

Emil Stjernholm

Distant reading televised public information: The communication of Swedish government agencies, 1978- 2020

2022

<https://doi.org/10.25969/mediarep/19199>

Veröffentlichungsversion / published version
Zeitschriftenartikel / journal article

Empfohlene Zitierung / Suggested Citation:

Stjernholm, Emil: Distant reading televised public information: The communication of Swedish government agencies, 1978- 2020. In: *NECSUS_European Journal of Media Studies*. #Materiality, Jg. 11 (2022), Nr. 2, S. 201–225. DOI: <https://doi.org/10.25969/mediarep/19199>.

Erstmalig hier erschienen / Initial publication here:

<https://necsus-ejms.org/distant-reading-televised-public-information-thecommunication-of-swedish-government-agencies-1978-2020/>

Nutzungsbedingungen:

Dieser Text wird unter einer Creative Commons - Namensnennung - Nicht kommerziell - Keine Bearbeitungen 4.0/ Lizenz zur Verfügung gestellt. Nähere Auskünfte zu dieser Lizenz finden Sie hier:

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Terms of use:

This document is made available under a creative commons - Attribution - Non Commercial - No Derivatives 4.0/ License. For more information see:

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Distant reading televised public information: The communication of Swedish government agencies, 1978-2020

Emil Stjernholm

NECSUS 11 (2), Autumn 2022

<https://necsus-ejms.org/distant-reading-televised-public-information-the-communication-of-swedish-government-agencies-1978-2020/>

Abstract

This study focuses on the Swedish public information programme *Anslagstavlan*, a unique audiovisual communication tool for Swedish government agencies since 1972 and an important part of Sweden's audiovisual cultural heritage. Whereas digital research methodologies for data-driven text analysis have been developed and established over the last decades, the use of digital tools in the analysis of audiovisual sources has only recently gained increased attention. In this article, automatic speech recognition algorithms are used to extract the speech from a large sample of spoken messages in *Anslagstavlan* (1978-2020) which are then explored using digital methods for text analysis. The article argues that automatic speech recognition and corpus analysis should be viewed as a useful tool to gain an overview of a larger corpus of audiovisual media and to notice patterns and trends that would not be visible by close reading alone.

Keywords: digital humanities, media history, automatic speech recognition, public information, television history

With a smile on his face, the muscular weightlifter Lennart 'Hoa-Hoa' Dahlgren walks to a day nursery to pick up his son, whereupon fast-cut scenes show him embarking on an intense wrestling match with a group of overjoyed kids. 'It is cool to split the parental leave', the narration concludes in a casual tone of voice.[1] The Swedish Social Insurance Agency's iconic 1978 campaign for progressive parental policy is merely one example of the type of public information that reached millions of Swedes through the television programme *Anslagstavlan* (*The Bulletin Board*). The construction and consolidation of the Swedish as well as the Nordic

welfare states was shaped not only by social reforms, economic productivity, and effective governance, but also by state interventions in the media system and government agencies' communicative practices.[2] In the post-war period, for example, utopian visions were often coupled with attempts at persuasion, and communicative activities were used to steer the private life of citizens and to demonstrate the importance of certain virtues.[3]

From the start in 1972, *Anslagstavlan* emerged as a means of soft power that government agencies utilised to influence and shape people's views on topics such as health, welfare benefits, and traffic security. The short messages were designed to be easily accessible with a direct and humorous mode-of-address, and in order to ensure that the circulation of public information was maximised, the programme was regularly broadcast alongside popular programmes such as M*A*S*H, Soap, and the daily sports programme *Sportspegeln*. Meanwhile, to its critics, *Anslagstavlan* served as the ultimate manifestation of the bureaucratic, overprotective 'nanny state'. While televised public information was most influential during the days of the public service broadcasting monopoly, lasting until the late 1980s and early 1990s, the programme is still being broadcast today 50 years after its premiere. *Anslagstavlan* thus constitutes a neglected part of Sweden's audiovisual cultural heritage and a prism through which shifts in the Swedish public information tradition can be studied.

Placed at the intersection of media studies and digital humanities, this article focuses on discourses developed in televised public information over time. As large numbers of digitised audiovisual media become available on different platforms, there are new ways today to explore and understand the development of narratives and discourses in television history. While digital research methodologies for data-driven text analysis have been developed and become established over the last decades, the use of digital tools in the analysis of audiovisual sources has only recently gained increased attention.[4] Much of the digital humanities scholarship on audiovisual media builds on key works within the field of film studies on the narrative and aesthetic codes of moving images.[5] In recent years, scholars have developed computational tools such as Cinemetrics, Lignes de Temps, Shotlogger, ImagePlot, and the Distant Viewing Toolkit, which cast the evolution of film and

television styles in a new light.[6] This impetus has resulted in several significant studies in which new approaches to moving image analyses are advanced.[7] Besides this, much emphasis in digital film studies has been placed on contextual factors such as production and circulation.[8] By contrast, this study focuses less on television style and more on computer-assisted corpus analysis, specifically analysing the public information messages and discourses promoted in the television program *Anslagstavlan*.

