approach to studying the influence of public figures on naming conventions. Here, we calculated how many of the peaks in prevalence graphs coincided with a peak in the first given name graph. Because the prevalence time series tend to demonstrate distinctive spikes, we again compared the observed proportions to a confidence interval obtained by bootstrapping. This time, the 90% bootstrap interval was based on comparing each public figure with 1,000 random first names and selecting the 0.05...0.95 of that distribution.

Also, because there is no single, straightforward definition for a peak, many different range-based definitions were studied. In other words, peaks were defined as the highest value in running ranges of 2 to 19 years, resulting in 18 different definitions. The results of these analyses are presented in Figure 5 below.

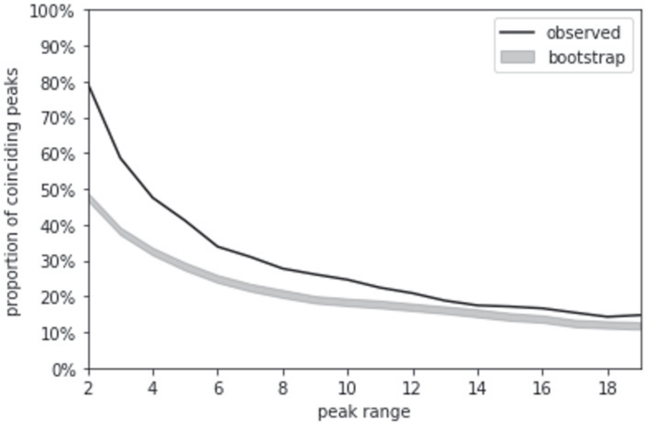

Figure 5. The proportion of coinciding peaks in newspaper occurrences and first given names in name record data. The grey area is 90% bootstrap confidence interval for 1,000 random public figures – first name pairings.

The vertical axis shows the proportions of coinciding peaks, the horizontal axis the different peak definitions running from 2 to 19. The black line represents the actual observed proportions for coinciding peaks between the public figure occurrences and their respective first names. The grey area represents the randomly obtained 90% confidence interval. As can be seen in Figure 5, the observed line stays well above the confidence interval with diminishing difference between them. Even until the peak range of ten years, there are roughly 10% additional coinciding peaks above what could be expected by chance alone. This not only corroborates Kiviniemi's observation (1982:173) but indicates that his estimation of 3% might even be conservative. The interpretation here would be that over the span of ten years, public figures' names would cause three-times as many peaks in popularity of names in first given name records compared to questionnaire data.

 NoSo 2023 | https://doi.org/10.59589/noso.32023.14404

5. Conclusions and discussion

Our aim was to use computational methods to investigate what role public figures have had in determining how first given names were chosen in Finland at the beginning of the 20th century. As far as we are aware, this kind of research has not been conducted before. Thus, one of our main focuses was learning if computational research methods can be used to conduct such a study.

The results corroborated many of the earlier observations made from smaller scale studies. There seems to be no over-encompassing influence of celebrity names influencing how people choose to name their children, or that are at least clear enough to register on such metrics as general correlation or Jensen-Shannon Divergence. It should be noted, however, that metrics such as correlation coefficients and JSD give more emphasis on how the overall probability mass is divided across the distributions than matching peaks, and so have limited interpretability concerning possible causal projections. However, the results regarding the material available is provisional: the number of occurrences of public figures in the newspaper data could possibly change drastically once all of the material becomes available. Nevertheless, the current results seem to give reason to assume this will be relatively unlikely, since there seems to be quite few dramatic changes in distributions that would correspond with different compositions of the dataset.

Furthermore, it is possible that selecting only specific subgroups of public figures (such as prominent cultural figures) might show different levels of correlation. Additionally, it seems that even names with relatively high levels of correlation did not match the given name statistics close enough to suggest causation.

A closer look at individual names shows that often, the influence of public figures on the naming conventions has been much more nuanced than can be studied using statistical methods alone. As noted earlier, part of the positive overall correlation is caused due the decline in foreign names (mainly Swedish ones) at the beginning of the 20th century, both in newspapers and in given names. Needless to say, this phenomenon was induced not only because of the examples set by the public figures but mostly because of the ongoing Finnicization. As the proportion of foreign names (see Figure 2) had decreased rapidly from the 1880s onwards, it is to be expected that their number also declined among the local public figures.

In a similar fashion, it is difficult to measure the influence of public figures in those cases where the name was common among the notables and ordinary people. The name *Juhani* is a good example. Its popularity started to increase from the mid-1920s onwards as indicated by the blue graph in Figure 6 below. The same figure also depicts the newspaper prevalence of car merchant *Juhani Korpivaara* (orange). Based on the Pearson coefficient, these two variables have a strong correlation. However, one can easily notice that the popularity of the name *Juhani* had already increased before any mentions of *Juhani Korpivaara* in the newspapers. Furthermore, the list of the most frequent public figures in the press contains many people whose first name was *Juhani*. Among these men, the author *Juhani Aho* received the most mentions. However, at the peak of his press appearances, the name had already been prevalent since the start of the 1910s, therefore it seems unlikely that his example would have been significant in increasing the popularity of the name *Juhani*. The conclusion is that there is no single public figure named *Juhani* who could be credited with initiating the name's popularity.

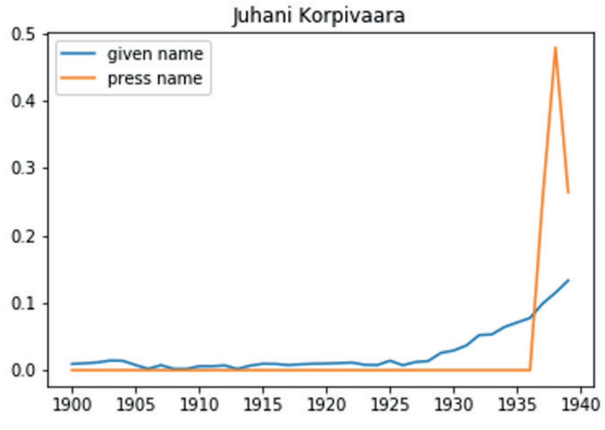

Figure 6. The prevalence frequency of *Juhani Korpivaara* in the press (orange) and the popularity of the name *Juhani* according to the Finnish first given name records (blue).

 NoSo 2023 | https://doi.org/10.59589/noso.32023.14404

On the other hand, the degree to which peak numbers of mentions of public figures in the newspaper data matched with peaks of their respective first names in the name records suggests that instead of overarching correlation, the interaction between these two elements is temporally a much more restricted phenomenon. Over a period of ten years, the peaks coincided roughly 50% more often than could be expected by chance. This observation also corresponds with previous studies, where famous people have been seen to generate peaks of popularity of their first names (Kiviniemi 2006:212–226). Based on our results, the earlier assumptions regarding the role of public figures in naming trends can be further elaborated by claiming that this phenomenon was not restricted to only the most famous individuals. In addition to this, it can be argued that often the influence of public figures caused relatively small and temporally limited popularity peaks.

The case of a male actor and director *Kosti Elo* (Figure 7) is a good example of a situation in which press prominence seems to have caused slight rises in the popularity of a name. The correlation appears to be very strong on multiple occasions. For example, during the first years of the 20th century, the rising press prominence of Kosti Elo is followed by an increase in the popularity of his first name *Kosti*. However, the situation is not as clear as it seems. The problem with names like *Kosti* is that their popularity as a first given name was low. The popularity of *Kosti* peaked in 1904, when 78 were given the name. The lowest point was in 1920, when the name was given to 20 people. With such small numbers, the role of chance must also be considered. On the other hand, as pointed out in Section 3.2, the peaks of press prominence and first given name popularity coincide more frequently than could be expected by chance. Thus, it is completely plausible that the relatively small and temporally limited increases in the popularity of a name, like in the case of *Kosti*, can be explained by the press prominence.

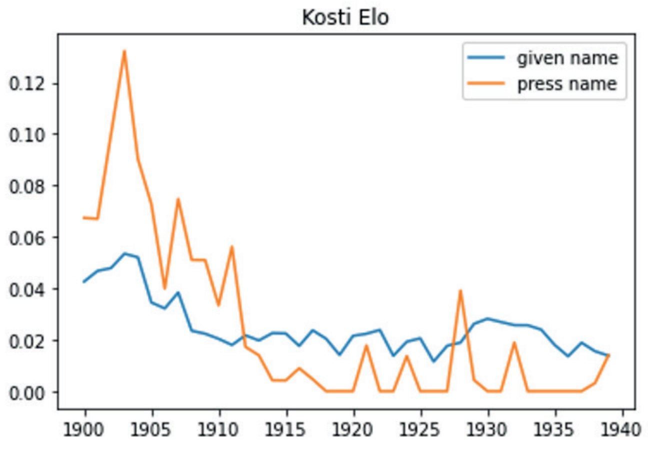

Figure 7. The prevalence frequency of *actor and director* Kosti Elo in the press (orange) and the popularity of the name *Kosti* according to the Finnish first given name records (blue).

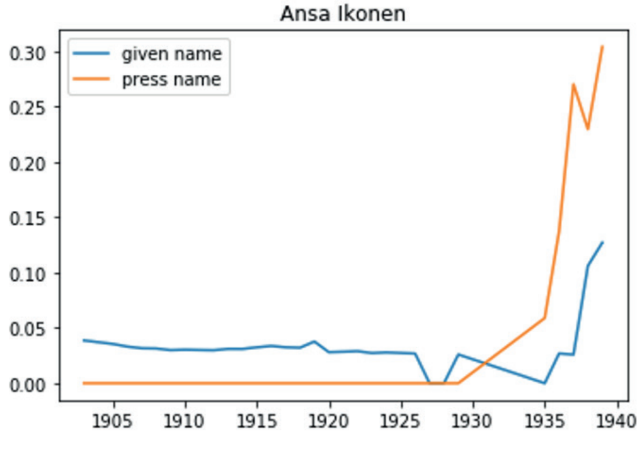

Figure 8. The prevalence frequency of actor *Ansa Ikonen* in the press (orange) and the popularity of the name *Ansa* according to the Finnish first given name records (blue).

There are also cases where the correlation is obvious and strong upon closer inspection. For example, Figure 7, shows how the famous Finnish actor *Ansa Ikonen* and her prevalence in the press (orange) is clearly connected to the growing popularity of the name *Ansa* (blue). We can see that the name *Ansa* was almost non-existent until 1935 and 1936 when Ansa Ikonen became one of the most famous actors in Finland.[7]

The above presented case of *Ansa Ikonen* seems to be a typical example of a situation where the press publicity of a public figure has influenced the naming conventions. What these cases have in common is the low starting level of popularity of the name. As previously stated, more common names like *Juhani* had many other possible factors causing the increase in the popularity. Preliminarily, it could also be argued that female names were more prone to the influences of publicity than the male equivalents. This idea would be in line with earlier notions according to which female name trends tend to vary faster than the male ones (Raunamaa 2020). However, this preliminary notion requires further investigation. The same can be said of the categories of the public figures: which types of public figures influenced naming conventions the most.

In the future, a more elaborate statistical analysis could be developed to take into account the general trend of each name and weighted influence of public figures against that. At the same time, a number of explanatory variables could be considered, such as the gender or professions of public figures. For example, multinomial regression could provide a statistical framework for such an analysis.

References

Ainiala, Terhi, Saarelma, Minna & Sjöblom, Paula. 2016. *Names in focus: An introduction to Finnish onomastics*. Helsinki: Finnish Literature Society.

BS = Biografiasampo. https://biografiasampo.fi/. See project homepage

7 Her real first given name was *Aili*.

https://seco.cs.aalto.fi/projects/biografiasampo/en/ for further information (accessed 20 June 2022).

Borin, Lars & Forsberg, Markus & Roxendal, Johan. 2012. Korp – the corpus infrastructure of Språkbanken. In *Proceedings of the eighth international conference on language resources and evaluation* (LREC'12), 474–478. Istanbul: ELRA.

Brennan, Claire. 2018. Digital humanities, digital methods, digital history, and digital outputs: History writing and the digital revolution. *History Compass* 16(10). https://compass.onlinelibrary.wiley.com/doi/abs/10.1111/hic3.12492 (accessed 20 June 2022).

Degaetano-Ortlieb, Stefania & Teich, Elke. 2018. Using relative entropy for detection and analysis of periods of diachronic linguistic change. *Proceedings of the second joint SIGHUM workshop on computational linguistics for cultural heritage, social sciences, humanities and literature*, 22–33. Santa Fe, New Mexico: Association for Computational Linguistics.

Endén, Rauno & Kokkonen, Jüri & Lyytinen, Eino & McAlester, Gerard. 1996. *Yleisradio 1926–1996: A history of broadcasting in Finland*. Helsinki: Finnish Historical Society.

Hyvönen, Eero & Leskinen, Petri & Tamper, Minna & Tuominen, Jouni. 2018. Semantic National Biography of Finland. In Mäkelä, Eetu & Tolonen, Mikko & Tuominen, Jouni (eds.), *Proceedings of the digital humanities in the nordic countries 3rd conference (DHN 2018)*, 372–385. (CEUR Workshop Proceedings 2084). http://ceur-ws.org/Vol-2084/short12.pdf (accessed 20 June 2022).

Kiviniemi, Eero. 1982. *Rakkaan lapsen monet nimet: suomalaisten etunimet ja nimenvalinta*. Espoo: Weilin + Göös.

Kiviniemi, Eero. 2006. *Suomalaisten etunimet*. Helsinki: Suomalaisen Kirjallisuuden Seura.

Kuikka, Martti. 2010. Opettajankoulutus eilen, tänään ja huomenna. *Koulu ja menneisyys* 48. 1–18.

Leibring, Katharina. 2016. Given names in European naming systems. In Hough, Carole (ed.), *The Oxford handbook of names and naming*, 199–213. Oxford: Oxford University Press.

Marjanen, Jani, et al. 2019. A national public sphere? Analysing the language, location and form of newspapers in Finland, 1771–1917. *Journal of European Periodical Studies* 4(1). 54–77.

NBF = National Biography of Finland. Helsinki: Finnish Literature Society, 1997–. https://kansallisbiografia.fi/ (accessed 20 June 2022).

Nyström, Staffan. 2016. Names and meaning. In Hough, Carole (ed.), *The Oxford handbook of names and naming*, 39–51. Oxford: Oxford University Press.

OSF = Official Statistics of Finland: Births [e-publication]. Helsinki: Statistics Finland. http://www.stat.fi/til/synt/tau_en.html (accessed 20 June 2022).

Ogilvie, Brian. 2016. Scientific archives in the age of digitization. *Isis* 107(1). 77–85.

Pechenick, Eitan Adam & Danforth, Christopher & Dodds, Peter Sheridan. 2015. Characterizing the Google Books corpus: Strong limits to inferences of socio-cultural and linguistic evolution. *PloS ONE* 10(10). https://doi.org/10.1371/journal.pone.0137041 (accessed 20 June 2022).

Raunamaa, Jaakko. 2020. Margareta, Katerina and Kristina: Female names in medieval Finland. *Studia anthroponymica Scandinavica* 35. 75–98.

Raunamaa, Jaakko. 2021. *Finnic anthroponymy in the Middle Ages: Perspectives on pre-Christian and early Christian personal names*. Helsinki: University of Helsinki. (Doctoral dissertation.)

Saari, Mirja. 2012. The development of Finnish into a national language. In Hüning, Matthias & Vogl, Ulrike & Moliner, Oliver (eds.), *Standard languages and multilingualism in European history*, 179–204. Amsterdam: Benjamins.

Salway, Benet. 1994. What's in a name? A survey of Roman onomastic practice from c. 700 BC to AD 700. *The Journal of Roman Studies* 84. 124–145.

Suomen tilastollinen vuosikirja 1903 = Suomen tilastollinen vuosikirja 1903. Helsinki: Tilastollinen päätoimisto.

Vilkuna, Kustaa & Mikkonen, Pirjo & Paikkala, Sirkka. 1988. *Uusi suomalainen nimikirja*. Helsinki: Otava.

Ylikangas, Heikki. 2002. *Käännekohdat Suomen historiassa: pohdiskeluja kehityslinjoista ja niiden muutoksista uudella ajalla*. Helsinki: WSOY.

Nordisk tidskrift för socioonomastik / Nordic Journal of Socio-Onomastics (71–96)

'The super-fats, the middlies and the barely-fats': Naming and describing the body in plus-size fashion blogs

Hanna Limatius

Abstract: This article investigates the practices of naming and describing bodies in plus-size women's fashion blogs. Building on Jeffries' (2007) work on the construction of bodies and bodily processes in women's magazines, I explore how bloggers use language to construct their identities through references to bodily characteristics. I investigate terms the bloggers use to describe the plus-size body, as well as the connotative features of these terms. In addition, I analyse the use of three terms that were used to describe bodies that are not plus-size: *thin*, *slim*, and *skinny*. The results demonstrate that, while the bloggers build counter-discourse to the mainstream media discourses that construct fatness as a negative characteristic, they also maintain particular hegemonic discourses on beauty, sexuality and gender.

Keywords: body positivity, fashion blogs, fat acceptance, gender, identity, othering, weight stigma

Hanna Limatius (University of Vaasa). 'The super-fats, the middlies and the barely-fats': Naming and describing the body in plus-size fashion blogs.

NoSo 2023 | https://doi.org/10.59589/noso.32023.14410

1. Introduction

The way we describe ourselves contributes to the construction of our identity. Particularly in social media and other online settings, the words that we choose for presenting ourselves are significant, as we lack some of the cues that contribute to self-presentation in face-to-face interactions. By using specific terminology to describe ourselves, we can discursively highlight particular aspects of our identity and downplay others (Goffman 1990). Similarly, we can use words to identify ourselves with other members of a social group or community – that is, the choices we make in describing ourselves online also affect our social identity (Gumperz & Cook-Gumperz 1982; Tajfel 1982).

In this article, I explore the ways in which identity is constructed online through practices of naming and describing the body (and parts of the body, see Jeffries 2007). Specifically, I study the ways in which the body is referenced in fashion blog texts written by plus-size women. I am particularly interested in the specific words that blog authors use to describe their bodies and the connotative features of these words. In the context of this article, I consider this type of language use related to bodies and body parts as naming, since it reflects the bloggers' preferred way of addressing their bodies, as well as their identification with similar others, i.e. the in-group. There is a broad variety of terms that can be used to describe a plus-size woman's body, yet the bloggers tend to favour certain words (Limatius 2020). The fact that plus-size people are an often-marginalized group, particularly in the context of fashion, is relevant here, as the process of naming their bodies online gives these women agency over their bodies. As Chibuwe, Mpofu and Bhowa (2021:3) put it, 'self-naming becomes an act of exercising power over the self'. However, as defining one's own identity typically also includes processes of defining the other, I am also interested in how plus-size fashion bloggers name bodies that are not plus-size. Therefore, I also investigate what type of words the bloggers use to describe bodies and body parts that are different to theirs and thus represent the out-group.

 NoSo 2023 | https://doi.org/10.59589/noso.32023.14410

I will use a qualitative study informed by methods of corpus-assisted discourse analysis (Baker 2006) to address the following research questions:

1. How are plus-size bodies named and described in the blog data?
2. How do the bloggers name and describe bodies that are different to theirs?
3. What kind of connotative features does the naming and describing of bodies have in the context of plus-size fashion blogging?

In order to answer these questions, I will first position my work in the broader context of research into language, gender and fatness. As I will demonstrate, the intersections of these topics are relevant for the practice of constructing identity through naming and describing the body. Similarly, as Pilcher (2016:766) states, the concepts of naming, identity and embodiment are also ‘inextricably intertwined’. After addressing the role of language in constructing identity specifically in the contexts of gender and fatness, I will move on to the concepts of naming and describing, explaining how I approach them based on Jeffries’ (2007) framework for studying women’s magazines. After providing theoretical background to the topic, I will introduce the data and the analytical method. Finally, the empirical section of the article will present the findings of the qualitative analysis along with my conclusions.

The main results of the analysis can be divided into three categories. First, the bloggers in the data name the self as an empowered agent, highlighting the fact that by choosing the language to describe themselves with, they are in control of their representation. Under this category, I discuss two more specific ways of using descriptive terminology as an empowering resource: reclaiming existing terms and innovating new terms. Second, the bloggers also construct a collective, social identity by naming the self as a member of a group. Here, I will address the use of inclusive pronouns, the use of the word *fat* as a noun, and the practice of naming the blogosphere as a shared ‘space’ or community. Finally, through an investigation of three terms that were used to describe the bodies of thinner women – the others – I will demonstrate how practices of naming and describing bodies are also used to construct difference. The results of the analysis provide new insights into the significance of naming bodies and body parts, particularly in the context of fashion blogs. My work shows that while many of the naming practices in blogs are comparable to those in women's magazines (Jeffries 2007), they are used to construct different ideas of 'normal' bodies that reflect the norms and values of the plus-size fashion blogging community.

2. Language, gender and fatness

When studying the language use of plus-size fashion bloggers, it is necessary to consider the societal context of weight stigma, particularly in relation to fashion. People with plus-size bodies are marginalized in contemporary Western societies, and their representation in the media is dominated by medicalized, moralizing discourses that construct fatness as a burden to society, a failure in self-control and a risk to both individual and national health (Dickins et al. 2011; Brookes & Baker 2021). In the fashion industry, thin bodies have long been particularly idolized and fat bodies particularly vilified, although there have been attempts to reconstruct these norms in recent years as a result of the popularization of the body positivity movement (Sastre 2014; Zavattaro 2021; Brathwaite & DeAndrea 2022). Social media has been in a pivotal role in the shift towards more body diversity in fashion imagery – according to Czerniawski (2021:3), 'The rise of [body positive] hashtag campaigns coincides with the decline in influence of mainstream magazines and traditional advertising campaigns'.

Over the past two decades, plus-size women's fashion blogs have become central spaces for producing counter-discourse to the one-sided representation of fat people in the mainstream media (Connell 2013; Scaraboto & Fischer 2013; Harju & Huovinen 2015; Limatius 2020). According to Cooper (2016:76), practicing 'fatshion' blogging 'encourages fat people's creative participation in spaces where they are usually excluded'. In addition to the aforementioned body positivity movement, other social justice movements such as the fat acceptance movement (Zavattaro 2021:283) and the health at every size movement (HAES, The Association for Size Diversity and Health 2020) have provided fat people with resources to oppose the 'obesity epidemic' focused discourse of the mainstream media. Language has an important role in the construction of these counter-discourses. According to LeBesco (2001:76), fat people can create and regulate 'a new social reality through the use of words', resulting in a 'more inhabitable subject position' than the discourses that characterize fatness in current Western media. The reappropriation of the word *fat* is one of the most concrete ways in which activists have harnessed language into creating counter-discourse. Previous research (Connell 2013; Scaraboto & Fischer 2013; Harju & Huovinen 2015; Limatius 2020; Zavattaro 2021) illustrates how bloggers and influencers have embraced the word purposely in order to lessen its stigma and to normalize it as a neutral descriptor.[1]

While fat people of all genders experience discrimination, women and homosexual men are more likely than heterosexual men to face strict beauty and body standards (Hartley 2001:67).[2] Moreover, weight stigma intersects with other types of marginalization, resulting in increased stigmatization of fat people of colour (Zavattaro 2021:284). Women's bodies in general, particularly in the context of sexuality, have traditionally been subject to societal and discursive control (Ho & Tsang 2005:524), which is also reflected in the portrayal of women's bodies in the media. For example, Brookes and Baker (2021:200) report that in the British press, women's obesity is discussed more

[1] It should be noted that outside of the body positivity and fat activist communities, this kind of neutral use has not yet been normalized, and *fat* continues to be a contested term. However, as LeBesco (2001:76) names academics as one group of people who can participate in transforming the subject position of fat people through language, I also use the term *fat* as a neutral word with no derogatory intent throughout this paper.

[2] However, as noted by Brookes and Baker (2021:178) unobtainable body ideals for men have also increased in recent years.

than men's, and the articles on women, obesity and weight loss also have a more aesthetic focus, where being 'beautiful' and fitting into specific items of clothing is highlighted. Discourses on womanhood and sexuality are inescapably tied to hegemonic views of bodies as 'beautiful', 'acceptable' or 'normal' – and, by default, also as 'ugly', 'unacceptable' or 'abnormal'. Media such as women's magazines (Jeffries 2007) and fashion blogs contribute to these discourses either by maintaining hegemonic ideals or (less commonly) by challenging them. As Jeffries (2007:21) puts it, references to the body construct the body. Since the body has a central role in the discursive construction of gender and sexuality, such references also contribute to our understanding of those concepts.

3. Naming and describing the body

In this article, I apply Jeffries' (2007:66) framework for studying the construction of women and their relationship to their bodies and bodily processes to fashion blog data. Jeffries (2007:65) investigated the 'naming and describing of women and their bodies' in women's magazines. I consider the framework suitable for blog data, as many parallels can be drawn between fashion blogs and women's magazines. Both media are aimed at an audience consisting predominantly of women, both offer examples and advice on how to fashion the body in different ways to their readers, and both typically have commercial motivations, although in the case of blogs these are not always as obvious. This means that in addition to expressing the author's opinions, the texts are also selling us specific ideas on how we should view and treat our bodies.

Jeffries (2007:63) considers the noun phrase as 'the basic unit of naming'. This means that when studying the ways in which the body is named, it is relevant to consider both the choice of head nouns as well as how that head noun is modified (Jeffries 2007:63). Thus, modifiers such as adjectives are also investigated as a 'part of the naming process' – which is why Jeffries (ibid.) talks about naming and describing the body. The processes of naming and describing are con-

 NoSo 2023 | https://doi.org/10.59589/noso.32023.14410

nected, and they both contribute to the textual construction of meaning and the incorporation of ideologies into texts (Jeffries 2007:63).

Jeffries’ (2007:78) analysis of naming conventions revealed tendencies toward particular ways of constructing the reader of women’s magazines. That is, the reader was either textually constructed as a specific kind of woman, or ‘identified with her body parts’ (Jeffries 2007:78). In the latter type of construction, sexual or sexualized body parts were of particular significance (ibid.). Similarly, Pilcher (2016:771–772) has also discussed nicknames that emphasize particular body parts and aspects of the body (e.g. *fatso* or *ginger*). According to her, such nicknames can be used either for highlighting belonging to a group or for othering individuals and marking them as outsiders. As I will go on to demonstrate in the analytical section, naming and describing the body is used for both purposes in the blog data.

Establishing categories of similar and different bodies through naming and describing results in ideas of what kind of body is seen as the norm. The notion of the ‘real’ or ‘normal’ body is created through (repeated) social performance, just like the notion of ‘real’ or ‘normal’ gender (Butler 1999). Indeed, a common theme in women’s magazines is the evaluation of the normality of the body, which Jeffries (2007:85) mentions as particularly noticeable in magazine articles aimed at teenagers. Throughout her data, Jeffries (2007:85) also notes the prevalence of adjectives that evaluated the body in a positive or negative way, particularly in a ‘before and after’ context.

The idea of the body as ‘natural’ or ‘unnatural’ – in other words, right or wrong – was also reflected in the presence of the ideology of ‘the unchanging body’ (Jeffries 2007:89) in women’s magazines. By this, Jeffries refers, for instance, to the expectation that the body should return to its ‘natural’ (i.e. attractive, socially acceptable) state after childbirth. According to Jeffries (2007:89), the body in women’s magazines is thus ‘portrayed as serving the individual, as able to let us down, and to some extent able to be “fixed”’. This results in the reader of the texts being left with the pressure to ‘create the perfect, natural and normal body by any means possible’ (Jeffries 2007:101). The idea of the body as something to be ‘fixed’ is particularly relevant in the case of fat bodies and their media representation – in addition to women's magazines, for example numerous reality television shows focus on 'fixing' fat bodies either through weight-loss or particular ways of dressing up (Zavattaro 2021:283). Since plus-size fashion blogs are authored by fat women themselves, they have more agency in creating their representation, which results in different discourses concerning the body. However, although these blogs may resist the hegemonic ideas of 'right' kind of bodies, they cannot escape such definitions completely. Similarly, while they rely on a different conceptualization of the 'real' body compared to Jeffries' (2007) data from women's magazines, many body positive marketing campaigns continue to exclude a variety of bodies and thus create their own norms and boundaries based on features of the body (Czerniawski 2021).

4. Data and method

In connection with a previous study (Limatius 2020), I compiled a corpus consisting of 20 UK-based blogs that focused on fashion from the perspective of a plus-size woman. The blog corpus, which was compiled in 2015, contains 7,776 blog posts (3,177,959 words) and 32,737 comments (1,000,049 words). All blogs featured in the corpus were publicly available at the time of data collection, and the authors of the blogs were informed about the inclusion of their blogs in the corpus.

As the goal of the present study is identifying discursive patterns in the use of specific terms, the analysis is based on methods of corpus-assisted discourse analysis (Baker 2006). A list of all descriptors that were used to describe plus-size women's bodies (or parts of their bodies) in the corpus was compiled using the Word List function in AntConc. The resulting list consisted of altogether 45 terms (Limatius 2020). In order to expand the investigation to include identification through difference, in 2021, I compiled another dataset that featured all occurrences of *thin*, *slim* and *skinny* in the blogs. After identifying relevant terms, I conducted a concordance analysis in order to examine the wider context in which the terms were used (Brookes & Baker 2021:32). Occurrences where the words were used to describe things other than bodies or body parts (e.g. ‘big surprise’) were excluded from the analysis. As the use of the terms needed to be critically considered in the specific context of the plus-size fashion blogging community, the main focus of the analysis is on a qualitative, discourse-analytical approach.

In her study on the naming and describing of women and women’s body parts, Jeffries’ (2007:66) presents a ‘checklist of textual features’ that were examined. In my analysis of fashion blog data, the following categories, adapted from Jeffries (2007:66), are of particular relevance: use and reference of pronouns, connotative features of chosen terms, and structural properties of noun phrases (e.g. the use of modifying adjectives). However, since descriptive words – mainly adjectives – were the starting point for my analysis, the analytical process also differs slightly from that of Jeffries’ (2007). In order to study the structural properties of noun phrases, I examined the nouns that the adjectives which had been established as descriptors for bodies in my previous work (Limatius 2020) modified (e.g. the nouns referred to as *fat*, *plus-size*, *big*, etc.). I also investigated the pronouns that were used to refer to the noun phrases that featured such modification (e.g. *I*, *we*, *they*). Finally, in order to study the connotative features of the descriptors chosen by the bloggers for specific bodies and body parts, I examined the broader context around the use of these descriptors through the lens of earlier literature in fat studies (e.g. LeBesco 2001; 2004; Cooper 2016; Czerniawski 2021) as well as previous research focusing on the community-specific linguistic and discursive practices of plus-size fashion bloggers (Connell 2013; Scaraboto & Fischer 2013; Harju & Huovinen 2013; Limatius 2020).

5. Findings

The results of the qualitative analysis of blog data can be divided into three categories: naming the self as an empowered agent, naming the self as a member of a group, and naming the self as different from others. In the following sub-sections, I will discuss this categorization and provide examples from the blog texts to illustrate different practices for naming and describing the body under each category.

5.1. Naming the self as an empowered agent

The bloggers in the data constructed the self as an empowered agent through naming and describing their body in specific ways. In this section, I will discuss two distinct practices that could be observed in the blog texts: assigning new meanings to existing terminology (i.e. reclaiming the terms), and creating innovative new terms.

5.1.1 Reclaiming existing terms

The most popular terms used for describing the plus-size body in the corpus were *fat* (1,869 occurrences), *plus-size* (901) and *big* (413) (Limatius 2020). These terms could be found in all 20 blogs, although there were differences in the frequency of their use between individual bloggers (for a more in-depth investigation of the frequencies, see Limatius 2019). The prevalence of *fat* is particularly notable due to its traditionally negative connotative features and the potential that plus-size fashion blogs have for challenging these connotations through different representations of fatness. Such resistance of dominant discourses can also be observed in previous literature on naming and the body. For example, Pilcher (2016:772) mentions the fact that derogatory nicknames can be re-appropriated 'as a source of resistance and empowerment', while Ho and Tsang (2005:524) discuss how women can (re)present themselves through personal narratives of their bodily relationships.

In the context of women's magazines, Jeffries' (2007) discussed the connotative features of adjectives that were used to modify nouns referring to the body. According to Jeffries (2007:84), the majority of these adjectives 'could easily be interpreted in the context as hyponyms of either *good* or *bad*' – thus, they were used to construct specific bodily features as positive or negative.

 NoSo 2023 | https://doi.org/10.59589/noso.32023.14410

1) And it may just have another little article by **your favourite fat girl**.... Buy your copy now!

In Example 1 from the blog data, *fat* is used by the author in a self-reference to ‘your favourite fat girl’. This noun phrase describes the author’s body as *fat*, but the other modifiers of the noun *girl* make it clear that the term is not used in a derogative sense. The use of *your* implies closeness between the blogger and the reader, while *favourite* has obvious positive connotations. Thus, *fat* is not a bad word for the blogger – the blog text has different ‘good or bad’ distinctions compared to those typically seen in the mainstream media (Jeffries 2007), where describing someone as a ‘fat girl’ would likely be interpreted negatively. This phenomenon is also demonstrated by the bloggers’ use of some words that are traditionally connotated more positively than *fat*:

2) I do however, **class myself curvy**, I also know I’m **fat too**, but with **a large bust a small(ish) waist** and **some rather large hips** I have what you might class as **a curvy figure**.

The blogger in Example 2 discusses the use of the word *curvy*. In most contexts, *curvy* could be interpreted as a more flattering term than *fat* – for example, plus-size fashion brands are more likely to market their clothing to ‘curvy women’ than to ‘fat women’. However, in the plus-size fashion blogging community, the term is often debated, as some bloggers and fat activists see it as linguistic sugar-coating that impacts the representation of ‘actually fat’ women negatively (Limatius 2017). *Curvy* (159 occurrences) was also notably less popular than *fat* (1,869 occurrences) in the corpus, with some bloggers not using it at all (Limatius 2019). The somewhat problematic nature of the term within the community is reflected in Example 2 – while the blogger names herself as *curvy*, she makes sure to mention that identifying with the term does not take away her fatness. Similarly, she describes her waist as *small(ish)* instead of merely *small*, and provides justification for why she *might* be classed as having a curvy figure. Describing the body as *curvy* is performed carefully, while being *fat* is stated as a fact. Thus, common descriptors of bodies that might also be found in women's magazines, for example, are used by the bloggers, but they use them in a way that is in line with the community's norms. *Fat* is not vilified, and the smallness of specific body parts is not idealized.

5.1.2 Innovating new terms

In addition to describing their bodies and body parts with relatively common words such as *fat* or *curvy*, the plus-size fashion bloggers named their bodies using more creative formations, such as wordplay and abbreviations. These more innovative terms included *fatshionista* (2 occurrences), *fatfashionista* (1), *inbetweenie* (2), *bootiful* (1), *boobilicious* (1), *bootylicious* (1), *booby* (3), *chubs* (2), *chubster* (8), *bigfatfatty* (1) and *BBW*, which is an acronym for 'big, beautiful woman'(21). While these terms were notably less common than constructions such as '*fat* + head noun', they are nevertheless worth discussing in the context of constructing identity through naming the body.

Jeffries (2007:70–71) noted a tendency in women's magazines 'for people to be identified by the body part under consideration' – for example, women were equated to their vagina or breasts ('Whether you're big, small, pointy, saggy or pert, chances are you wish your boobs were different (*Bliss*)'). According to Jeffries (2007:71), this practice was unique to discussions of either sexual or reproductive body parts, or weight, which she states is 'usually considered to be a factor in (sexual) attractiveness'. Thus, in women's magazines, women were only equated to body parts with sexual connotations, and not other parts of their body.

Some of the innovative terms in the blog corpus display a similar pattern – they are related to traditionally sexualized body parts, such as breasts (*boobilicious*) and buttocks (*bootiful*). The term *BBW* also has sexual connotations. While not explicitly linked to any particular sexual body part, *BBW* is a term widely used in the porn industry and to convey sexual attraction to women with larger bodies.

NoSo 2023 | https://doi.org/10.59589/noso.32023.14410

3) You should all be aware by now that I am on a mission to prove that **short, fat, old birds** can be **bootiful** too.

In Example 3, the blogger names herself as a *short, old, fat bird*, who despite these presumably negative characteristics, can still be *bootiful*. Here, we see that the discourse on bodies and beauty that is constructed in the blogs is still affected by hegemonic ideas of what constitutes an attractive – and thus socially acceptable – woman. The blogger’s statement implies, for example, that older women are usually not considered beautiful (or sexual). Moreover, while the term *bootiful* may be merely intended as humorous word play, it nevertheless connects the notion of beauty to a traditionally sexualized part of a woman’s body. The blogger’s language use highlights her agency, since she states her ‘mission’ to defy hegemonic beauty standards, but at the same time it contributes to the discourse where all women, regardless of age and size, should strive to be as ‘beautiful’ as possible.

In addition to naming the body after sexualized body parts, there were also creative word formations in the data that stemmed from the most popular descriptor in the corpus, *fat*. Some of them were amalgamations of *fat* and *fashion*, like *fatshionista*, which has an established role in the language of the community (Gurrieri & Cherrier 2013; Scaraboto & Fischer 2013), while others appeared to intensify the notion of fatness (*bigfatfatty*). Some, namely *chubs* and *chubster*, are plays on another descriptor related to fatness, *chubby*:

4) For **us chubs** it really is hit and miss in there, although they go up to a size ‘20’ and yes I mean a ‘20’ some garments are more generous than others.

In an earlier study, I noted an element of casualness and humour in the ways in which the bloggers referred to the bodies of other plus-size women, perhaps in order to highlight the fact that using words traditionally considered as derogatory was not intended as an insult (Limatius 2017). While the use of *fat* and other terms that highlight fatness instead of ‘curves’ in a positive, reclaimed sense was common within the community, some bloggers preferred more ‘neutral’ language (e.g. the term *plus-size*) (Limatius 2020). Some women may interpret specific words that refer to body parts as derogatory while others experience them as empowering (Ho & Tsang 2005:527). In Example 4, the blogger describes a particular group through identifying a shared bodily feature – namely, ‘chubbiness’. Such language use appears playful and comparable to the use of nicknames, which, according to Pilcher (2016:771–772) can signal affection and belonging. However, one needs to be aware of the possible different interpretations of naming practices within the group in order to successfully create feelings of community and belonging instead of hierarchies and divisions.

5.2 Naming the self as a member of a group

While many of the practices of naming and describing the body in the corpus were focused on the body of the individual blogger herself, the data also contained references to plus-size women as a group with shared bodily characteristics. In such examples, the blogger was named as one member of the group, and the bodies and body parts of other bloggers or blog readers were also referenced. In this section, I discuss three features of the bloggers’ language that were of particular relevance for constructing the in-group: the use of inclusive pronouns, the use of the word *fat* as a noun, and naming the blogging community as a shared space.

5.2.1 Pronoun use

As demonstrated in Examples 3 and 4 in the previous section, the bloggers often identified themselves as a part of a larger collective of people that was named after shared characteristics, like *short, fat, old birds* in Example 3 or *us chubs* in Example 4. This type of language use can be seen to construct social identity (Gumperz & Cook-Gumperz 1982; Tajfel 1982), as it identifies the individual bloggers as members of a specific social group and implies closeness or attachment to

 NoSo 2023 | https://doi.org/10.59589/noso.32023.14410

this group (Benwell & Stokoe 2006:25). One typical way of establishing and naming this in-group is via pronoun use:

5) I am really trying to wear more skirts and I find that pencil skirts and midi skirts can be dressed up and down and [they are] very flattering for **us plus sized girls**.

Jeffries (2007:70) refers to the readers and writers of women’s magazines’ articles ‘splitting up the world’ based on particular bodily characteristics. In Example 5, identifying as plus-size creates such a division – there is an implication that women who are not plus-size (i.e. not part of ‘us’) need not worry about their skirts being ‘flattering’. By naming *us plus sized girls* as a group, the blogger simultaneously identifies her own body in relation to particular fashions, and offers advice to her readers on how to dress in a ‘flattering’ way. A similar practice was also present in Jeffries’ (2007:69) data from women’s magazines, where answers to readers’ letters typically featured the use of the first-person plural pronouns *we* and *us*. The advisors included themselves in the same category as the letter-writers in order to construct generalizations that presented the writers and their bodies as ‘normal’, and thus acceptable.

5.2.2 *Fat* as a noun

Another way of naming oneself as a member of a group in the corpus was reflected in the use of the word *fat* as a noun instead of an adjective. This practice, interestingly, occurred particularly when the word was used in its plural form *fats* (21 occurrences) in reference to a group:

6) Even better would be a girly get together with **all my fabulous fats** and I may have found the perfect place.

In Example 6, the blogger labels a particular group with the noun phrase *all my fabulous fats*. The expression is affectionate, referring to a group of friends, all of whom happen to be fat. Yet, fatness is seen as such a defining shared characteristic that the group is named after it. According to Jeffries (2007:72), 'the use of a nominal form indicates a characteristic which is represented as more permanent and intrinsic to the person than the adjective'. Jeffries (ibid.) theorizes that as nouns can be considered labels, they have a starker effect than adjectives – nouns are 'likely to sound like a categorization', while an adjective 'simply notes a particular characteristic'. In light of this view, we can consider labelling oneself as ***a** fat* to be a stronger and more permanent marker of identity as describing the self or others with the adjective *fat*:

7) I don't want to drive a wedge between **the super-fats**, **the middlies** and **the barely-fats**.