This article is the result of a collaboration with the National Library of Sweden's digital humanities lab, KBLab. Although the National Library of Sweden's audiovisual collection is frequently used in media studies scholarship, no previous research project has utilised automatic speech recognition to analyse material from their collection. In this study, then, automatic speech recognition algorithms were used to extract the speech from videos in the dataset. Following this, the digital text analysis methods were used to perform corpus analysis in an effort to identify trends and patterns in audiovisual public information which would not be detectable through qualitative analysis. This study thus takes a novel approach to Swedish television history. The following key research questions will be addressed: What type of information did government agencies communicate through television and how did the main themes shift over time? What do these shifts say about the development of the Swedish public information tradition? Given that many challenges remain in turning sound and vision into data,[9] this study also constitutes a contribution toward methodological development.

Mapping government agency communications

Book historian Robert Darnton argues that every age is an age of information – moulded by the communication systems available at the time.[10] While the concept of public information (*samhällsinformation*) first occurs in the public debate of the 1950s in relation to the introduction of television,[11] it becomes more pervasive in the public debate of the 1960s and 1970s. At this time, public information about topics such as consumers' rights,[12] health,[13] and social problems[14] was frequently portrayed as a way to deepen citizen knowledge about their rights

and responsibilities in society. The media landscape of the early 1970s saw television emerging as the dominant medium, something which a Swedish Government Official Report on the future of public service radio and television took particular notice of.[15] With the expansion of the public sector, the report noted, there was an increasing demand for public information on the responsibilities of society – and television was seen as a promising ‘new medium’ in this context. A wide assortment of societal issues in dire need of increased public awareness were highlighted: ranging from rules and regulations, social issues, employment and social affairs, health issues, environmental issues, to issues concerning local associations or minority groups. As Monika Djerf-Pierre and Lennart Weibull observe, the prevailing notion was that there were so many positive facets of society that simply ‘deserved better treatment on radio and television’.[16]

Within film and media studies, the past decade has seen a growing interest in the broad spectrum of ‘useful’ audiovisual media that were produced with a specific purpose in mind such as to inform, to educate, or to sell.[17] However, televised public information spots – what in the US context is usually labelled *public service announcements* and in the UK context *public information films* – remains an under-researched phenomenon. Along with trailers and commercials, television scholar John Ellis notes, public service announcements constitute ‘interstitials’, which is ‘a whole class of television output’ that carries messages, builds anticipation and delay, and interlaces with other programs.[18] While some research exists on producers of televised government information spots, focusing on influential organisations and institutions such as non-profit The Advertising Council (1942-) in the US and the Central Office of Information in the UK (1946-2011),[19] less emphasis has been placed on this particular type of state interference in the Scandinavian media system.[20] Notably, the perceived need for expanded public information was not limited to Sweden, but dedicated programs with short public information spots also circulated elsewhere in Europe, such as in Denmark (*OBS - Oplysning til Borgerne om Samfundet*) and the Netherlands (*Postbus 51*) during the 1970s.

DISTANT READING TELEVISED PUBLIC INFORMATION: THE COMMUNICATION OF SWEDISH GOVERNMENT AGENCIES, 1978-2020



Fig. 1: An example of televised public information on Anslagtavlan. The spot was commissioned by the Swedish Social Insurance Agency in 1978. Captions translated by the author.

Following a Swedish Government Official Report on the expansion of public information, the Swedish Parliament assigned the Swedish Radio Administration (SR) new directives that granted government agencies the right to communicate messages on television – following the so-called ‘channel split’ (*kanalklyvningen*) – resulting in a new administrative entity within SR called the Working Committee for Public Information (*Beredningsgruppen för myndighetsmeddelanden*).[21] While much research has been devoted to the Swedish public service model, wherein regulated independence from the government has been described as a cornerstone,[22] Swedish television’s function as a communication tool for the government authorities has received little attention in previous scholarship.