In Example 7, we are clearly dealing with a categorization, since the blogger is referring to specific groups of *fats* who can be placed at different points on a spectrum of fatness. Thus, the social identity of being one of *the fats* appears divided into smaller groups of people based on their 'level' of fatness. People who are slightly overweight might be considered as what the blogger refers to as *barely-fats* – they do not fit the ideal beauty standards maintained by the fashion industry and mainstream media and may participate in practices of fashion blogging and blog reading to find alternative representations. However, the challenges they face are different than the challenges of those who are at the higher end of the fat spectrum. These people are not just marginalized in terms of fashion, but likely face discrimination on several levels of their everyday lives. Moreover, for the *super-fats*, finding fashionable clothing may be virtually impossible (Gurrieri & Cherrier 2013:279). Acknowledging distinctions between different groups by naming them, the blogger displays her awareness about such issues, and identifies the plus-size blogosphere as a diverse community.

5.2.3 Blogs as an in-group space

Another way of highlighting groupness in the data was naming the space where the plus-size blogging community interacted. The blog-

 NoSo 2023 | https://doi.org/10.59589/noso.32023.14410

ging community was typically named as an empowering space that the bloggers had ‘discovered’ or ‘joined’:

8) I am not sure how long I have been using the word fat for, but I’ve certainly been using it a lot more since I started blogging and joined **the ‘fatosphere’** (or the **#psbloggers crew**, hehe).

Again, the terms used by the blogger in Example 8 are rooted in the common descriptors for women’s bodies – the *fatosphere* contains the word *fat*, while *#psbloggers crew* is derived from *plus-size bloggers*. *Fatosphere*, like *fatshionista*, is a well-established part of the lingo of plus-size fashion bloggers. Gurrieri and Cherrier (2013:279) define the *fatosphere* as

> A loosely interconnected network of online resources aimed at creating a safe space where individuals can counter fat prejudice, resist misconceptions of fat, engage in communal experiences and promote positive understandings of fat.

They also refer to the bloggers creating an alternative space for resisting mainstream beauty ideals and constructing new ones (Gurrieri & Cherrier 2013:283). It is meaningful that this place is named, and that the term *fat* is incorporated into its name. Referring to the network of blogs and the communal ties formed between bloggers as the *fatosphere* marks it as a space specifically for those with fat bodies. Similarly, the use of *#psbloggers crew* limits the community to those who use the hashtag #psbloggers on social media in order to identify themselves with the group of bloggers who all share the characteristic of being plus-size. In terms of social identity, naming the group after bodily characteristics separates the in-group from the others. Through such language use, the bloggers demonstrate that the space occupied by the group is not for everyone.

5.3 Naming the self as different from others

Finally, as establishing difference between oneself and the other is also a relevant practice in identity construction (Grad & Martín Rojo 2008:12), I examined words that the bloggers used when referring to women who were not plus-size. The three words that were investigated were *thin* (108 occurrences), *slim* (129 occurrences, including verbs *to slim/slim down*), and *skinny* (93 occurrences).

5.3.1 The usage of *thin*

In the corpus, the word *thin* was typically contrasted with *fat* – women were divided into fat women and thin women, and people more generally into fat people and thin people. This was a very concrete example of the bloggers 'splitting up the world' (Jeffries 2007:70) based on features of the body. The words could also be used to describe opposite ends of a spectrum of bodies:

> 9) All bodies are good bodies – **thin and fat and everything in between**.

In Example 9, the blogger contributes to a typical body positivity discourse, evaluating *all bodies* as 'good'. Here, *fat* and *thin* are named as the extremes, while all other types of bodies are generalized into *everything in between*. Although *thin* is positioned as the opposite of *fat*, neither of these terms are connotated negatively.

However, there were also instances of negative evaluation of thin bodies in the corpus. In Jeffries' (2007:97) data from women's magazines, the discussion on body parts was often 'concerned with what is wrong with them', which resulted in frequent use of intensifying adjectives that evaluated these body parts negatively. A similar evaluation was present in the blog data in the context of *thin*, as intensifying adjectives were often used to modify the term. I found instances of *very thin*, *rail thin*, *extremely thin*, *unhealthily thin*, *exceptionally thin*, and *really thin* in the data. By modifying the word in such a way, the bloggers constructed *thin* as problematic and not something to aspire to:

10) I don’t think that **being unhealthily thin** is a goal to aspire to and I’m certainly not wishing to be **really thin**.

In Example 10, the blogger talks about thinness in connection with a discussion on weight loss. As I discovered in an earlier study, plus-size fashion bloggers who were in the process of losing weight often provided justifications for it, for example by stating health-related concerns (Limatius 2017). In this way, the bloggers strove to prove to their readers that they were still body positive and had not simply succumbed to the pressures of hegemonic beauty ideals. However, at the same time, this type of reasoning contributed to the dominating discourse of fat as unhealthy. Example 10 demonstrates another way of making weight-loss more acceptable within the community – a reassurance that the blogger does not ‘wish to be *really thin*’. This type of language use also contributes to ideas of weight and health as inescapably connected by emphasizing that being *too* thin is not ‘healthy’. Thus, the blogger problematizes thinness in the same way the mainstream media typically does with fatness.

5.3.2 The usage of *slim*

Out of the three terms investigated, *slim* appeared to have the most positive connotations in the data. It was used to describe *friends*, as well as more general groups, such as *people*, *women*, *girls*, *readers*, and *models*. Nevertheless, a difference between *slim* and *fat* was still emphasized:

11) Before I begin, I’d like to give examples to **the slim readers amongst us** to what seeking healthcare whilst fat is like.

In Example 11, the blogger simultaneously acknowledges *slim readers* as part of her blog’s following, and draws a boundary between the slim and the fat. There is an assumption that those readers who are slim have not experienced prejudice from healthcare professionals, while those who are fat share this experience. Again, people – here, blog readers – are divided into two named groups, one consisting of insiders (*us*) and the other of outsiders (those *amongst us*) who need more information in order to understand the in-group.

While all bloggers in the data identified as plus-size, some of them also used *slim* in self-description when referring to particular body parts in contrast with the rest of one's body:

12) I'm **apple shaped with slim legs**.

Jeffries (2007:94) noted a contrast between the use of 'big' adjectives that premodified personal nouns and 'small' adjectives that were 'more likely to function as the object of a possessive verb such as *have*' (emphasis added) in women's magazines. For instance, people were described as 'bigger-chested' (pre-modifier), but as '[having] small breasts' (object of a possessive verb). According to Jeffries (ibid.), the former use 'indicates a more intrinsic feature of the body than the mere 'possession' of small breasts'. Something similar can be noted in Example 12 – *apple shaped* is a more prominent feature of the blogger than her possession of *slim legs*. A body part could be named as *slim* in contrast to other parts of the body, but the slimness of these individual body parts did not shape the identity of the blogger in the same way as fatness did.

5.3.3 The usage of *skinny*

The connotations of otherness were the most obvious in the references to *skinny* in the corpus. A skinny body was portrayed as something foreign and unattainable by the bloggers:

13) I always felt that the spa was a place for **the skinny and beautiful** and **my extra large backside** was not a welcome presence amongst them.

Naming a group of people that the writer cannot identify with as *the skinny and beautiful* links these two features together and creates an implication they go hand in hand (i.e. those who are not skinny cannot be beautiful). At the same time, the blogger reduces herself to a specific body part (Jeffries 2007:70); namely, her *extra large backside*.

The otherness of *skinny* is also present in the next example, where a blogger talks about ‘feeling skinny’:

> 14) I went home, despite looking heavily pregnant, **feeling skinny in a size medium**, haha.

In Example 14, the blogger acknowledges *skinny* as something that is not a reality for her, but rather a feeling associated with wearing oversized clothing. A similar way of speaking about *fat* is a common phenomenon in everyday English – it is not unusual for a straight-size person to state they are ‘feeling fat’ as a euphemism for feeling unattractive, uncomfortable, or bloated. In the same way as most people who use such language do not actually identify as fat, the blogger in Example 14 makes it clear she does not consider herself to be *skinny* by describing herself as *looking heavily pregnant*. Rather, using the word is her way of naming the feeling of fitting into fashion smaller than her usual clothing size.

Finally, *skinny* was also found in the context of insults directed at smaller women in the corpus. While this type of use was very uncommon overall, there were a few instances where *skinny* had offensive connotations:

> 15) I agree with you about **these skinny birds** saying how disgusting it is to be fat, but they should remember real men like meat it is only dogs that like bones. xx (comment in a blog post)

We can consider *these skinny birds* in Example 15 in comparison to *all my fabulous fats* in Example 6. According to Jeffries (2007:81), the use of demonstratives is significant – there is a ‘psychological distancing’ involved especially in the use of distal deictic terms, whereas possessive adjectives were used about more positive features of bodies in Jeffries’ data. In addition, it is notable that here the nickname *bird* appears to be connotated differently compared to ‘short, fat, old birds’ who ‘can be bootiful too’ in Example 3. The same nickname can be affectionate when used in reference to plus-size women, and distancing when used in reference to thinner women, who appear as the 'outsiders'[3] (see Pilcher 2016:771–772). It is also notable that the commenter in Example 15 refers to male approval as a defining factor for 'good' bodies. Again, while fatness is constructed as a quality that is not inherently negative, the commentary nevertheless maintains hegemonic discourses on gender, sexuality and size.

6. Conclusions

The UK-based fashion bloggers in the data constructed representations of the plus-size body both through reclaiming existing, traditionally negatively connotated terms and through innovating new terms. Both of these practices demonstrate that the plus-size fashion bloggers are aware of the agency they have in affecting their representation through language. The bloggers take advantage of the blogs as their own space, where they can challenge and dismantle discourses that traditionally construct the fat body as an inherently 'bad' body. They also utilize naming and describing the body for community-building purposes, by showing solidarity towards others who share their bodily characteristics, and including themselves in the same social group as their presumed readership based on the shared experience of living in a fat body. Thus, words that are used to describe the fat body have different connotative features in the blogs than they do in mainstream fashion media.

Although describing one's own body (or similar bodies) was more common in the data, highlighting the differences between particular groups of people through naming and describing was also significant in terms of identity construction. The analysis of how the bloggers named other people and bodies as *thin*, *slim*, or *skinny* revealed

[3] It is worth noting that the derogative use of skinny in Example 15 occurs in a comment from a blog reader, and while the commenter claims to state their agreement with the blogger, they are not quoting the original blogger word for word. The majority of the instances of derogative language use towards smaller women – or indeed, offensive language in general – appeared in the commenting section instead of the actual blog texts.

a division between us and them – the fat and the thin. While thin bodies could have positive connotations – for example, they could be beautiful bodies, or the bodies of friends and allies – their position as outsiders (those who are not part of *us*, but exist *among* us) was nevertheless present. In addition, there were more negative connotative features, such as the connection between being *too* thin and being unhealthy or unattractive.

As LeBesco (2001:78) has stated, the process of reclaiming fatness through language use is not without its challenges, because new discourses are still bound to construct hierarchies and boundaries between groups of people based on their bodily characteristics:

> However, I am aware that the process of gaining the upper hand, or reconstituting fat identity to change its current status as spoiled, will in turn produce its own subset of unthinkable, unliveable, and abject bodies. Subjects are constituted by the processes of excluding and abjecting, so it is necessary to reflect on how these processes shape fat identity.

The analysis presented in this study hopefully sheds some light onto such processes of shaping the fat identity. The plus-size fashion blogosphere, as illustrated by the language use of its members, is a space with great variety. While the use of *fat* in general appears frequent – specifically in its non-derogatory, reclaimed sense – we can also observe the presence of different sub-categories of fat bodies: *curvy, plus-size(d)*, and *apple-shaped*, *super-fats* and *barely-fats*, to name a few. Bloggers can feel ‘fabulous’ because of their fat bodies, or despite them. The fat body, while unifying the group through shared experience, is approached in different ways by the group members. This type of variety also inevitably leads to the exclusion of some identities in particular contexts, as LeBesco (2001:78) warned.

Finally, it is important to note that despite the potential that blogging has for producing counter-discourse and agency, there are still features in the blog texts that contribute to maintaining hegemonic discourses on bodies and beauty. Although the bloggers construct different representations of ‘right’ kind of bodies compared to Jeffries’

(2007) data from women's magazines, these ideas are not completely absent. While thinness is not valued in the same way it was in women's magazines, the language used in blogs still reflects the idea that it is valuable for women to be beautiful and desirable. Discussing clothing that is 'flattering' on a plus-size body and mentioning 'feeling skinny' when fitting into smaller clothing than usual, for example, illustrate that the mainstream beauty standards are not completely absent from the blogs. In addition, the connotative features of some of the creative descriptors for the body such as 'bootiful' contain a problematic assumption that a 'good' or 'acceptable' (female) body is a body that is sexually attractive (to men). Due to the presence of such discursive features, plus-size fashion bloggers and their ways of constructing the body still need to be examined critically, much like the body positive advertising campaigns that make claims of inclusivity and diversity, but often end up featuring models that only represent a very narrow idea of the plus-size body (Czerniawski 2021).

One central limitation of the present study pertains to the data. The original blog corpus was compiled in 2015, and since then, both blogs as a genre and the plus-size fashion industry have gone through significant changes. As Czerniawski (2021:3) points out, social media has changed the ways in which people consume fashion, and this has also resulted in new types of fashion-related content creation. Short-video format content produced on TikTok and Instagram Reels, for example, has become popular among body positive activists and plus-size fashion influencers. As these platforms facilitate viral content, they also provide the content creators with a broader and more diverse audience. All of these changes in content creation potentially affect the representation of marginalized bodies, which is why it is important to study the naming and describing of bodies and body parts in these new social media environments in the future. The present study on fashion blogs, as well as Jeffries' (2007) work on women's magazines, can function as useful starting points for such investigations.

References

The Association for Size Diversity and Health. 2020. https://asdah.org/health-at-every-size-haes-approach/ (accessed 25 April 2022).

Baker, Paul. 2006. *Using corpora in discourse analysis*. London: Continuum.

Brathwaite, Kyla N. & DeAndrea, David C. 2022. BoPopriation: How self promotion and corporate commodification can undermine the body positivity (BoPo) movement on Instagram. *Communication Monographs*, 89(1). 25–46, https://doi.org/10.1080/03637751.2021.1925939

Brookes, Gavin, & Baker, Paul. 2021. *Obesity in the news: Language and representation in the press*. Cambridge: Cambridge University Press.

Bucholtz, Mary, & Hall, Kira. 2005. Identity and interaction: A sociocultural linguistic approach. *Discourse Studies* 7(4–5). 585–614.

Butler, Judith. 1999. *Gender trouble: Feminism and the subversion of identity*. New York: Routledge.

Chibuwe, Albert, & Mpofu, Phillip, & Bhowa, Kudakwashe. 2021. Naming the ghost: Self-naming, pseudonyms, and identities of phantoms on Zimbabwean Twitter. *Social Media + Society*, July-September 2021. 1–12. https://doi.org/10.1177/20563051211035694

Connell, Catherine. 2013. Fashionable resistance: Queer “fa(t)shion” blogging as counterdiscourse. *Women’s Studies Quarterly* 41(1/2). 209–224.

Cooper, Charlotte. 2016. *Fat activism: A radical social movement*. Bristol: HammerOn Press.

Czerniawski, Amanda M. 2021. “Real” bodies in plus-size fashion. *Fat Studies*. https://doi.org/10.1080/21604851.2021.1913827

Dickins, Marissa & Thomas, Samantha L. & King, Bri & Lewis, Sophie & Holland, Kate. 2011. The role of the fatosphere in fat adults’ responses to obesity stigma: A model of empowerment without a focus on weight loss. *Qualitative Health Research* 21(12). 1679–1691.

Goffman, Erving. 1990. *The presentation of self in everyday life*. London: Penguin.

Grad, Hector, & Martín Rojo, Luisa. 2008. Identities in discourse: An integrative view. In Dólon, Rosana, & Todolí, Júlia (eds.), *Analysing identities in discourse*, 3–28. Amsterdam: John Benjamins.

Gumperz, John J., & Cook-Gumperz, Jenny. 1982. Introduction: Language and the communication of social identity. In Gumperz, John J. (ed.), *Language and social identity*, 1–21. Cambridge: Cambridge University Press.

Gurrieri, Lauren, & Hélène Cherrier. 2013. Queering beauty: fatshionistas in the fatosphere. *Qualitative Market Research: An International Journal* 16(3). 276–295.

Harju, Anu, & Huovinen, Annamari. 2015. Fashionably voluptuous: normative femininity and resistant performative tactics in fatshion blogs. *Journal of Marketing Management* 31(15–16). 1602–1625.

Hartley, Cecilia. 2001. Letting ourselves go: Making room for the fat body in feminist scholarship. In Braziel, Jana Evans & LeBesco, Katherine (eds.), *Bodies out of bounds: Fatness and transgression*, 60–73. Berkeley: University of California Press.

Ho, Petula Sik Ying, & Tsang, Adolf Ka Tat. 2005. Beyond the vagina-clitoris debate – From naming the genitals to reclaiming the woman's body. *Women's Studies International Forum* 28. 523–534.

Jeffries, Lesley. 2007. *Textual construction of the female body: A critical discourse approach*. London: Palgrave Macmillan.

LeBesco, Katherine. 2001. Queering fat bodies/politics. In Braziel, Jana Evans & LeBesco, Katherine, (eds.), *Bodies out of bounds: Fatness and transgression*, 74–87. Berkeley: University of California Press.

LeBesco, Katherine. 2004. *Revolting bodies: The struggle to redefine fat identity*. University of Massachusetts Press.

Limatius, Hanna. 2020. *Communities of Empowerment: Exploring the Discourse Practices of Plus-Size Fashion Bloggers*. Tampere: PunaMusta Oy.

Limatius, Hanna. 2019. 'I'm a fat bird and I just don't care': A corpus-based analysis of body descriptors in plus-size fashion blogs. *Discourse, Context & Media* 31. https://doi.org/10.1016/j.dcm.2019.100316

Limatius, Hanna. 2017. 'There really is nothing like pouring your heart out to a fellow fat chick': Constructing a body positive blogger identity in plus-size fashion blogs. *Token: A Journal of English Linguistics* 6. 23–49.

Pilcher, Jane. 2016. Names, bodies and identities. *Sociology* 50(4). 764–779.

Sastre, Alexandra. 2014. Towards a radical body positive. *Feminist Media Studies* 14(6). 929–943. https://doi.org/10.1080/14680777.2014.883420

Scaraboto, Daiane, & Fischer, Eileen. 2013. Frustrated fatshionistas: An institutional theory perspective on consumer quests for greater choice in mainstream markets. *Journal of Consumer Research* 39(6). 1234–1257.

Tajfel, Henri. (ed.). 1982. *Social identity and intergroup relations*. Cambridge: Cambridge University Press.

Zavattaro, Staci M. 2021. Taking the social justice fight to the cloud: Social media and body positivity. *Public Integrity* 23. 281–295. https://doi.org/10.1080/10999922.2020.1782104

Nordisk tidskrift för socioonomastik / Nordic Journal of Socio-Onomastics (97–130)

Lunatics and idiots: The effect of semantic shifts on naming and labelling in the British Parliament, 1950–2000

Minna Nevala & Jukka Tyrkkö

Abstract: This article studies the sense development, semantic shifts and use of words referring to people with mental illness in public discourse in the latter half of the 20th century. The focus is on the process of labelling or naming, which often reflects a more prevalent, societal attitude either in favor or against particular group memberships. The results show that while the old terms underwent a semantic change around the Second World War, their use continued in the latter half of the 20th century. Terms such as *lunatic* and *idiot* were used as intensifiers, for comedic purposes, and as distancing devices in intergroup relations and were no longer referential to medically diagnosed mental health conditions.

Keywords: people with mental illness, labelling, parliamentary discourse, group membership, social identity, representation

Minna Nevala (University of Helsinki) & Jukka Tyrkkö (Linnaeus University). Lunatics and idiots: The effect of semantic shifts on naming and labelling in the British Parliament, 1950–2000.

NoSo 2023 | https://doi.org/10.59589/noso.32023.14413

1. Introduction

1) In section sixty-seven, for the words **'lunatic or idiot, fatuous or furious person'** there shall be substituted the words '**person suffering from mental disorder** within the meaning of the Mental Health (Scotland) Act, 1960'. (Lord Craigton, 12 July 1960)[1]

The above extract (1) is from a parliamentary discussion about the Mental Health Act of 1960 which was designed as the first comprehensive revision of Scottish mental health law in over 100 years. The Bill was to establish, among other things, the Mental Welfare Commission for Scotland, which was to serve as an independent central body 'with the duty of exercising protective functions in respect of mentally disordered persons' (Mental Health Bill, HL Deb 5 July 1960 vol. 224 cc1011–26). The extract shows an example of how the previously used terms *lunatic* and *idiot* were to be changed into a more medically based term *person suffering from [a] mental disorder.*

The history of using terms such as *lunatic* and *idiot* goes far back. As early as the Victorian era, the English were fascinated by social and moral degradation, and digging up 'the dirt' on others (O'Reilly 2014). As with prostitutes, criminals and the sexually deviant, people with mental illness were classified as a 'filthy' part of society and a problem to be dealt with. At the time, the public use of derogatory terms was widely accepted (Houston 2000), visiting mental asylums like Bedlam was considered public entertainment (Arnold 2008), and people could be locked up in asylums on the request of legal guardians without any sort of medical consultation.

Over the decades that followed, both the legal and public perspectives on mental health issues progressed: psychiatric challenges became normalized as medical matters, healthcare standards improved, and the rights of patients and sufferers were codified in law. Along the way, the language of public mental health discourse

[1] In all examples, direct labelling is marked in bold.

underwent various stages of development, with derogatory and dismissive terms gradually being replaced by medical terminology, and the focus shifting from people with mental illness as an anomaly to be expelled from civil society to being a personal and private challenge that requires and deserves help from society.

In this study we discuss the sense development, semantic shifts and use in public discourses of words referring to people with mental illness during the latter half of the 20th century (from 1940 to the present day). We identify the relevant lexical items using the *Historical Thesaurus of the Oxford English Dictionary* and then investigate their use both quantitatively and qualitatively in parliamentary debates, as represented by the Hansard Corpus (Alexander, forthcoming). The quantitative trends are further contrasted with contemporaneous data drawn from Google Books (British), which allows us to examine the differences of lexical richness and frequency developments within the sense-family of *mental health* in the two registers (see Nevala & Tyrkkö forthcoming). By examining the complete lexical field together, we are able to identify turning points in the public discussion and dig deeper into the linguistic processes involved. One of these processes that we focus on is labelling or naming, which often reflects a more prevalent, societal attitude either in favor of or against particular group memberships. The different definitions and descriptions related to negatively evaluated groups like this are seen as a basis for creating and spreading public stereotypes (cf. Hintikka & Nevala 2017; Nevala 2019).

The study is structured as follows, we start by charting the sociocultural background for labelling people with mental illness in 1950–2000, as well as defining what we mean by the language and the representation of deviance. We then present our data and method. We provide examples of our data before presenting the results, followed by our conclusions.

2. Labelling people with mental illness

Labelling, naming and name-calling all belong to the general field of *socio-onomastics*, that is, the study of how speakers assign names to others (see e.g. Ainiala & Östman eds. 2017). Depending on the circumstances, the relationship between the individuals involved, and the intended and perceived purpose of the act, the act of associating a new lexical sign to someone can be a positive or a negative activity. A distinction can be made between *labelling*, where a person is typically assigned into a category or group of some kind (*British*, *patient*, *queer*), and *naming*, where a person is given a byname that is intended to either replace or supplement their official name (*Big Jimmy*, *Crooked Hillary*). Both labelling and naming are necessary linguistic acts, but they also have the potential of being used in a hurtful or derogatory manner (see, e.g. Tyrkkö & Frisk 2020). In particular, when labelling is used in a way that hides a person's identity and replaces it with membership in a group (such as nouns referring to race, ethnicity, sexuality, or gender identity), the person's autonomy and subjecthood are challenged and they are typically stigmatised as the 'other'. The 'verdictive force' of nicknames and labels, as described by Adams (2009:84), thus making them a powerful means of harnessing social control over others (see also Adams 2008; Croom 2013).

Medical conditions and illnesses are considered sensitive subjects in most societies, with psychiatric conditions perhaps being the most stigmatized, as they are often seen as depriving a person of their social standing and independence. Consequently, the way the public label someone as a person with a mental condition is usually considered offensive and derogatory, and various linguistic strategies have been proposed as a way of mitigating such effects. As an example, Price (2019:3) argues that 'the [British] press use identity-first forms (identified as stigmatizing by mental health advocates) to refer to people with mental illness (e.g. 'a schizophrenic') more often than person-first forms (such as 'a person with schizophrenia')'.

In the present study, we focus on labelling with common nouns that refer to individuals with a mental health condition. More specifically, we focus on nouns such as *lunatic* or *idiot*, which were until the

NoSo 2023 | https://doi.org/10.59589/noso.32023.14413

early decades of the 20th century entirely neutral and acceptable (see Section 3). However, these terms have since undergone a semantic *substitution* (also *narrowing*, or *restriction*) that renders them unfit to use in reference to individuals with mental health conditions (Stern 1931). More specifically, this type of change can be described as a case of *taboo-induced replacement* (Hock & Joseph 2009:220–221), where words that become culturally inappropriate are replaced by others, and the primary referents of the original words may undergo a change of some kind.

2.1 A brief look into mental health in British society

Societies have long aimed at restricting the activities and rights of people with mental illness. Establishing clear lines between ‘the sane’ and ‘the insane’ has also meant seeking to legitimize the institutionalizing of people with mental illness. From the eighteenth century onwards, ‘mentally ill’ was used as a label that covered a wide range of real and supposed conditions from low intelligence and personality disorders to psychiatric illnesses. In Britain, around 5,000 people were held in specialized lunatic asylums before the year 1800, and it has been estimated that as many people with mental illness occupied various workhouses, bridewells and jails (Porter 2002:95; Scull, MacKenzie & Hervey 2014). In 19th century England, patient numbers climbed from approximately 10,000 to ten times the amount by the year 1900 (Porter 2002:112).

It was not, however, those who were clinically attested as mentally ill who were institutionalized – virtually all socially marginalized were considered mentally ill to some extent. The bulk of this group, such as paupers, criminals and streetwalkers, were seen as not only dangerous and monstrous but also as insane and idiotic (Porter 2002:92). Furthermore, women with their madness or hysteria, or ethnic or other minorities, were often described as ill and lacking reason (see e.g. Froelich 2016 on early modern women; Potts & Weare 2018 on modern women). Those imprisoned for homosexual acts would sometimes claim mental illness in order to be released, but unless the person was then released into the care of a private physician, there was a danger of being confined to an asylum instead (Janes 2014).

Nineteenth-century British society often saw people with mental illness as ‘filthy’, which was reflected in the widely accepted use of derogatory terms in public discourses. Although mental illnesses were increasingly recognized as inherently medical issues, those that had them were also seen as deviant and perverse ‘monsters’ who constituted a threat, therefore needing to be locked up and kept away from respectable and proper citizens (Nevala 2019; see also Potts & Weare 2018 for the 21st-century use of the term *monster*).

Many mental institutions were known for their cruelty and corruption, and the idea behind the madhouses was to brutally tame those with mental illness through physical restraint, bloodletting, ice cold baths, purges, and vomits (Porter 2002:100). Visiting notorious mental asylums like Bedlam (Bethlem Royal Hospital in London) was considered public entertainment, and the well-to-do flocked in to see the psychotic mad. Gradually, over the course of the nineteenth century, practical psychiatry was transformed into a tool for restoring people with mental illness to health, instead of merely keeping them away from respectable society.

Both World Wars seem to have been a decisive turning point when it came to changes in general attitude to mental illnesses, with the prevailing discourses casting mental conditions more as personal challenge than as anomalies. The change can be observed in the overall lexis of mental health, and perhaps particularly so when it comes to the words that were acceptable to use in reference to individuals with mental conditions (Nevala & Tyrkkö forthcoming).

By the 1950s, developments in psychiatric care had permanently moved away from thinking of mental illness as something characteristic of socially marginal groups. The general attitude also changed in terms of patient autonomy. In the Victorian era, the medical staff treated mental patients essentially as objects, i.e. rarely listening to their requests and barely addressing or treating them as human beings (Scull, MacKenzie & Hervey 2014:30–31; see also Suzuki 1999). Patients in asylums had no say in what treatment was issued, a situa-

 NoSo 2023 | https://doi.org/10.59589/noso.32023.14413

tion that persisted into the early 20th century; electric shock treatment was introduced to treat, for example, schizophrenia in the 1920s and 1930s. The development of psychiatric medicine in the early 1950s gave people with mental illness more agency, enabling many patients to maintain a 'normal' life with medication.

In many present-day societies, the segregation of 'the insane' from 'the sane' is no longer seen as a straightforward norm, and modern psychiatry is opting more for integration instead of segregation. Nevertheless, attitudes toward people with mental illness, particularly as portrayed in the media, still show remnants of the past negative discourses in the form of stereotyping (Birch 2011:84). For instance, Balfour (2019) found that people with schizophrenia are often represented in the British press as enacting violent crime. This seems to be a part of a broader tendency of representing people with mental illness in a contradictory way as both insane and morally culpable (Balfour 2020:538; Cross 2014). By representing people with mental illness as criminals, newspapers suggest to their readers that they are morally responsible for their crimes, and therefore blameworthy. However, people officially diagnosed with mental illnesses tend to receive reduced sentences on the grounds of diminished responsibility. In other words, they can be accused of the crime, yet not be held responsible for their actions.

2.2 The representation of deviance: Labels, names and social stereotypes

We have chosen a fairly broad definition of deviance from mental health norms. We include, firstly, such individuals that were seen by 'respectable society' as being disruptive, immoral, or dangerous in the eyes of the law (criminal mental illness), and secondly, those who were less or not so (medical state of mental illness). In general, the line between 'decent' society and the collective outcasts has, of course, been historically fluid. While there are certain common denominators between people with mental illness across the diachrony, there are also groups within those with mental illness whose marginalization is and has been time or culture specific.

In discourse analysis, moral evaluation is 'linked to specific discourses of moral value' (van Leeuwen 2008:110). Representation of a particular individual or a group can manifest itself by the use of adjectives such as *healthy*, *normal*, or *natural*. This is what van Leeuwen (2008:109) calls 'moral legitimization' – it consists of the processes of evaluation, abstraction, and comparison. People are categorized on the basis of positive and negative values into different sociocultural groups. As Van Dijk (2009:141) states, giving attributes to the self and others concerns the interactional and societal context. This means that defining is not only governed by macro-level norms or shared knowledge, but it is also produced in micro-level interactions and situations. According to Gumperz and Cook-Gumperz (1982:3), different ideologies people support or maintain enter into face-to-face interaction and discourse practices where 'subconscious and automatic sociolinguistic processes of interpretation and inference' can lead to different outcomes. No interaction is thus value-free, despite it always being assessed according to an individual's norms and values. As Ochs (1993:289) understands it, social identity is usually something not explicitly encoded in language use, but rather a social meaning inferred in act and stance meanings. Social representation can be seen to evolve and vary in social interaction in response to the acts and stances of other interlocutors, but also according to the speaker's own attitude towards each interactional situation (Ochs 1993:298).

Often the values and norms of a particular group are also manifested in the negative labelling of other groups or their members, often by means of creating and maintaining negative impressions using 'labels of primary potency' (Allport 1986). This means that certain categorical and often simplistically binary characteristics, such as male/female, heterosexual/homosexual, or normal mental health/mentally ill, carry more perceptual potency than others, and signal a difference from what is considered mainstream (e.g. moral distinctiveness). The division between what could be called neutral, positive, and negative labelling and naming reflects a more prevalent, societal attitude either in favor of or against particular group memberships, as in the case of people with a mental illness versus those without one. Thus, for example, 'sane' people can relegate others with mental illnesses to an out-group by creating and using negative terms and attributes. Studies have shown that specific naming strategies exist for social outsiders such as criminals (Mayr & Machin 2012:57). Clark (1992:224) calls this process of extreme negative labelling *fiend naming*. When criminals and criminally insane people are referred to as monsters, they are depicted as being so evil and alien that they have to be placed outside decent society as abnormalities. Negative labelling then becomes a strategy based on the notion that no 'normal' person would be capable of such 'monstrous' behavior. As mental conditions are typically associated with unorthodox non-normative behaviors, references to mental conditions are an equally important means of signalling that a person or entity is strange or 'not normal' and therefore need not be taken into account when decisions are taken, or plans made.

In comparison, 'normal' people are often labelled positively with what could be called *angel naming*, respectively (cf. Nevala 2016 for the representation of crime victims). For example, this effect could be achieved by using the adjective *poor* and an age-specific noun alone, such as 'the poor girl', or by highlighting some positive quality of a person when that quality has no relevance to the situation at hand. In the British parliamentary data (see Nevala & Tyrkkö forthcoming), we have observed a strong diachronic trend towards more humane treatment of people with mental health conditions, with increasing voices calling for treatment in medical, rather than penal, institutions. For example, in late 19th-century parliamentary discussions, there already was a clear distinction between the terms *pauper lunatic criminal* and *pauper lunatic* (see Section 4.2 for further discussion of *lunatic*).

What is considered a mental disorder or illness has in the past been stereotypically labelled with attributes associated with emotionality, unreliability, hysteria, deviousness etc. Potts and Weare (2018:43) discuss critical labelling of women with mental illnesses and note that in many court cases involving criminals with mental illnesses, certain labels are used that actually index cultural (mis)understandings about mental health disorders. One such label is the colloquial term *mad-woman*, which has been used to cover any medicalized term, such as 'psychopathic disorder' or 'antisocial personality disorder' (see also Froelich 2016 for the use of *madwoman* in Early Modern English). In our data, *mad(man)* is also used, alongside with words that entered vocabulary between 1870 and 1950, such as *crazy, crackpot, looney, deficient, defective, dotty, barmy, natural, nut-house* and *bin* (see Nevala & Tyrkkö, forthcoming). Notably, when a generic noun such as *lunatic* or *madwoman* is used for labelling someone as being mentally ill, the underlying message is that it is not important to be specific, that all people with a mental condition belong to the same group and it is natural to think of them as a collective, rather than as individuals.

3. Materials and methods

The present study analyses public discourse in the British House of Commons, with particular focus on a set of nouns that underwent a semantic shift during the early decades of the 20th century, from neutral labels for people with mental illness to offensive slurs or ostensibly comedic references. We are interested in discovering how rapidly this shift took place, whether both the old and the new meanings co-existed in the discursive space of the parliamentary floor, and what pragmatic meanings are given to the terms.

Although political language is the object of study, we would not define the approach taken as critical discourse analysis, as our aim is not to uncover expressly political or otherwise opaque motivations behind the use of these terms – even if the terms may have been used to describe political opponents. Consequently, instead of focusing on the 'critical', we see the study more from discourse-historical and lexico-semantic perspectives, and argue that the semantic shifts are evidence of a wider sociocultural change that discourse analytical methods can help bring into focus. As the primary data come from two very large corpora, the study can be described as belonging to the field of corpus-assisted discourse studies (CADS) (see e.g. Partington 2010 and Partington et al. 2013). The comparison of quantitative findings from the multi-billion-word Google Books corpus and the Hansard Corpus allows us to both contextualize the latter in light of the former, and to address the oft-cited shortcomings of purely qualitative discourse analysis, namely anecdotal evidence and cherry-picked data. Similarly, the detailed qualitative analysis of the Hansard data from the latter half of the twentieth century ensures the validation of the interpretations given to the observed trends.

3.1 The parliamentary record and the Hansard Corpus

The records of the British Parliament, conventionally called the Hansard, have been public documents since the early 19th century (see Vice & Farrell 2017 and Alexander forthcoming, for the early history of the Hansard and the linguistic implications thereof). The present study focuses on data from the House of Commons.

An important distinction needs to be made between the earlier Hansard, up to and including the year 1910, and the later Hansard. In its earlier years, the Hansard relied on journalists' notes and other written material, while from 1910 the Hansard has been based on transcriptions by the Parliament's own transcribers. Consequently, it is reasonable to state that the text type of the Hansard changes around 1910, going from a combination of reported speech and narrative description, with substantial proportions of official documents mixed in, to transcriptions of debates, described by the Parliament's own official descriptions as a 'substantially verbatim' record of what was actually said in the House (see Hiltunen et al. 2019). It is well known that the transcribers quietly correct features of spoken language that are deemed irrelevant, such as repetitions, false starts, and so on, as well as utterances that go against the official rules of 'unparliamentary' language, as defined in the Erskine May, the official description of parliamentary conventions named after the author of the first edition, Thomas Erskine May. The reliability of the Hansard as a record and the linguistic implications thereof have been discussed by several scholars including Slembrouck (1992), Mollin (2007), and Alexander (forthcoming), with the general consensus view being that while some spoken-like features may not be entirely accurately recorded, the Hansard is otherwise a reliable dataset that allows large-scale analyses of the topics and discursive practices of the British Parliament (see e.g. Tyrkkö 2019; 2020).

The Hansard record was compiled into a linguistic corpus by the SAMUELS project at the University of Glasgow. The 1.6-billion-word *Hansard corpus* is currently available via both the English Corpora website (http://www.hansard-corpus.org) and the Hansard at Huddersfield website (https://hansard.hud.ac.uk/site/index.php). The authors used an XML-annotated version of the corpus provided by colleagues in Glasgow (for more details, see Nevala & Tyrkkö forthcoming).[2] Queries were run for the seven key terms (see Section 3.3.) and the standardized frequencies were calculated as normal. It is worth noting that unlike most corpora, which are samples drawn from and representative of a population, the Hansard is a complete record of the parliamentary debates and thus, statistically speaking, the corpus is the population rather than a sample.

3.2 Google Books

Google Books is a massive archive of digitized books compiled and made available by Google. It can be easily accessed via the Google Books service, which allows varying levels of preview access to individual books. It can also be accessed via the Google Ngram viewer, which provides quantitative data on word and phrase frequencies in the various language and/or country-specific iterations of the archive; in the case of English, there are separate datasets for British and American English. The latter data can also be downloaded or accessed using the Google Ngram viewer API (Application Programming Interface), although this is often technically challenging

[2] The authors are grateful to Marc Alexander and Fraser Dallachy at the University of Glasgow for access to the corpus.

 NoSo 2023 | https://doi.org/10.59589/noso.32023.14413

due to the amount of data involved. Finally, it can be accessed using the Google Ngram viewer API (Application Programming Interface), which allows local scripts to make calls to a server and to receive the requested data in return. The reliability and representativeness of Google Books has been discussed widely in academia since at least Michel et al. (2011; cf. Laitinen & Säily 2018). It is our understanding that the general consensus is that despite its shortcomings, Google Books is a highly valuable resource when it comes to forming overall 'culturomics' style analysis of lexical prominence over time.

In the present study, we used the 32-billion-word Google Books British (v. 2019) collection and a custom script for making the API calls.[3] Data provided by Google on the overall unigram counts for each year were used to calculate the standardized frequencies of the seven terms of interest (see Section 3.3). Due to the amount of data and the lack of direct access to the concordance lines, we relied on the overall frequencies of the query terms without discerning between word classes or different uses of the terms.

3.3 Repertoire of relevant items and principles of classification

The lexis of mental health is extensive in English, and it goes without saying that it would be impossible to carry out a CADS-style mixed-methods study of the full repertoire of terms available at any given time, let alone over a timespan of two hundred years. For the present study, we focused purely on the terms of interest by drawing on a previous study by Nevala & Tyrkkö (forthcoming), in which we used the *Historical Thesaurus of English* to identify all known synonyms of several fields of mental health terminology and carried out a quantitative canvassing of the full range of terms in the Hansard. That investigation showed that relatively few terms were used with any regularity, and the majority either very rarely, or not at all. In the parliamentary context, this can be explained by the combination of a

[3] The tool was written in Livecode 9.6.5. For more details, see Tyrkkö & Mäkinen (2022).

somewhat restricted range of topics, the nature of the debates, and to some extent the rules of conduct governing those debates, which proscribe against name-calling.

In the present study, our main research question concerns the use of terms that were once used in reference to people with mental health conditions, but during the 20th century turned into derogatory or colloquial expressions. By examining the complete lexical field, we have identified turning points in public discourse and explore the linguistic processes involved. One of these processes that we are particularly interested in is labelling or naming, which often reflects a more prevalent, societal attitude either in favor of or against particular group memberships. This focus limited the range of possible lexical items, leading us to select seven terms for quantitative analysis, and to further examine four terms in the qualitative study.

The quantitative part of the study compares standardized frequencies of the relevant word forms observed in Google Books Britain and the Hansard, focusing primarily on showing the overall trends. In the qualitative part, we first carry out a collocation analysis of the four terms to gain an understanding of their distributional semantics. We then discuss the terms' usage with manually selected examples identified through close reading of concordance lines.

4. Findings

The findings are presented in two parts. We begin by discussing the overall frequency changes in the use of the seven terms of interest over the 19th and the 20th centuries, to highlight changing lexical practices over time. While focusing on the Hansard, we start by presenting contrastive data from Google Books in order to show the extent to which parliamentary discourse differs from generic language use. Second, we focus on the latter half of the 20th century in parliamentary language and provide a qualitative analysis of the semantic shift that affected the terms in question and the pragmatics of using them in parliamentary debates.

4.1 Trajectories of lexical frequency

We begin the analysis with a look at what happened to the lexical frequencies of the seven terms over the last two-hundred years in both general English written texts, as represented by Google Books British (Figure 1), and in parliamentary language, as represented in the Hansard Corpus (Figure 2). The axes in the two figures are scaled similarly for clarity.

Beginning with the data from Google Books, it appears that the relative usage of the seven terms has remained largely stable for most of the timeline, with a general decline from the 1930s onward. *Lunatic*, *idiot* and *madman* have consistently been the most common terms, followed by *fanatic, maniac*, and *imbecile*. *Moron*, by contrast, has been very rare and only really in use in the 20th century; consequently, due to the relatively much lower frequency of *moron*, the item cannot be seen in the visualization.

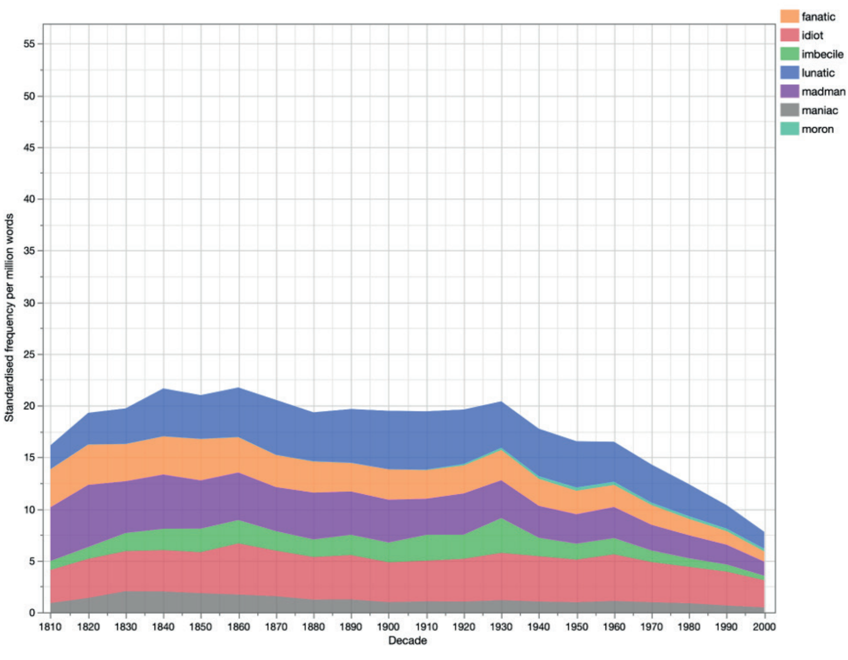

Figure 1: Trajectories of lexical frequency in Google Books British.

By contrast, the use of the seven terms has been much more varied in parliamentary discourse, owing largely to the oscillating nature of legislative topics, as well as the fact that for much of the timeline, the terms under investigation were considered either medical or generic terms for people with mental illness. Consequently, at times when mental health issues and the care of psychiatric patients were discussed, the frequency of references rose sharply, while at other times the frequencies dropped in equal measure. *Lunatic*, which was the generic term for people with mental illness, has been by far the most frequent term from the 1830s onwards, with *fanatic*, *idiot*, and *imbecile* seeing infrequent but steady use. *Madman* and *maniac*, two fairly common terms in generic English, have rarely been used in parliamentary discourse.

The extreme frequency of *lunatic* during the 19th century are explained by the continued debates in Parliament on how the care and containment of persons with mental illnesses ought to be arranged in society. As discussed in Nevala & Tyrkkö (forthcoming), the term *lunatic* was not only frequent in institutional contexts such as *lunatic asylum*, but also used as a generic term that could be used as a collective term in reference to a wide range of different sufferers. Premodifiers such as *criminal* and *pauper* were also frequently used with *lunatic* when referring to specific subgroups of mental patients; *criminal lunatics* were violent offenders who needed to be incarcerated, while *pauper lunatics* were mental patients whose care could not be paid for by themselves or their families.

 NoSo 2023 | https://doi.org/10.59589/noso.32023.14413

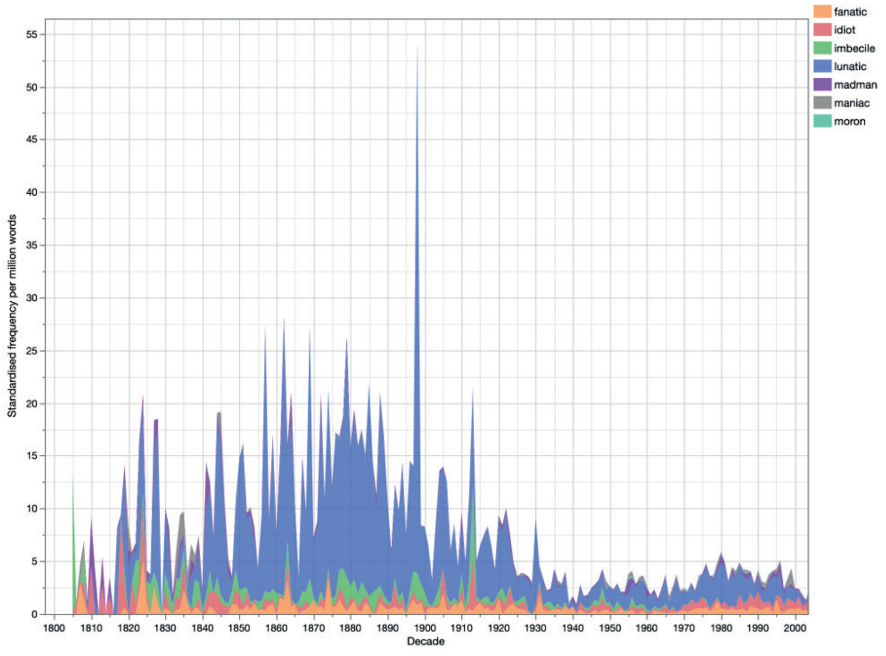

Figure 2: Trajectories of lexical frequency in the Hansard Corpus.

A notable difference between the two datasets is that while there was a clear overall reduction in the use of *lunatic* in particular, but also of the other terms, in parliamentary language since the 1940s, the decline in the use of the same terms was much more gradual in general English. The overall aggregate frequency of the seven terms remained at a plateau of c. 2.5–5 hits per million words throughout the 20th century in parliamentary discourse, while their frequency in the Google Books data decreased from approximately 20 to 7.5 hits per million words in the same period. This suggests that attitudes to the terms changed at different rates in Parliament than in general written English. When we visualize the rates of change by standardizing the frequencies for each term and then comparing the diachronic plots, we can see that the developments look somewhat similar (Figure 3).[4]

[4] Standard scores of z-scores are calculated from the mean and sample standard deviation. Standardization allows us to observe relative trends that are easily obfuscated by the magnitudinal differences.

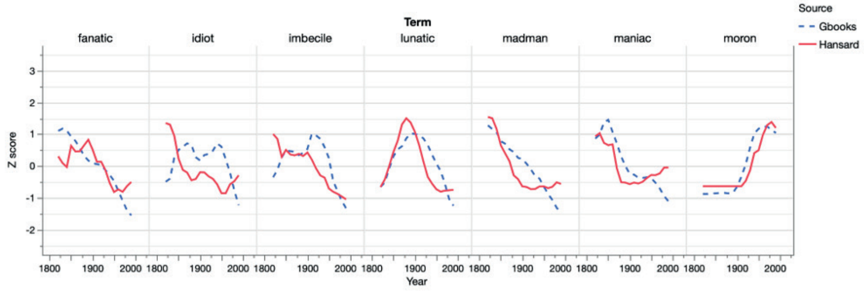

Figure 3: Comparison of diachronic trends over time by z-score.

Correlating the decade means we can see that only *lunatic*, *madman* and *moron* show a statistically significant correlation between the Hansard and Google Books, hence parliamentary language and general written British English followed a somewhat similar developmental pattern regarding these three terms.[5] In contrast, *idiot*, *imbecile*, *fanatic* and *maniac* do not show a significant correlation between the two populations (Table 1).

As discussed in Nevala & Tyrkkö (forthcoming), the 1930s appear to have been a turning point for societal attitudes to mental health. As modern psychiatric care developed and gained ground, a lexical shift took place that rendered the old terms archaic and unacceptable in polite use. Notably, while such shifts may take a considerable time in general language use, they may be observed much more swiftly in context-governed discourses (such as parliamentary debates) where the speakers are experienced in the rhetorical use of language and concerned about the impressions they make on the audience. Perhaps of equal importance, when terms disappear from formal use in official documents, titles of bills, and names of institutions, the need to use them referentially declines almost overnight.

[5] R^2, or the coefficient of determination, indicates the proportion of the variance in the dependent variable (here, the Hansard) that is explained by the independent variables (here, Google Books). The F test is a statistic that indicates the ratio of explained variation and unexplained variation, and the p-value, calculated from the critical F-value, indicates whether we may reject the null hypothesis that data are not linear.

Table 1. Regression statistics for each term between the Hansard and Google Books Britain.

Term	R2	F-value	p-value
lunatic	0.419	12.98	**
idiot	0.154	3.28	ns
imbecile	0.03	0.56	ns
madman	0.495	17.6	***
fanatic	0.057	1.1	ns
moron	0.676	37.6	***
maniac	0.125	2.58	ns

(*** = p<0.001; ns = non-significant)

However, as Figure 2 also shows, while there was a substantial reduction in the use of these terms in Parliament from the 1930s onward, their use did not end entirely. Given that none of the seven terms have been considered inoffensive in the context of psychiatric discourse for well over half a century, the question is what situations give rise to their use in the parliamentary context? To answer the question, in Section 4.2 we turn to a qualitative analysis of parliamentary language during the period 1950–2000.

4.2 Lunatics to maniacs in the post-war period 1950–2000

In this section, we will concentrate on four specific terms used for people with mental illness in parliamentary discussions and debates in the post-war period, namely, *lunatic, idiot*, *madman* and *maniac*. As stated in the quantitative analysis (Section 4.1), all four terms have been among the most commonly used terms throughout the period being studied in the data from Google Books, but used more variably in the data from the Hansard. For example, *lunatic* and *idiot* have appeared in steady use, whereas *madman* and *maniac* have rarely been used in parliamentary discourse. All four terms have been more prominent in general English from the 1930s onwards, and are studied here in their micro-level context in order to show what kind of connotations and meanings they have developed throughout their years in use in the latter part of the 20th century.

We begin the qualitative analysis by identifying the most statistically significant collocates of the four words to establish an overall sense of their use in context. The collocates were derived from the BYU English Corpora interface to the Hansard corpus (https://www.english-corpora.org/hansard/). Using a window of three words to the left and right of each keyword, we extracted the fifty collocates with the highest 'mutual information (MI) score' (an association strength measure widely used in corpus linguistics) and then visualized the shared and unique collocates of each term as a network graph using the open source network analysis tool Cytoscape (https://cytoscape.org) (Figure 4). The MI score is used as edge weight in an edge-weighted spring-embedded layout (a force-directed network paradigm that positions the nodes by minimizing the sum of forces in the network, see Kamada & Kawai 1989). Collocates with a negative connotation are marked in red.

While the technical sense of *lunatic* (in relation to parliamentary discussions on laws of lunacy, etc.) started to disappear during the first decades of the 20th century, the term persisted in colloquial use to refer to someone not fully in control of their mental faculties. The collocates show that references to the old meaning persisted to some extent as well, with the collocate list featuring terms such as *prisons*, *Broadmoor*, *asylums*, and *criminal*. However, references to lunatics were increasingly made as expressions of exasperation over someone else's lack of judgement or inability for rational behaviour. The term in itself seems to have been used without much connotation to violence, hence premodifiers such as *criminal* or *violent* that were frequently used with *lunatic* when referring to specific subgroups of people with mental illness.

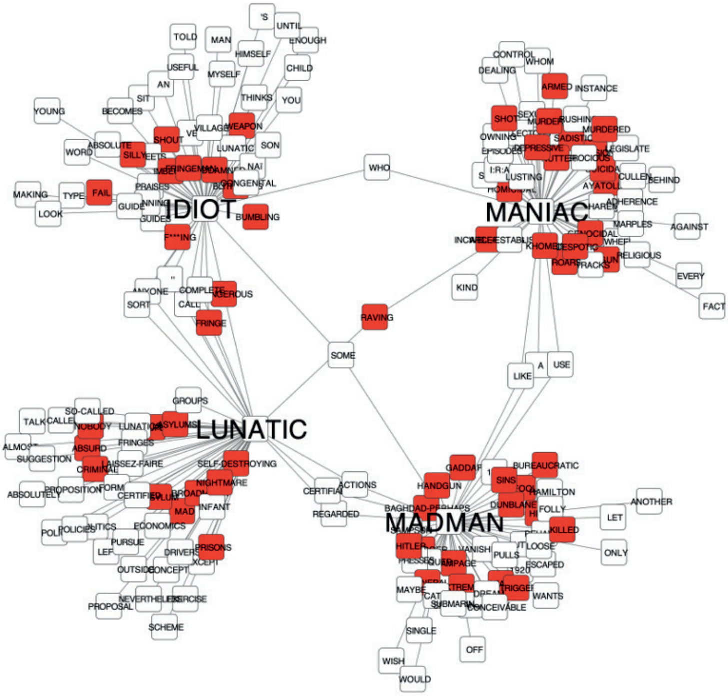

Figure 4: Network graph of shared and unique collocates of *idiot*, *lunatic*, *madman* and *maniac*.

Example (2) is an excerpt from Mr Nicholson's contribution to a 1953 debate on Anglo-American relations, where he uses *lunatic* as a general term for a person who asks irrational questions, someone less dangerous than 'a Russian agent'.

2) May I ask whether the hon. Member who asked the Question is a Russian agent or merely **a lunatic**?
(Mr Nicholson, 25 June 1953)[6]

This semantic shift continued for the remainder of the century. As with the previous example, (3) the word *lunatic* is used not in refer-

[6] In all examples, direct labelling is marked in bold and contextual evidence is underlined.

ence to psychiatric illness but as a characteristic that can be applied as a label to people, situations and acts alike; the collocate list includes terms such as *policies*, *proposal*, *scheme*, and *politics*. The speaker, Mr Sproat repeats the term, labelling an entire chain of events as lunatic: the Lord President is illogical, which leads to 'a lunatic premise', which in turn leads to 'these lunatic proposals', which ends up in 'everything thereafter [being] lunatic' or '[the kind of] lunacy'. Being a lunatic is once more equalled with being illogical, not rational. Here, the use of lunatic also shows a juxtaposition between 'lunatic people' making foolish decisions and 'rational people' wanting to make wise decisions.

> 3) I was about to deal with it in one crisp sentence. That is the kind of **lunacy** into which we shall be led by **these lunatic proposals**. If the Lord President <u>never was logical</u>, if he starts from **a lunatic premise**, everything thereafter will be **lunatic**. That is what is happening with the Bill.
> (Mr Sproat, 1 February 1977)

Whereas in the previous example the use of *lunatic* was targeted to a specific person or group, the example in (4), taken from the debate on the Firearms Amendment Bill, shows the use of the term as a label for any person who wants to buy a gun and is able to do so. Mental health is not validated in any way, medically or otherwise, therefore the premise for being a lunatic here is insanity and violent behavior. The speaker, Lord Campbell, refers to the Bill being 'flawed' and 'bad', meaning that if the Bill is not good, then a particular group of people – lunatics owning a gun – are placed outside the group of 'good and (mentally) sound people' like him.

> 4) I make only a short intervention. I find myself in some difficulty because I consider the whole conceptual basis of the Bill to be <u>flawed</u>. I do not agree with it. I take the

view that these guns are freely available and **any lunatic** can obtain one. The Bill is <u>a bad Bill</u>.
(Lord Campbell of Alloway, 16 January 1997)

In addition to *lunatic*, there were other terms that could have different connotations and uses in our data. Such terms include *idiot* and *madman*, and of these two, *idiot* seems to have been generally considered as the more serious condition. The significant collocates of *idiot* include various terms that reference lack of mental capacity, such as *bumbling*, *blithering*, *lunatic* and *imbecile*. Examples (5) and (6) show instances of *idiot*, first with a more technical or medical meaning and secondly in a general use. In (5), the speaker talks about 'a helpless imbecile', emphasizing the fact that the first girl cannot take responsibility for her actions, partly because she is only five years old, partly because of her lack of mental faculties. The second girl is even more mentally deficient and not (criminally) responsible in the eyes of the law, being underaged and 'a microcephalic idiot'. In (6), being an idiot denotes being a person who foolishly calls people with learning difficulties using descriptors like 'thick'. The excerpt clearly shows the general attitude of the 1990s, where learning disabilities were associated with mental deficiency, but at the same time the speaker uses name-calling and labelling when he draws the line between a preferable and acceptable behavior and something that should not be approved of.

5) Let me drag out into the full light of day five cases on the waiting list in Essex, one in my own constituency. The first is a girl of five, **a helpless imbecile**, <u>destructive and mischievous</u>. Her mother has been in <u>a mental hospital</u> and threatens to murder the child. The second is a girl of 13, **a microcephalic idiot**, living in a family of 10 in the care of grandparents who are in poor health. She has been known <u>to knock her grandmother down</u>.
(Mr Braine, 5 November 1952)

6) When Tim Bakewell was six years old he was diagnosed by **some idiot** as thick. I believe that term is quite often applied to people who turn out to be dyslexic. He was considered to be bone idle because he did not appear to be able to read. People believed that that must be his fault in some way. Now, at the age of 11, he receives treatment in a special unit within an ordinary school.
(Lord Renwick, 4 April 1990)

Interestingly, as we saw in example (5), the use of *idiot* often connotes violent behavior. The speaker refers to one 'idiot' girl who is 'destructive', the other one 'knocking her grandmother down'. Both descriptions are mentioned as if they were only appositions, not attributes denoting real danger.

However, in examples (7), (8) and (9), real violence is very much in focus. The term *madman*, or *mad person*, appears to denote danger and destruction in many cases in the second half of the 20th century, in relation to crimes involving shootings, stabbings or wars. The list of collocates includes terms such as *extremist*, *handgun*, *rampage*, *hell*, and *bureaucratic*, as well as several proper nouns such as *Hitler*, *Stalin*, *Baghdad* and *Gaddafi*. In (7), *madman* is equated with paranoid and violent behavior, which under war conditions may result in destroying the world 'with one blow'. The example is from 1955, but it could easily be from the present, as the similar reference to *a madman* being in control of nuclear weaponry is still used today.

7) It made it inevitable in war that a head of a State who found himself hemmed in by the forces used against him should be a paranoic, a war criminal, **a madman**. I ask the question, how is it possible in a modern war – prepared as we must be today, with the things we are called on to do now to destroy with one blow literally millions, maybe tens of millions – for a man to sit in control of such a situation and not be **a madman**? Indeed, even to contemplate the possibility of doing it as we are now makes one ask which of us is not mad at this moment

to be seriously considering settling the problems of the world by <u>destroying the world</u>?
(Mr Hudson, 10 March 1955)

In (8), the juxtaposition between 'legitimate, decent and honest people' and 'one awful, awful madman' is used in the discussion about gun control in Northern Ireland in 1996. The parliamentary debate in question took place after a school shooting in the Dunblane Primary School, where 16 children and their teacher were shot dead. The 'madman' in question, Thomas Hamilton, is described as 'awful and deranged', but even more so by way of the descriptions given to members of the 'decent' society, for example, people who are 'legitimate', 'decent', 'honest', 'innocent' and 'honourable' – everything the perpetrator supposedly is not.

8) Regrettably, the Home Secretary's statement means that the sins of <u>this awful, deranged individual</u> will be paid for by the many thousands of <u>legitimate, decent and honest people</u> who enjoy not a so-called sport, as my right hon. and learned Friend the Member for Putney (Mr Mellor) put it, but a <u>leisure activity</u> that has great historic traditions and which they enjoy. The measures that he announced will be received with great dismay. He has got it completely the wrong way round. He is affecting many, <u>many innocent people</u> because of <u>the awful, awful acts</u> of **one awful, awful madman**. I think that he is wrong, and that his legislation will not be accepted by the majority of <u>decent, honourable people</u> in this country.
(Mr Carlisle, 16 October 1996)

Also in (9), the discussion about gun permits refers to the Dunblane massacre. Here, the term used is *maniac*, but just as in (8), the turn describes two opposite sides, i.e. someone who 'takes some mad action that causes great harm and injury' and members of the public who are the victims of that action. Also in the data as a whole, the collocates of *maniac* include terms such as *sadist*, *shot*, *homicide, gen-*

ocide, and *gun*. The term *maniac* denotes mental instability and violent behaviour, the public abiding to rules and regulations and being decent in the eyes of the law.

9) The point about **the maniac** is that it is not obvious that someone is **a maniac** until he takes some mad action that causes great harm and injury to other members of the public. It is a matter of having regulations and rules that protect the public in between the point where **a maniac or a mentally unstable person** thinks of taking some action and actually takes that action.
(Mr Henderson, 18 February 1997)

Sex and sexuality are also associated with the term *maniac*, but notably not with the other terms under investigation. The list of significant collocates of maniac includes terms such as *sex*, *sexual* and *lusting*, particularly during the 1950s and 1960s. Example (10) comes from the debate on the 1966 Sexual Offences Bill, which proposed to decriminalize homosexual acts between consenting adults. Leo Abse, the speaker, was a Welsh Labour MP with a noted interest in psychoanalysis and one of the main promoters of the Sexual Offences Act 1967, which decriminalized homosexual acts between consenting adults in England and Wales.

10) But when we are talking of all these men whose relationships, in private, are with adults who give their consent, we know that to talk of sending these men to prison is, as has been said again and again, as therapeutically useless as incarcerating **a sex maniac** in a harem: This is the position, and we all know it: There is no rehabilitative element at all in the punishment that is meted out.
(Mr Abse, 19 December 1966)

In example (10), we see Mr Abse arguing that if the objective of punishment by incarceration is to be therapeutic, there is no such benefit to be found in incarcerating gay men in all-male prisons, the hetero-

 NoSo 2023 | https://doi.org/10.59589/noso.32023.14413

sexual analogy of which is the harem mentioned in the text. The term *sex maniac* was a relatively recent coinage in British English at the time. According to data from the Google Books corpus, the concept emerged in the 1930s and its frequency increased markedly from the 1940s to the 1960s. The term occurs 15 times in the Hansard corpus, ten of which during the 1960s.

5. Discussion

The main focus of the present study is on semantic change in mental health terms and the use of formerly neutral but subsequently discredited terms for naming and labelling in the British Parliament. The data show that a subset of commonly used terms (*lunatic*, *idiot*, *imbecile*, etc.) underwent a process of semantic substitution around the 1930s and the 1940s, which rendered them largely inappropriate to use in formal and official contexts. Naturally, the shift was neither immediate nor complete, and we find some instances of these terms being used in seemingly serious manner up until the 1980s (example 11).

> (11) The honourable Gentleman listed some practical difficulties. He said that there is no test of capacity to vote for anyone else, but that we are creating a test of capacity for the voluntary mental patient. However, <u>one cannot appear on the register if one suffers from a legal incapacity to vote</u>, for example, if one is **an idiot or a person of unsound mind**.
> (Sir Mayhew, 18 October 1982)

Notwithstanding some occasional instances of archaic usage, the primary use of such terms underwent a gradual semantic change, shifting towards the colloquial as a result of advances in psychiatry and societal attitudes to mental health care. Rather than suggesting medically validated mental health issues, they came to be used in an exaggerated and sometimes jocular manner in reference to ideas, concepts and actions that the speaker considered ill-conceived, extreme or offensive. Notably, the terms themselves were occasionally the topic of explicit discussion, as in example (12):

> (12) It is a serious point seeking clarification for the future guidance of honourable Members. Quite rightly and making a factual statement of the truth, my honourable Friend the Member for Renfrew, West, called the honourable Member for Worcestershire, South, an idiot. For the guidance of honourable Members, <u>is</u> **'idiot'** <u>a parliamentary expression</u>, because it may be that many honourable Members will choose to use it on future occasions?
> (Mr Lewis, 15 October 1972)

Semantic changes take time, and most of the terms discussed here retain some connection to the field of mental health. Consequently, while the use of these terms in naming and labelling can no longer be interpreted as a serious claim that the object of reference is the product of, or associated with, an actual mental health condition, the implication remains that there is an association between things the speaker does not approve of, or understand, and mental health. Likewise, such naming practices often serve as intensifiers, highlighting a person's incompetence and lack of reason.

What is particularly interesting here is the pragmatic purpose of such labelling in the parliamentary context. While it seems self-evident that many, if not most, of such references are intended to characterize the referent, we also find instances of self-characterization both directly and by means of disassociating oneself from the object. The former are almost invariably jocular in nature, as in example (13):

> 13) I thank my right honourable Friend for that answer. He will know that I am **a technological imbecile**, and that many customers of the Government do not have the skills to deal with such technology. What efforts have been made to make it user-friendly?
> (Mr Rooney, 17 December 1997)

NoSo 2023 | https://doi.org/10.59589/noso.32023.14413

By contrast, the latter usage is perhaps of greater interest, because it has more subtle rhetorical implications. For example, when the leader of another country is labelled as a *madman*, the speech act signals that the person's actions and arguments are so far removed from the speaker's own rational perspective that they must have a mental health condition. This notion is reinforced through the common use of mental health terms in connection with violence and threatening behaviour, as observed in the present study.[7] More significantly, the implied presence of such a condition would seem to suggest that the person is incapable of rational debate and their positions cannot be understood or meaningfully discussed (cf. Clark 1992). They are simply 'mad'. However, at the same time, the speaker is also making it clear to the audience that the speaker is the opposite of mad: rational, clear-headed, someone who can be trusted. In political language use, such as in parliamentary debates, there is considerable value in convincing the audience that you are a reasonable and rational person while at the same time labelling your opponents' views as something that cannot even be discussed rationally.

So we return to the premise of social, and societal, polarization through labelling and naming practices. As introduced in Section 2.2, the division between neutral/positive and negative labelling reflects a prevalent attitude either in favor of or against people with mental illness versus 'normal' people. A person with mental illness, whether labelled as *lunatic* or *idiot*, can stereotypically be dangerous, morally deviant and irrational in their behavior – something a respectable member of society does not want in their 'civil' in-group. This moral juxtaposition is not restricted to the macro-level attitudes we hold about people with mental illness in society at large, but it is also expressed in the way we reach micro-level situational conclusions about people who mentally deviate from us in our daily social interaction.

[7] On the complicated issue of mental health conditions and violence, see e.g. DeAngelis (2021).

6. Conclusion

This study of naming and labelling with mental health terms in the British Parliament has shown how the old generic terms that underwent a semantic change around the Second World War continued to be used at a diminished frequency in the latter half of the 20th century. The terms were no longer referential to medically diagnosed mental health conditions, rather they were primarily used as intensifiers, for comedic purposes and as distancing devices. The latter helping speakers to portray political opponents, civil servants, foreign leaders and others as extremists and incompetent, whose actions and opinions are less than rational, and at times dangerous. Generally, the notion of deviance has been used as a basic means to make a division between the in-group of the 'sane' and the out-group of the 'insane'. Although in principle any type of behavior can be a called and considered lunatic or idiotic, the underlying moral stereotype may denote socially unacceptable, erratic conduct and even violence. When it comes to the actual labels used, cultural (mis)understandings about people with mental illness follow suit.

References

Adams, Michael. 2008. Nicknames, interpellation and Dubya's Theory of the State. *Names. A Journal of Onomastics* 56(4). 206–220.

Adams, Michael. 2009. Power, politeness and the pragmatics of nicknames. *Names. A Journal of Onomastics* 57(2). 81–91.

Ainiala, Terhi & Östman, Jan-Ola (eds.). 2017. *Socio-onomastics: The pragmatics of names*. Amsterdam: John Benjamins.

Alexander, Marc. (forthcoming) The language of the British Hansard. In Korhonen, Minna, Kotze, Haidee & Tyrkkö, Jukka (eds.), *Exploring language and society with big data: Parliamentary discourse across time and space*. Amsterdam: John Benjamins.

Allport, Gordon. 1986. The language of prejudice. In Escholz, Paul, Rosa, Alfred & Clark, Virginia (eds.), *Language awareness*, 261–270. New York: St. Martin's Press.

Arnold, Catherine. 2008. *Bedlam: London and its mad*. New York: Simon & Schuster.

Balfour, James. 2019. “The mythological marauding violent schizophrenic”: Using the word sketch tool to examine representations of schizophrenic people as violent in the British press. *Journal of Corpora and Discourse Studies* 2. 40–64.
Balfour, James, 2020. Representation of people with schizophrenia in the British press. In Friginal, Eric & Hardy, Jack A. (eds.), *The Routledge handbook of corpus approaches to discourse analysis*, 537–553. London: Routledge.
Birch, Michael. 2011. *Mediating mental health: Contexts, debates and analysis*. London: Routledge.
Clark, Kate. 1992. The linguistics of blame: Representations of women in the Sun reporting of crimes of sexual violence. In Toolan, Michael (ed.), *Language, text and context: Essays in stylistics*, 208–224. London: Routledge.
Croom, Adam M. 2013. How to do things with slurs: Studies in the way of derogatory words. *Language and Communication* 33(3). 177–204.
Cross, Simon. 2014. Mad and bad media: Populism and pathology in the British tabloids. *European Journal of Communication* 29(2). 204–217.
DeAngelis, Tori. 2021. Mental illness and violence: Debunking myths, addressing realities. *Monitor on Psychology* April/May. 31–36.
Froehlich, Heather. 2016. “Thus to make poor females mad”: Finding the ‘mad woman’ in Early Modern drama. In Nevala, Minna, Lutzky, Ursula, Mazzon, Gabriella & Suhr, Carla (eds.), *The pragmatics and stylistics of construction and characterisation*. Online, *Studies in Variation, Contacts and Change in English* 17. Helsinki: VARIENG. http://www.helsinki.fi/varieng/series/volumes/17/froehlich/ (accessed 11 May 2022)
Geeraerts, Dirk, Grondelaers, Stefan & Bakema, Peter. 1994. *The structure of lexical variation: Meaning, naming, and context*. Berlin: De Gruyter.
Gumperz, John J. & Cook-Gumperz, Jenny. 1982. Introduction: Language and the communication of social identity. In Gumperz, John J. (ed.), *Language and social identity*, 1–21. Cambridge: Cambridge University Press.
Hiltunen, Turo, Räikkönen, Jenni & Tyrkkö, Jukka. 2020. Investigating colloquialization in the British parliamentary record in late 19th and early 20th century. In Hiltunen, Turo & Loureiro Porto, Lucía (eds.), *New perspectives on democratization* (Language Sciences 79), 1–16. https://doi.org/10.1016/j.langsci.2020.101270.
Hintikka, Marianna & Nevala, Minna. 2017. Representations of prostitutes and prostitution as a metaphor in nineteenth-century English newspapers. *Journal of Historical Sociolinguistics* 3(2). 219–240.

Hock, Hans Henrick & Joseph, Brian D. 2009. *Language history, language change, and language relationship: An introduction to historical and comparative linguistics*. Berlin: Mouton de Gruyter.

Houston, Robert Allan. 2000. *Madness and society in eighteenth-century Scotland*. Oxford: Clarendon Press.

Janes, Dominic. 2014. Oscar Wilde, sodomy, and mental illness in late Victorian England. *Journal of the History of Sexuality* 23(1). 79–95.

Kamada, Tomihisa & Kawai, Satoru. 1989. An algorithm for drawing general undirected graphs. *Information Processing Letters* 31(1). 7–15.

Laitinen, Mikko & Säily, Tanja. 2018. Google Books: A shortcut to studying language variability? In Nevalainen, Terttu, Palander-Collin, Minna & Säily, Tanja (eds.), *Patterns of change in 18th-century English: A sociolinguistic approach*, 223–233. Amsterdam: John Benjamins.

Marjanen, Jani et al. 2019. A national public sphere? Analysing the language, location and form of newspapers in Finland, 1771–1917. *Journal of European Periodical Studies* 4(1). 54–77.

Mayr, Andrea & Machin, David. 2012. *The language of crime and deviance: An introduction to critical linguistic analysis in media and popular culture*. London: Continuum.

McEnery, Anthony & Baker, Helen. 2019. Language surrounding poverty in early modern England: A corpus-based investigation of how people living in the seventeenth-century perceived the criminalized poor. In Suhr, Carla, Nevalainen, Terttu & Taavitsainen, Irma (eds.), *From data to evidence in English language research*, 223–257. Leiden: Brill.

Michel, Jean-Baptiste, Shen, Yuan Kui, Aiden, Aviva Presser, Veres, Adrian, Gray, Matthew K., The Google Books Team, Pickett, Joseph P., Hoiberg, Dale, Clancy, Dan, Norvig, Peter, Orwant, Jon, Pinker, Steven, Nowak, Martin A. & Aiden, Erez Lieberman. 2011. Quantitative analysis of culture using millions of digitized books. *Science* 331(6014). 176–182.

Mollin, Sandra. 2007. The Hansard hazard: Gauging the accuracy of British parliamentary transcripts. *Corpora* 2(2). 187–210.

Nevala, Minna. 2016. Solidarity in evaluation: The portrayal of murderers and their victims in late nineteenth-century newspapers. In Nevala, Minna, Lutzky, Ursula, Mazzon, Gabriella & Suhr, Carla (eds.), *The pragmatics and stylistics of construction and characterization*. Online, *Studies in Variation, Contacts and Change in English* 17. Helsinki: VARIENG. http://www.helsinki.fi/varieng/series/volumes/17/nevala/ (accessed 11 May 2022).

Nevala, Minna. 2019. Two miserable creatures or those atrocious criminals? Evaluative reference in the Mannings murder reporting. In Lutzky, Ursula & Nevala, Minna (eds.), *Reference and identity in public discourses*, 19–41. Amsterdam: John Benjamins.

Nevala, Minna & Tyrkkö, Jukka. (forthcoming) From criminal lunacy to psychiatric illness: The changing discourses of mental health in the British Parliament. In Korhonen, Minna, Kotze, Haidee & Tyrkkö, Jukka (eds.), *Exploring language and society with big data: parliamentary discourse across time and space*. Amsterdam: John Benjamins.

Ochs, Elinor. 1993. Constructing social identity: A language socialization perspective. *Research on Language and Social Interaction* 26. 287–306.

O'Reilly, Carole. 2014. 'Dirt, death and disease': Newspaper discourses on public health in the construction of the modern British city. *Journal of Historical Pragmatics* 15(2). 207–227.

Partington, Alan. 2010. Modern diachronic corpus-assisted discourse studies (MD-CADS) on UK newspapers: An overview of the project. *Corpora* 5(2). 83–108.

Partington, Alan, Duguid, Alison & Taylor, Charlotte. 2013. *Patterns and meanings in discourse: Theory and practice in corpus-assisted discourse studies* (CADS). Amsterdam: John Benjamins.

Porter, Roy. 2002. *Madness: A brief history*. Oxford: Oxford University Press.

Potts, Amanda & Weare, Siobhan. 2018. Mother, monster, Mrs, I: A critical evaluation of gendered naming strategies in English sentencing remarks of women who kill. *International Journal for the Semiotics of Law-Revue Internationale De Sémiotique Juridique* 31(1). 21–52.

Price, Hazel. 2019. *The discursive construction of mental illness in UK newspapers (1984–2014): A critical corpus linguistic analysis*. University of Huddersfield.

Slembrouck, Stef. 1992. The parliamentary Hansard "verbatim" report: The written construction of spoken discourse. *Language and Literature* 1. 101–119.

Stern, Gustaf. 1931. *Meaning and change of meaning*. Bloomington: Indiana University Press.

Suzuki, Akihito. 1999. Framing psychiatric subjectivity: Doctor, patient and record-keeping at Bethlem in the nineteenth century. In Forsythe, Bill & Melling, Joseph (eds.), *Insanity, institutions, and society, 1800–1914: A social history of madness in comparative perspective*, 115–136. London: Routledge.

Sweetzer, Eve. 1991. *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure*. Cambridge: Cambridge University Press.

Tyrkkö, Jukka. 2019. Kinship references in the British Parliament, 1800–2005. In Lutzky, Ursula & Nevala, Minna (eds.), *Reference and identity in public discourses*, 97–124. Amsterdam: John Benjamins.

Tyrkkö, Jukka. 2020. The war years: Distant reading British parliamentary debates. In Hansson, Joacim & Svensson, Jonas (eds.), *Doing digital humanities: concepts, approaches, cases*, 169–199. Växjö: Linnaeus University Press.

Tyrkkö, Jukka & Frisk, Irina. 2020. Crooked Hillary, Lyin' Ted and failing New York Times: Nicknames in Donald Trump's tweets. In Schneider, Ulrike & Eitelmann, Matthias (eds.), *Linguistic inquiries into Donald Trump's language: From 'fake news' to 'tremendous success'*, 109–129. London: Bloomsbury Academic.

Tyrkkö, Jukka & Mäkinen, Ilkka. 2022. Culturomic explorations of literary prominence using Google Books: A pilot study. *Knygotyra* 78, special issue "Bestsellers, Digital Technologies and their Limitations in Book History" (ed. Martyn Lyons). 111–138.

van Dijk, Teun A. 2009. *Society and discourse: How social contexts influence text and talk*. Cambridge: Cambridge University Press.

van Leeuwen, Theo. 2008. *Discourse and practice: New tools for critical discourse analysis*. Oxford: Oxford University Press.

Vice, John & Farrell, Stephen. 2017. *The history of Hansard*. London: House of Lords Hansard and the House of Lords Library.

Nordisk tidskrift för socioonomastik / Nordic Journal of Socio-Onomastics (131–154)

Appellativized first names as terms of abuse: The case of the three pejorative Finnish expressions *uuno*, *tauno*, and *urpo**

Maria Sarhemaa

Abstract: In the article, three appellativized Finnish personal names are analysed in their use as a derogatory part of communication. In support of the argumentation, two kinds of data are used: corpus data from the internet forum Suomi24 ('Finland24') and internet questionnaire data. The methods of qualitative corpus and questionnaire research are adapted, and the approach is loosely cognitive. It is evident that the three expressions are used pejoratively and to express negative emotions. Thus, their usage makes the communication impolite or emphasizes impoliteness. By showing the meanings and usages of the three expressions, it is argued that appellativized proper names have a specific, complex semantic profile.

Keywords: socio-onomastics, cognitive onomastics, personal names, first names, appellativization, deonymization, terms of abuse, derogatory terms

---

Maria Sarhemaa (University of Helsinki). Appellativized first names as terms of abuse: The case of the three pejorative Finnish expressions *uuno*, *tauno*, and *urpo.*

* The research for this article was supported by a grant from the Kone Foundation in Finland.

NoSo 2023 | https://doi.org/10.59589/noso.32023.14416

1. Introduction

Using first names as appellatives or adjectives is a common and productive phenomenon in contemporary colloquial Finnish. In this article, I analyse appellativized first names as a part of communication, more precisely, as terms of abuse. In this study, I look at first names which have lost their properhood, i.e. have been appellativized, and are used pejoratively or as terms of abuse. Even though appellativized expressions are widespread not only in Finnish but in various languages, there are few studies on the use of appellativized first names as terms of abuse.

I focus on three appellativized expressions, *uuno*, *tauno*, and *urpo*, which are originally Finnish male names but can be used colloquially as appellatives or adjectives with pejorative meanings. I have chosen these three expressions, as they are widely known among Finnish speakers and they were appellativized in different decades, with *uuno* being the oldest and *urpo* the latest expression as an appellative. I discuss the meanings and usages of these three expressions. I also address the semantic profile of the appellativized proper names as terms of abuse and how insulting they are. I argue that appellativized personal names have a specific semantic profile, and their meaning is thus more complex than that of general terms of abuse.

I use two kinds of data to support my argument, namely, corpus data from a Finnish internet forum and online questionnaire data, using qualitative corpus and questionnaire research methods with a loosely cognitive linguistic approach. By drawing on material collected from an internet forum, I can determine how the expressions are used authentically in informal contemporary language. On the other hand, the questionnaire data provides semantic intuition of native speakers. Naïve speakers may not be capable of giving detailed information about the differences of the expressions or exact semantic descriptions, nevertheless, a large sample can help clarify semantic issues (Langacker 2008:86).

The analysis illustrates that *uuno*, *tauno*, and *urpo* have common meanings, but they also have individual meanings and usages. They are used to express negative emotions, and their usage is either a part of impolite communication or makes the communication impolite.

The structure of the article is as follows. First, the context is outlined (Section 2), followed by the data and the method (Section 3). Thereafter, the meanings and usages of the three expressions are analysed using the empirical data (Section 4). The approach then becomes more theoretical, with a discussion of the semantic profile of the appellativized proper names used as terms of abuse (Section 5). Finally, the findings are summarised and discussed (Section 6).

2. Outlining the context

2.1 Appellativization of *Uuno*, *Tauno*, and *Urpo*

Appellativization is the process in which a proper name acquires an appellative meaning, i.e. appellativizes. The terminology of the phenomenon varies, and the terms *deproprialisation* ~ *deproprialization*, *deonymisation* ~ *deonymization* ~ *deonimization* and *antonomasia* have been used in addition to appellativization (see e.g. Harvalík 2012:12; Reszegi 2018:25).

A previous study of mine (Sarhemaa 2021), shed light upon the background and the motivation of the appellativization of *Uuno*, *Tauno*, and *Urpo*. Of these, *uuno* is the oldest expression as an appellative with the first examples of its use being from the 1910s (Paunonen 2000). During that time, *Uuno* was a common first name, but after the appellativization its popularity slowly started to decrease. According to Hämäläinen (1956:207–208), in the 1920s, *uuno* was a common expression in army slang with the meanings of 'stupid', 'silly', 'incompetent', 'beginner'. There were also compounds with the head *uuno*, for example *talliuuno* (literally 'stable' + *uuno*) 'horse-drawn vehicle driver', 'horseman'. The motivation of the appellativization of *Uuno* remains unclear. Nevertheless, Paunonen (2018) finds it possible that the origin of the pejorative *uuno* might relate to the founder of the Finnish elementary school system, Uno Cygnaeus (1810–1888). It is possible that school, which many pupils considered as boring, got the name of the founder as its symbol, and the pupils started to call teachers *uuno*. Afterwards the expression *uuno* [stupid] spread to broader use.

The first name *Tauno* originates in Karelia[1] and is probably based on the homonymous adjective *tauno*, which in Karelian language has the meaning of ‘gentle’, ‘good-tempered’ (Vilkuna 2005; Saarelma 2007). The adjective has been used in Finnish dialects too, especially from the Karelian Isthmus. Therefore, it could be possible that the adjective *tauno* has went through a change of meaning and gotten the meaning of ‘stupid’ in colloquial Finnish. However, I find this unlikely. Hence, the motivation for the appellativization needs to be discovered elsewhere. More plausible explanation is given by Paunonen (2018). According to Paunonen (2000), the first examples of the name being used to mean ‘stupid’ in Helsinki slang are from the 1950s. Paunonen (2018) assumes that the motivation behind the meaning ‘stupid’ lies on the expressions with somewhat similar sound structure and pejorative meaning, such as *taulapää* [blockhead], *tauhka* [junk], and *taukki* [stupid] (*Taukki* has also been a hypocorism of *Tauno*). Most of those named *Tauno* were born before the appellativization happened, in 1915–1929. That means that in the assumed time of the appellativization *Tauno* was mainly a name of middle-aged men. This may have a bearing on the appellativization, as young people likely found the name old-fashioned.

Urpo has never been a common first name, nor does it appear to be gaining popularity; it was already a rare name in the 1970s, when it is assumed to have been appellativized (Paunonen 2000). Its appellativization may lie on sound symbolism, that is, the sound structure of *urpo*, more precisely, the consonant cluster *rp*. It can be considered as a phonestheme, which is a submorphemic, affective, meaning-carrying entity (Bolinger 1950:130; Hinton et al. 1994:5; see e.g. Abelin 2015; Kawahara et al. 2018; Shih & Rudin 2021; for more about sound

[1] Karelia is an area located in two sides of the border of Finland and Russia, most of it being situated in Russia. In the area, besides Finnish and Russian languages, also Karelian language is spoken. Karelian is a Finno-Ugric language which is closely related to Finnish.

 NoSo 2023 | https://doi.org/10.59589/noso.32023.14416

symbolism in onomastics). According to Paunonen (2018), an affective verb *urputtaa* [shoot one's mouth off] which is an older slangy expression than *urpo* with the same phonestheme, may also have motivated the appellativization.

2.2 Impoliteness and terms of abuse

Using terms of abuse is a way to express impoliteness in language, although there is no consensus about the exact definition of impoliteness. Bousfield & Locher (2008:3) establish a minimalist definition 'Impoliteness is behaviour that is face-aggravating in a particular context' (see Bousfield & Locher 2008:3–8 for more exhaustive information about impoliteness). Culpeper (2011:19–21) discusses several definitions of impoliteness and summarizes that its key notions are face (threatening), social norms, intentionality, and emotions. Linguistic impoliteness has been researched from different angles and disciplinary perspectives, for example in discursive pragmatics (Garcés-Conejos Blitvich & Sifianou 2019) and from cross-cultural perspective (Culpeper et al. 2014).

According to Culpeper (2010:3242–3243), the conventionalised impoliteness formulas include – not exhaustively – insults, pointed criticisms and complaints, challenging or unpalatable questions and/or presuppositions, condescension, message enforcers, dismissals, silencers, threats, and negative expressions. Of course, these are not all equally impolite, some of them might even seem inoffensive and require specific context to have the effect of impoliteness (Culpeper 2010:3243). The same phrase may be extremely impolite in one context, but entirely innocuous in another. Muikku-Werner (2005) has studied linguistic jeering – a type of impoliteness – and concluded that jeering is a continuum with malicious wickedness at one end and friendly teasing at the other. The latter is a positive phenomenon which expresses community spirit and being a part of 'us' (Muikku-Werner 2005:274).

In this article, I only focus on one way of expressing impoliteness in language, namely appellativized expressions used as terms of abuse. When speaking about abusive terms, the terminology varies. They have been called terms of abuse (e.g. De Raad et al. 2005; Van Oudenhoven et al. 2008; Beirne 2020), derogatory words (e.g. Norri 2000) and (dysphemistic) terms of insult (Allan & Burridge 1991:27–28). I do not address the differences with these terms. I use '(appellativized) terms of abuse' to refer to appellativized expressions used in a derogatory meaning.

According to Allan & Burridge (1991:27–28), dysphemistic terms of insult can metaphorically compare people with animals (e.g. calling someone a *rat*), include tabooed body parts or sexual behaviour (e.g. *motherfucker*), or mental or physical inadequacy (e.g. *idiot*). De Raad, Oudenhoven & Hofstede (2005) compared the use of terms of abuse in three languages, Spanish, Dutch, and German, and found that the meanings of the terms of abuse vary in these languages. They argue that there are cultural differences in abusive language and that the different terms of abuse violate distinct cultural values and may also display fewer known taboos in a given culture (De Raad et al. 2005:163).

The dysphemism of a term of abuse might lie in the way the expression is used. Hence, it is not necessarily a property of the word itself (Allan & Burridge 1991:28). The expressions I analyse are appellativized first names. Therefore, it is an interesting question whether the dysphemism in this case is a property of the word or not. Indeed, it is not a property of the homonymous first name, rather it can be a property of the appellativized form. The pejorative nature of the expressions will be discussed in the coming sections.

De Raad et al. (2005) distinguish the use of terms of abuse according to the person in which the sentence is used. First-person expressions can be used for self-disclosure and to rouse oneself, and third-person expressions to exchange negative information. However, if we look at terms of abuse as a method of hurting someone with verbal abuse, De Raad et al. (2005) propose that second-person expressions are the most interesting. They may be used to express breach of expectations, as well as to elicit a certain reaction. Using terms of abuse also tells something about the person uttering them. They

 NoSo 2023 | https://doi.org/10.59589/noso.32023.14416

may be used to affect another person and make them despicable. (De Raad et al. 2005:153–154.) The use of terms of abuse may thus be seen as representing confrontation between ‘us’ and ‘them’. Pälli (2003) studies human groups as discursive formations and adapts the terms ‘in group’ and ‘out group’, originally used in the framework of social psychology. Human beings have groups with which they identify (in groups), but also groups with which they do not identify (out groups). Negative features relate to out groups, and the members of them are seen as similar with each other (Pälli 2003:41). In Section 4.1 I analyse the use of *uuno*, *tauno*, and *urpo* in different grammatical persons in the corpus data.

3. Data and method: Corpus and questionnaire

I analyse two data sets in this study. The first was collected from the Suomi24 Sentences Corpus (Aller Media Ltd. 2014) provided by The Language Bank of Finland. This corpus includes all the discussion forums of the Suomi24 (‘Finland24’) website from the years 2001 to 2016. People use the Suomi24 forums to discuss various topics, such as human relations, entertainment and politics. I collected a total of 1,500 sentences in which appellativized *uuno*, *tauno*, or *urpo* have been used (500 sentences each). The language used on the Suomi24 is informal and the participants remain anonymous. Moderators may remove comments or threads, but this only usually happens if user asks them to do so (Lagus et al. 2016:9). Only slightly moderated discussion forums allow various expressions and free the relatively anonymous users from the fear of expressing extreme views (Sobkowicz & Sobkowicz 2012:448–449). Abusive and impolite language is thus evident on Suomi24 as can be seen in the upcoming section. Discussions on Suomi24 include hate speech and even death threats (Harju 2018:63), however, there are no examples of the latter in the data.

The second data set was compiled through an online questionnaire. I used the E-lomake online form to design and publish a questionnaire. E-lomake is a form builder software that enables survey creation and management. First, a pilot questionnaire was available to a limited number of respondents over a four-day period. As no changes were made to the questionnaire after the pilot, the responses of the pilot are also included in the data. After the pilot, the questionnaire was available to the public online for one week during autumn 2017. I posted it as a public post on my Facebook profile and asked people to share it. It was shared 71 times and completed by 542 respondents and a total of 531 responses were included in the data. One participant was excluded as they were underage and it was not possible to obtain parental consent and ten respondents were excluded from the data as they were not native speakers of Finnish. Evaluating the meaning of colloquial expressions requires native-like language skills, and the number of the non-native speakers was not big enough for comparing their answers with the native speaker group. The age distribution of the respondents was 18–86 years, with the average age being 41 years. The majority (85.3%) of the respondents were women, 14.7% were men, 0.2% non-binary, and one did not state their gender. In the questionnaire, the respondents were asked about the meanings of *uuno*, *tauno*, and *urpo*, and how pejorative they find them. They were also asked to fill one of the expressions in sentences that had been adapted from the Suomi24 data. The entire questionnaire (in Finnish) is presented in an appendix to Sarhemaa (forthcoming).

I adapt the methods of qualitative corpus and questionnaire research. I categorize the corpus data according to the meanings and usages of the expressions and analyse them qualitatively. Using the numerical questionnaire data, I estimate how the informants have rated the pejorativity of the expressions. I have compiled frequency lists of the open-response items from the questionnaire data using #LancsBox (Brezina et al. 2020). By analysing them, I outline the meanings of *uuno*, *tauno*, and *urpo*. I use a loosely cognitive linguistic approach, i.e. I adapt appropriate methodology and concepts of cognitive semantics. Furthermore, I utilize the cognitive concepts of conceptualization and domain and base my analysis on them.

4. Meanings and usages of *uuno*, *tauno*, and *urpo*

4.1 Corpus data

According to dictionaries, *uuno*, *tauno*, and *urpo* are to some extent synonymous, and mean ‘dumb’ and ‘stupid’, or other similar pejorative adjectives. The extensive dictionary of Helsinki slang (Paunonen 2000) gives additional meanings to these expressions, including ‘slow-witted’ and ‘fool’ for *uuno*, ‘crazy’ and ‘dull’ for *tauno*, and ‘weird’ for *urpo. The Dictionary of Contemporary Finnish* (Kielitoimiston sanakirja) adds ‘jerk’ for *uuno* and *urpo*, and ‘yokel’ for *uuno*. It does not, however, include *tauno* as a search term. In summary, the basic meaning of *uuno*, *tauno*, and *urpo* is pejorative. Thus, they can be described as somewhat synonymous even though the dictionaries describe also differences between them. Nevertheless, my analysis is based on the impossibility of synonymity. Even if two expressions designate the same situation, they are not semantically identical because they are structuring it through different images (Langacker 1991:35). Hence, one of my research questions is the meanings and usages of these expressions: What do they mean in contemporary colloquial language? How do they differ?

I analysed 1,500 sentences from the Suomi24 data as per the meanings and usages of *uuno*, *tauno*, and *urpo*. It was often difficult to determine whether the meaning of the expression was ‘silly’, ‘stupid’ or something similar, so I refer to these meanings simply as pejorative. Therefore, the corpus data are used particularly to study the uses of the expressions, whereas their meanings are distinguished especially by using the questionnaire data. In addition to pejorative meaning, however, the expressions in the corpus data are sometimes used in the meaning of ‘ordinary person’, ‘average Joe’. In this case, the pejorative connotation is often questionable if the sentence or the broader context does not include something which clearly carries pejorative meaning, such as derogatory modifiers.

When the expression is used pejoratively, it may refer to another participant in the same thread, to the writer themself, or to someone else. In the latter case, the referent can be either generic or a specific person who is not participating in the conversation. I refer to this group simply as a general pejorative meaning. The data also include sentences in which the verb *uunottaa*[2] [scam, cheat] derived from *uuno* is used, or include playful modifications based on the names of two Finnish politicians[3] and the expression *urpo*, and have a pejorative meaning. However, I do not discuss these two latter cases here, because the *uunottaa* verb is not used as a term of abuse, and the specific forms derived from *urpo* only refer to two individuals, instead of being used generally abusively. Figure 1 shows the distribution of meanings, and the referents of *uuno*, *tauno*, and *urpo* in parentheses in the corpus data all added together (excluding the latter cases).

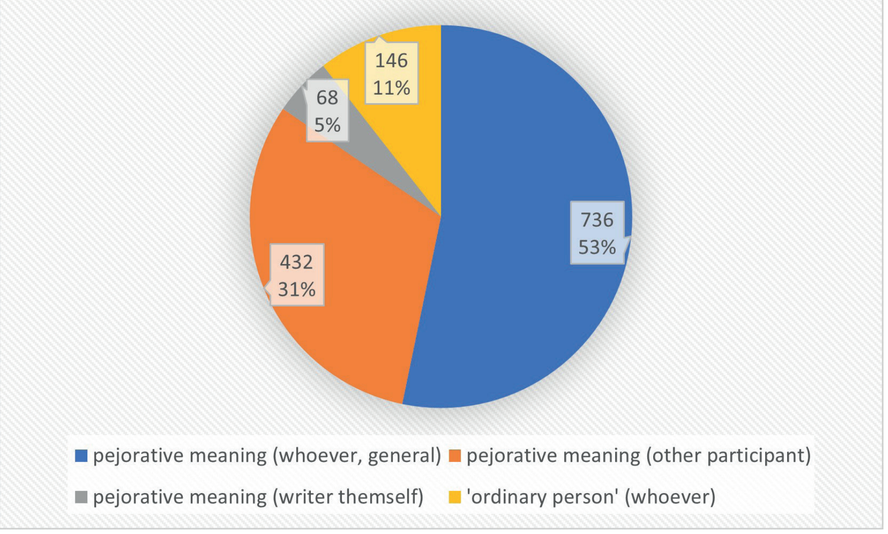

Figure 1. Meanings and referents of *uuno*, *tauno*, and *urpo* added together in the corpus data.

[2] *Uuno-tta-a* contains the stem *uuno*, a verbal derivational affix -*tta*- and an infinitive ending *a*. The meaning of the derivational affix -*tta*- is to make the entity into an item expressed by the stem (VISK § 317).

[3] These politicians are Petteri Orpo, the chair of the Finnish National Coalition Party and Jutta Urpilainen, European Commissioner. They both have formerly served as ministers in Finland. The sound structure of their surnames is somewhat similar to the word *urpo*, which motivates the pejorative modifications of their names such as *Petteri Urpo* and *Jutta Urpolainen*.

 NoSo 2023 | https://doi.org/10.59589/noso.32023.14416

In more than half of the sentences (53%), the expression has a general pejorative meaning and the referent can be anyone (Example 1), and almost one-third (31%) is used pejoratively to refer to another participant (Example 2). Hence, it is rare (5%) for somebody to refer to themself pejoratively using these expressions (Example 3). In 11% of the sentences the expression has the meaning of 'ordinary person', which is usually emphasized with a modifier such as *tavallinen* [ordinary] (Example 4). In these cases, the referent can be anyone.

1. Kohta joku uuno taas kyselee ihan samaa.
   [Soon some *uuno* will ask the same again.]

2. Hanki urpo itsetunto!
   [Get some self-respect, *urpo*!]

3. Olen ihan uuno ja aloittelija, joten kaipaisin neuvoja alusta loppuun.
   [I am such an *uuno* and a novice, so I would like advice from start to finish.]

4. Noinpa se tavallinen tauno haksahtaa linuksiin.
   [This is how an ordinary *tauno* slips up using Linux.]

As stated above, the three expressions are most often used in a general pejorative meaning. This holds true not only when the expressions are counted together but also when we look at each of them separately. However, when studying the other usages, we can find differences between the three expressions.

Uuno is the expression which most often refers to oneself and can be used for self-disclosure (see De Raad et al. 2005:153). In Example 3, the writer explicitly mentions that they are a novice, which can be understood as an attempt to disclose their inexperience and as a form of apology. In the data, the sentences in which the appellativized expression refers to oneself are often a part of a thread where someone is asking for advice and apologizing for their question. An apology can be seen as endeavour to avoid unpleasant responses and to ease tension in relation to the other participants of the conversation. Since *uuno* is the most common expression in this use, it seems to be the most likely expression to use for easing tensions when constructing social relations.

Typical for *tauno* is the use of it in the meaning of 'ordinary person'. One reason for that might be its use with the alliterative adjective *tavallinen* [ordinary] (see Example 4) or its colloquial forms *tavis* or *tavan*. Indeed, *Tauno* is often used with a modifier that emphasizes the meaning of 'ordinary'. The modifier, however, is not always *tavallinen* or its derivative. Other examples are *takametsän* [backwoods] and *naapurin* [neighbour]. Nevertheless, it is rare for *tauno* to be used in this meaning without any modifier. When *tauno* is used in the meaning of 'ordinary person' is not necessarily always pejorative, or at least its pejorativity is questionable.

Of the three expressions, *urpo* is the one which is most often used in a pejorative meaning referring to other participants of the discussion. Example 2 is an insult for another writer, and in it *urpo* is by far used as a term of abuse.

In addition to insults, applying Culpeper's (2010:3242–3243) conventionalised impoliteness formulas introduced in Section 2.2, it can be shown that in the corpus data *uuno*, *tauno*, and *urpo* are used in sentences which express all impoliteness formulas except threats. This implies that they are not formidable enough to express threat, but instead they can express impoliteness in various ways.[4]

As mentioned above, when studying the terms of abuse, second-person expressions are the most thought-provoking (see De Raad et al. 2005:153–154). By directly addressing a person with a term of abuse, the speaker may indicate that the other person belongs to the out group, in other words, is not a member of the speaker's in group. This kind of use is common in the corpus data (Examples 5–7). Other people can be referred to by either using the verb in second-person singular (5) or second-person plural (6) or by explicating the reference

[4] Threats might also be less frequent on internet forums than the other formulas mentioned. This may be one reason why there are no threats in the data.

in a different way, such as the writer addressing themself to someone by using the appellativized expression (7).

5. Oletko niin uuno, ettet muka tiedä, että kaikki valtamedia on näitten uutta siionistista maailmanjärjestystä ajavien konkkanokkien omistuksessa?
   [Are you such an *uuno* that you pretend you don't know that all mainstream media is owned by these large-nosed people furthering the new Zionist world order?]

6. Vittu te ootte taunoja, ei helkkari vieköön :D
   [Shit, you are such a *taunos*, fucking hell :D]

7. Ryyppäminen on vanhanaikaista urpot.
   [Boozing is old-fashioned, (you) *urpos*.]

In Example 5, the writer's in group believes in a Zionist conspiracy theory. The writer seems to find it impossible that someone is such an *uuno* that they do not know that the conspiracy exists. By using the word *uuno*, the writer on the one hand expresses that the other person belongs to the out group, and on the other hand suggests that the person might not have the necessary mental abilities to understand their claim. In Example 7, the writer indicates that drinking heavily is not desired behaviour in their in group (at least anymore), whereas the *urpos* of the out group still do that. In Example 6, however, the writer does not elaborate on why the other people are considered to be *taunos* belonging to the out group, so it is necessary to look at the larger context. Example 6 was taken from a thread about cars. The original poster writes that they have bought a premium car and the others start to discuss about what constitutes a premium car and what does not. Example 6 is part of a message where the writer suggests that the original poster does not see that the other participants of the discussion are taunting them. Thus, the writer tries to show to the original poster that they have been placed in the out group by the others, but at the same time the writer also excludes the original poster from their own in group.

The sentences in which a term of abuse is used to address another person are pejorative, and occasionally the pejorativity is emphasized by using swear words. In Example 6, there are two swear words which emphasize the pejorativity. However, an emoticon is used at the end of the sentence, which tends to make the sentence less insulting (see Huffaker & Calvert 2005). All the three expressions, *uuno*, *tauno*, and *urpo* may be used highly pejoratively, in rude and impolite sentences with swear words, but also in less pejorative sentences in which the only hint of abuse is the appellativized expression itself. For instance, Example 7 would not really be impolite without the word *urpot* at the end. That is to say, the appellativized expressions may be used as a part of impolite communication, but they can also be used to make the communication impolite.

Thus, it can be concluded that by using appellativized expressions as terms of abuse it is possible to represent confrontation and impoliteness, as well as to convey other people's belonging to the out group, as it is by using other terms of abuse. However, the semantic profile of appellativized expressions is different from that of the other terms of abuse. This will be discussed further in Section 5.

4.2 Questionnaire data

In the questionnaire, the respondents were asked to fill in *uuno*, *tauno* or *urpo* in 20 sentences, which were partly formed by using the sentences from the corpus data. In the responses to 16 sentences, the most often used expression was *urpo*. This implies that *urpo* is the most widely known and used expression of these three. *Uuno* was the most often chosen expression in the two sentences, which refer to oneself, and *tauno* in the two sentences, in which the missing word has the meaning of 'ordinary person'. One of these sentences included the word *tavallinen* [ordinary] and the other the word *naapurin* [neighbour] as a modifier. These findings are in line with the analysis of the corpus data.

The respondents were also explicitly asked about the meanings of the three expressions (e.g. *What does urpo mean*?). The most often mentioned feature for *urpo* and *uuno* was *tyhmä* [stupid]. For *tauno*, however, this was only the third most often mentioned feature. *Tauno* was most often mentioned to be *tavallinen* [ordinary], and the second most often *yksinkertainen* [simple]. The corpus data show that *tauno* is indeed used pejoratively, however, the questionnaire data confirms that besides that, a central meaning of *tauno* is 'ordinary'. *Hidas(älyinen)* [slow(-witted)], *hyväntahtoinen* [benevolent] and *hölmö* [fool] are examples of other common features questionnaire respondents often connected to tauno. This implies that *tauno* is rather a kind and decent expression despite its pejorative meaning. Someone called *tauno* seems thus stupid because of their mental features, such as slow-wittedness. In other words, their stupidity is unintentional.

Hölmö [fool], *yksinkertainen* [simple], and *hidasälyinen* [slow-witted] are features also connected with *uuno*. Other features the respondents mentioned were for example *tietämätön* [ignorant] and *ajattelematon* [unthinking]. This shows that a person called *uuno* may act foolishly because they do not pay attention to their own behaviour. Therefore, the stupidity of *uuno* is unintentional, as it is with *tauno* too.

Urpo, on the other hand, is often connected with features such as *idiootti* [idiot], *ärsyttävä* [irritating], *ääliö* [jerk]. Also, the word *tahallaan* [on purpose] recurs in the answers. Consequently, *urpo* is someone acting stupidly on purpose, not because of their mental properties.

The respondents were also asked about how negative they find the three expressions, i.e. the pejorativity of the expressions. The answers for the questions about the pejorativity demonstrate that *urpo* is the most pejorative of the three expressions. The respondents were asked to give the expressions a value between 1 and 5, 1 being the least negative and 5 the most. The average values were *urpo* 4.56, *uuno* 3.60 and *tauno* 3.18. This shows that *urpo* is the most pejorative of these expressions, and the difference in the pejorativity between it and the second most pejorative expression *uuno* is greater than the difference between *uuno* and *tauno*. The answers to the other question about pejorativity confirmed this finding; respondents were asked to rank the expressions based on their negativity, with the expression with the most negative tone receiving the value 1 and the least negative tone 3. The average values were *urpo* 1.25, *uuno* 2.15 and *tauno* 2.61. This clearly confirms that *urpo* is the most pejorative and consequently the most insulting of these expressions, whereas *tauno* is the least pejorative and the least insulting.

Respondents were also asked if they knew someone called *Uuno*, *Tauno* or *Urpo*, and if they thought it influenced their answers. The answers naturally greatly vary between the respondents. Many respondents mentioned the effect of a public figure with the first name *Uuno*, *Tauno* or *Urpo*, or a fictional ragged movie character Uuno Turhapuro. One respondent mentioned having older relatives called *Tauno*, and had it not been for the negative connotation of the name they would have named their son *Tauno*. Another respondent said that they know someone called *Urpo* and feel pity for him, whereas some wrote that they do not consider the appellativized form and the first name as a same word at all. One respondent wrote that they know people with these names but when thinking about these people they do not associate them with the homonymous expressions. Interestingly, however, when they meet an unknown person called *Urpo*, they start thinking if the person is also *urpo*.

In summary, for some people the link between the name and the appellativized expression is clear and evident, while others do not see a connection between them or at least do not consider it significant. It is also possible that if a person does not know anyone called *Urpo*, the interpretation will be affected exclusively by the stereotypes around the name and the appellativized homonym. However, if a person knows someone called *Urpo*, the name bearer's personal features may have a greater influence than the connotations of the appellativized expression (see Reszegi 2022:212). In other words, *Urpo* becomes a part of the person's in group and the connotations of the name are thus less negative.

So far, I have used an empirical approach and analysed the meanings and the use of the three appellativized expressions and compared them with each other. I will now proceed to a more theoretical analysis. Namely, I will discuss the appellativized expressions as a vocabulary group in contrast to the other terms of abuse, and analyse their semantic profile.

5. The semantic profile of the appellativized terms of abuse

In this section, I argue that the appellativized terms of abuse have a specific semantic profile. By examining them as one category, I can form an overall view about appellativized expressions as terms of abuse.

When using appellativized expressions as terms of abuse (see the examples in 4.1), the appellativized expression could theoretically be replaced by some other term of abuse, such as the Finnish equivalent for *stupid*, *fool* or *idiot*. However, the writer has chosen to use the appellativized expression, i.e. they have conceptualized the state of affairs in a way that it can plausibly be described by appellativized expression (for conceptualization, see Langacker 2008:27–30). This implies that the meaning of such an expression is wider and more multifaceted than the meaning of a more conventional adjective or noun. In the framework of cognitive grammar, the basis of a linguistic meaning is a set of cognitive domains that an expression invokes (Langacker 2008:44). I argue that the domain matrix of an appellativized expression is more complex than that of an average term of abuse. In many contexts, all three expressions, *uuno*, *tauno*, and *urpo*, could be translated into English using the above-mentioned adjectives. However, their semantic import would be different and none of them would be their exact equivalent. The reason for this is that the adjectives characterize the feature (e.g. stupidity) directly, whereas an appellativized personal name characterizes its referent and conceptualizes the meaning indirectly, without specifying the exact feature. The appellativized personal names are thus used as cultural indexes, since cultural, contextual, and encyclopaedic knowledge is required in greater deal when interpreting the meaning of such an expression (see Reszegi 2021:126; 2022:212). As stated in Section 4.1, they are also used to imply confrontation between ‘us’ and ‘them’. According to Allen (1983:311), English uses *John* in the same way, as ‘pejorative denomination for outsiders’. For example, *John Doe* is used for an average or unknown person and *dumb John* for an easy target or victim (Allen 1983:311; see also Sjöblom 2006:72). Of course, it would be possible to refer to these with a suitable adjective or appellative but the use of an appellativized proper name makes the interpretation more complicated, as its semantic profile is more complex.

What follows, then, is that the interpretation might also vary between different readers or hearers even though the expressions have fixed semantic features (see Reszegi 2021:126). That is, when interpreting such an expression, not only are the connotations of the appellativized expression evoked in the mind of the hearer or reader, but also the connotations of the proper name. In other words, different sociolinguistic factors may affect the interpretation, for example, the expected age of the bearers of the homonymous name. In terms of cognitive grammar, the domain matrix of the homonymous proper name is also evoked in the mind when interpreting such an expression (for domains, see Langacker 2008:44–54). Connotations of the proper names vary to a greater extent than connotations of widely known and conventional adjectives or appellatives, as people may or may not know bearers of the name, which subsequently may influence how they are interpreted. This is based on the fact that this kind of connotations is highly subjective. Whether the reader does or does not perceive that the expression has proper name origin, it might also affect the connotations and interpretation. Moreover, language learners may be unaware of the word’s origin as a proper name, instead they may only be familiar with the expression as a term of abuse. Hence, the role of the word as a first name cannot affect the way it is interpreted as a term of abuse.

It is possible that people who know someone with a homonymous name do not use these expressions, especially if their relationship is close – or may at least find the use of them more impolite than other people. The same holds true also for illnesses used as terms of abuse.

For example, in Dutch the word *kanker* 'cancer' is a common term of abuse, however, people with cancer or who have a relative with cancer may find the use of this expression harsh (De Raad et al. 2005:154). As mentioned in Section 4.2, some respondents knew people called *Uuno*, *Tauno* or *Urpo*, but only a part of them think that it influenced their answers. This illustrates that some people see the first name and the corresponding appellative in a way as the same but polysemous word, whereas the others see a stricter border between the proper name and the homonymous appellative and do not consider that they are connected.

It is an interesting question whether the proper name and the appellativized form actually have something in common, except the homonymity and the origin (*urpo* < *Urpo*). Nyström (2016) discusses names with homonymous common noun origin, such as Swedish male name *Björn*, which originates from the word *björn* 'bear' (see also Khoa 2022 for Vietnamese proper names with homonymous common name origin). In this case, the relation between the proper name and the common noun is opposite than in the case of appellativization. Nevertheless, there are similarities in the processes of interpreting such names. Not only does the name *Björn* denote someone bearing this name but also activates various processes in the brain with connotations of the homonymous noun (Nyström 2016). Nyström (2016) also argues that proper names and common nouns should be seen as two communicating and integrated parts of the mental lexicon and that in their interpretation, connotations have a significant role. I agree and, consequently, argue that the link between the appellativized expression and the original first name goes both ways. Hence, the connotations of the first name may affect the interpretation of the appellativized expression's meaning, and in turn, the appellativized expression's meaning may affect the idea of how we view the first name. The popularity of the first names *Uuno*, *Tauno*, and especially *Urpo* has remained relatively low after the appellativization (Sarhemaa 2021:108–112). The reason why the name givers avoid these names might be their negative connotations, which are due to the pejorative appellativized homonym.

6. Conclusions

In this article, I have analysed three appellativized Finnish male names, namely *Uuno*, *Tauno*, and *Urpo*, which all have pejorative meanings. It has been shown that they have both common and individual meanings and uses; despite the ostensible synonymy they are not synonymous. Specific for *uuno* is that it is used to refer to speakers themselves more often than the other two expressions. In particular, it is used to ease tensions and for self-disclosure. Besides the pejorative meaning, *tauno* also has another central meaning, namely 'ordinary person'. The questionnaire respondents mentioned the meaning of 'ordinary person' more often than the pejorative meaning, however, according to the corpus data, *tauno* is more often used in pejorative meaning without the meaning of 'ordinary person'. One reason for that might be that the language on an anonymous internet forum is often offensive. Special for *urpo* is that it is used in pejorative meaning referring to other people participating in the discussion more often than the other two expressions. The questionnaire data imply that someone called *urpo* acts foolishly on purpose, whereas the stupidity of *uuno* and *tauno* is unintentional.

The findings about the pejorativity of the expressions are confirmed by the analysis of both the corpus and the questionnaire data. The most pejorative and hence the most insulting of these expressions is indisputably *urpo*. *Tauno* is the least pejorative and the least insulting – possibly because of its another central meaning 'ordinary person' – and *uuno*'s pejorativity lies between the other two expressions. Even though there are prominent differences in the pejorativity, all three expressions can be used abusively, and the use of them can make communication impolite or emphasize impoliteness.

It has been argued that even though the appellativized terms of abuse can be used in a same way as the other terms of abuse, their semantic profile is more complex. This is because the interpretation of appellativized expressions also evokes the domain matrix of the homonymous proper name in the hearer's or reader's mind. That is why the connotations of the proper name might also affect the interpretation of such an expression. Furthermore, the appellativized

NoSo 2023 | https://doi.org/10.59589/noso.32023.14416

expressions might have an impact on our impression of the first names lying behind them.

The object of this study, appellativized personal names as abusive part of communication, is under-studied. Yet, it is a relevant topic and needs further research. Personal names are indeed personal, to put it another way, a name bearer has a personal relationship to their name and a name is close to a person's identity. Therefore, it is good to keep in mind that someone may be offended if the homonym of their name is used as a term of abuse.

When researching the meanings of appellativized expressions we touch upon the question of the meaning of proper names. The question whether proper names have a meaning – other than lexical or etymological – is a widely and vividly discussed topic among the onomastics (see e.g. Coates 2006; Sjöblom 2006; Van Langendonck 2007; Nyström 2016; Reszegi 2021 & 2022; Khoa 2022). In the future, it would be worthwhile to study the meaning of proper names in relation to the appellativization and the semantics of appellativized expressions. It is possible that the connotative meanings of the proper names are more commonly shared between different people if the name has an appellativized homonym. This is because of a two-way link between the name and the appellativized expression. Another question is, of course, whether these connotative meanings can be seen as a property of the name itself, or are they only related to the appellativized expression. To conclude, there remains a need for further research to discuss these issues more thoroughly.

References

Abelin, Åsa. 2015. Phonaesthemes and sound symbolism in Swedish brand names. *Ampersand* 2. 19–29. https://doi.org/10.1016/j.amper.2014.12.001

Allan, Keith & Burridge, Kate. 1991. *Euphemism and dysphemism: Language used as shield and weapon*. Oxford: Oxford University Press.

Allen, Irving Lewis. 1983. Personal names that became ethnic epithets. *Names* 31(4). 307–317. https://doi.org/10.1179/nam.1983.31.4.307

Aller Media ltd. 2014. *The Suomi 24 Sentences Corpus* (2016H2) [text corpus]. Kielipankki. Retrieved from http://urn.fi/urn:nbn:fi:lb-2017021505 (accessed 30 April 2017)

Beirne, Piers. 2020. Animals, women and terms of abuse: Towards a cultural etymology of *con(e)y*, *cunny*, *cunt* and *c*nt*. *Critical Criminology* 28. 327–349. https://doi.org/10.1007/s 10612-019-09460-w

Bolinger, Dwight L. 1950. Rime, assonance, and morpheme analysis. *WORD* 6(2). 117–136. https://doi.org/10.1080/00437956.1950.11659374

Bousfield, Derek & Locher, Miriam A. 2008. Introduction: Impoliteness and power in language. In Bousfield, Derek & Locher, Miriam A. (eds.), *Impoliteness in language: Studies on its interplay with power in theory and practice*, 1–13. Berlin: De Gruyter Mouton.

Brezina, Vaclav, Weill-Tessier, Pierre, & McEnery, Anthony. 2020. *#LancsBox* v. 5.x. [software]. Available at: http://corpora.lancs.ac.uk/lancsbox.

Coates, Richard. 2006. Properhood. *Language* 82(2). 356–382. http://www.jstor.org/stable/4490160

Culpeper, Jonathan. 2010. Conventionalised impoliteness formulae. *Journal of Pragmatics* 42. 3232–3245.

Culpeper, Jonathan. 2011. *Impoliteness: Using language to cause offence*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511975752

Culpeper, Jonathan, et al. 2014. Impoliteness and emotions in a cross-cultural perspective. *SPELL: Swiss Papers in English Language and Literature* 30. 67–88. http://doi.org/10.5169/seals-583882

De Raad, Boele, Van Oudenhoven, Jan Pieter & Hofstede, Merle. 2005. Personality terms of abuse in three cultures: Type nouns between description and insult. *European Journal of Personality* 19(2). 153–165. https://doi.org/10.1002/per.540

Garcés-Conejos Blitvich, Pilar & Sifianou, Maria. 2019. Im/politeness and discursive pragmatics. *Journal of Pragmatics* 145. 91–101. https://doi.org/10.1016/j.pragma.2019.03.015

Harju, Auli. 2018. Suomi24-keskustelut kohtaamisten ja törmäysten tilana. *Media & viestintä* 41(1). 51–74. https://doi.org/10.23983/mv.69952

Harvalík, Milan. 2012. Appellativisation and proprialisation: The gateways between the appellative and proprial spheres of language. In Felecan, Oliviu (ed.), *Name and naming. Synchronic and diachronic perspectives*, 10–17. Newcastle upon Tyne: Cambridge Scholars Publishing.

Hinton, Leanne, Nichols, Johanna & Ohala, John. 1994. Introduction: Sound-symbolic processes. In Hinton, Leanne, Nichols, Johanna & Ohala, John (eds.), *Sound symbolism*, 1–12. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511751806.001

Huffaker, David A. & Calvert, Sandra L. 2005. Gender, identity, and language use in teenage blogs. *Journal of Computer-Mediated Communication* 10(2). https://doi.org/10.1111/j.1083-6101.2005.tb00238.x

Hämäläinen, Simo. 1956. Kynttilä otti uunon vihtorin: Erisnimien käytöstä yleisniminä. *Virittäjä* 60. 202–227.

Kawahara, Shigeto, Noto, Atsushi & Kumagai, Gakuji. 2018. Sound symbolic patterns in Pokémon names. *Phonetica* 75(3). 219–244.

Khoa, Nguyen Viet. 2022. Revisiting semantic issues of proper names: A Vietnamese perspective. *Names* 70(1). 41–49. https://doi.org/10.5195/names.2022.2378

Kielitoimiston sanakirja. 2021. Helsinki: Kotimaisten kielten keskus, online publication 35. URN:NBN:fi:kotus-201433. This publication is updated regularly. Last update 11. November 2021. https://www.kielitoimistonsanakirja.fi (accessed 22 February 2022)

Lagus, Krista, Pantzar, Mika, Ruckenstein, Minna & Ylisiurua, Marjoriikka. 2016. *Suomi24: muodonantoa aineistolle*. (Valtiotieteellisen tiedekunnan julkaisuja; Nro 10). Helsinki: University of Helsinki.

Langacker, Ronald. 1991. *Concept, image, and symbol: The cognitive basis of grammar*. Berlin: Mouton de Gruyter.

Langacker, Ronald. 2008. *Cognitive grammar: A basic introduction*. Oxford: Oxford University Press.

Muikku-Werner, Pirkko. 2005. *Ilkeilyn kahdet kasvot*. Tampere: Osuuskunta Vastapaino.

Norri, Juhani. 2000. Labelling of derogatory words in some British and American dictionaries. *International Journal of Lexicography* 13(2). 71–106. https://doi.org/10.1093/ijl/13.2.71

Nyström, Staffan. 2016. Names and meaning. In Hough, Carole (ed.), *The Oxford handbook of names and naming*. Oxford: Oxford University Press. Online edition. https://doi.org/10.1093/oxfordhb/9780199656431.001.0001

Paunonen, Heikki. 2000. *Tsennaaks Stadii, bonjaaks slangii: Stadin slangin suursanakirja*. Helsinki: WSOY.

Paunonen, Heikki. 2018. Personal email 2.2.2018.

Pälli, Pekka. 2003. *Ihmisryhmä diskurssissa ja diskurssina*. Tampere: Tampere University Press.

Reszegi, Katalin. 2018. On the proper-name-to-appellative transformation. *Rivista Italiana di Onomastica* 24(1). 25–46.

Reszegi, Katalin. 2021. The mental and neural representation of names: A cognitive and economic point of view. In Leibring, Katharina, et al. (eds.), *The economy in names: Values, branding and globalization*, 119–132. Uppsala: Department of Archives and Research, Institute for Language and Folklore.

Reszegi, Katalin. 2022. *Kognitív szemléletű névtudományi vizsgálatok*. Debrecen: Debrecen University Press.

Saarelma, Minna. 2007. *Nimi lapselle*. Helsinki: Minerva.

Sarhemaa, Maria. 2021. Tavan tauno uunoilee urpokaupungissa: Nimien Uuno, Tauno ja Urpo appellatiivistuminen ja appellatiivien käyttö Suomi24-keskustelupalstalla. *Sananjalka* 63. 103–129. https://doi.org/10.30673/sja.107278

Sarhemaa, Maria. Forthcoming. Onko uuno urpompi kuin tauno? Kolmen etunimen appellatiivistuminen kyselytutkimuksen valossa. *Virittäjä* 127(2).

Shih, Stephanie S. & Rudin, Deniz. 2021. On sound symbolism in baseball player names. *Names* 69(1). 21–34. https://doi.org/10.5195/names.2021.2245

Sjöblom, Paula 2006: A cognitive approach to the semantics of proper names. *Onoma* 41. 63–82.

Sobkowicz, Pawel & Sobkowicz, Anton. 2012. Two-year study of emotion and communication patterns in a highly polarized political discussion forum. *Social Science Computer Review* 30(4). 448–469.

Van Langendonck, Willy. 2007. *Theory and typology of proper names*. Berlin: Mouton de Gruyter.

Van Oudenhoven, Jan Pieter, et al. 2008. Terms of abuse as expression and reinforcement of cultures. *International Journal of Intercultural Relations* 32. 174–185.

Vilkuna, Kustaa. 2005 [1976]. *Etunimet*. 4., updated edition. Helsinki: Otava.

VISK = Hakulinen, Auli, Vilkuna, Maria, Korhonen, Riitta, Koivisto, Vesa, Heinonen, Tarja Riitta & Alho, Irja 2004: *Iso suomen kielioppi*. Helsinki: Suomalaisen Kirjallisuuden Seura. http://scripta.kotus.fi/visk (accessed 12 April 2022).

NoSo 3·2023

Nordisk tidskrift för socioonomastik / Nordic Journal of Socio-Onomastics (NoSo) är en plattform för forskning som utforskar namns roll i samhället och i människors sociala interaktion. Tidskriften har en bred ansats som bygger på såväl historiskt som modernt material, olika typer av metoder, och teoretiska likaväl som praktiska frågeställningar. Inriktningen är nordisk på så sätt att forskningen som presenteras ska vara intressant ur ett nordiskt perspektiv.

Nordisk tidskrift för socioonomastik / Nordic Journal of Socio-Onomastics (NoSo) is a platform for research that explores the role of names in societies and in social interaction. The journal has a broad approach, including historical as well as contemporary data, different methods, and both theoretical and practical research questions. The focus is Nordic in the sense that the research presented should be of interest from a Nordic point of view.

ISSN 2004-0881
Nordisk tidskrift för socioonomastik
Högskolan i Halmstad
Box 823
SE-301 18 Halmstad
Sweden