The approach of this study is exploratory and inductive. As digital historian Heather Froehlich argues: ‘Corpus analysis is a form of text analysis which allows you to make comparisons between textual objects at a large scale ... It allows us to see things that we don’t necessarily see when reading as humans.’[23] Famously, Franco Moretti has argued that distant reading ‘allows you to focus on units that are much smaller or much larger than the text: devices, themes, tropes – or genres

and systems'.[24] Since the audiovisual dataset used for this study is more extensive than most traditional film and television studies corpora, computer-aided analysis can open up for new perspectives. As Eric Hoyt et al argue in a chapter on the interpretation of media history's big data, '[t]he true strength of computational approaches is encouraging researchers to read their evidence in different ways – in quantitative ways, yes, but also ways that highlight the marginal, the unexpected, and the forgotten.'[25] While linguistic scholarship on film and television narratives is a growing subfield,[26] such studies tend to focus on canonical examples rather than on the 'useful media' that Charles R. Acland and Eric Hoyt call 'the great unread' of media studies.[27] Drawing on these epistemological perspectives, this study aims to highlight continuity and change in the tradition of televised public information.

Methods and material

As stated, the primary dataset in this study is the recurring television segment *Anslagstavlan*. The program aired twice a week with each spot lasting between 30 seconds and one minute, and each program containing between three to five messages. While statistics concerning circulation are scarce, spots usually circulated between two to four weeks before they were substituted. The digitised content is held at the National Library of Sweden (Kungliga Biblioteket, KB). During the past five decades, some 4,000 *Anslagstavlan* programmes have been preserved, equivalent to approximately 330 hours of television, from which a smaller sample of 500 programmes are utilised in this analysis. Although SVT began broadcasting the programme in 1972, this study takes its starting point in 1978, the year the legislation concerning the deposit of audiovisual material to KB changed, making legal deposit of audiovisual material mandatory. From this year on, there is almost full coverage, whereas the preserved, digitised material has considerable gaps from the period 1972–1978. The complete textual dataset of these programmes is of mid-range quantity, 177,088 tokens in total, making it suitable for corpus analysis (rather than tools designed to handle vast amounts of data).

The cultural technological awareness of speech recognition has been shaped by popular culture, primarily through science fiction films such as Stanley Kubrick's

2001: A Space Odyssey, in which the intelligent computer HAL is able to understand spoken words and respond in a natural-sounding voice, and the intelligent machines in the *Star Wars* franchise, R2D2 and C3PO. Today, automatic speech recognition is one type of automation integral to what has been labelled 'digital society'.^[28] In this sense, speech recognition technologies make it possible for machines to provide a range of useful services for both commercial companies and the public sector. This development has not gone unnoticed by film and broadcasting archives. As Adelheid Heftberger points out, the interplay between audiovisual archives and researchers working with digital methods is of growing importance today.^[29] The present study has been conducted in collaboration with technical staff at the National Library of Sweden's digital humanities lab, KB-Lab. Currently, KB is developing new automatic speech recognition models trained on historical Swedish media material. A key long-term ambition is to make the vast quantities of audiovisual material deposited at KB computable and searchable, increasing the availability and usefulness of the audiovisual archive to scholars and students alike.^[30] Internationally, similar ambitions have been voiced with regard to broadcasting archives, such as the Netherlands Institute for Sound and Vision's Media Suite.^[31] The Wav2Vec model used for this particular study, labelled Fine-tuning XLSR-Wav2Vec2, has primarily been trained on Swedish language radio material and common voice data by the KB computer scientists.^[32] In this sense, the performance of the speech-to-text model on televised public information was unknown at the start of the project.

Previous research employing automatic speech recognition shows that the data yielded tends to require time-consuming and costly manual cleaning.^[33] Accordingly, during the first preparatory phase, the text lists produced by automatic speech recognition were manually reviewed in order to detect errors and defects. During this process, numerous anomalies were detected, such as the mislabelling of noise as words, patchy processing of dialogue heard in the background of the chief narration, the occurrence of non-Swedish words or phrases, and the misprocessing of children's speech. In this sense, similar to the process of OCR scanning of newspapers,^[34] the automatic speech recognition process generated digital noise and misinterpretations, highlighting that the speech-to-text rendition does not encapsulate all relevant information. With the help of a transcriber, the above-

mentioned anomalies were remedied through a manual cleaning operation. Moreover, in order to make the results as useful as possible, a Swedish-language list of stop words was used.[35]

In the analysis, emphasis is placed on stylistic analysis and corpus comparison, which was carried out using the tool AntConc.[36] While AntConc offers a range of analytical perspectives, primary focus has been placed on word frequency, keyness, keywords in context, and collocates. Word frequency simply ranks every unique word type by its frequency in the corpus. Keyness can be described as the frequency of a word in a text when compared with its frequency in a reference corpus. In this case, a smaller corpus (e.g., one decade of public information spots) has been compared with all available transcripts of public information spots. Notably, few programmes from the 1970s have been preserved, and thus the keyness of these terms corpus is not representative of the entire decade. AntConc also allows the scholar to explore keywords in context. By typing a manual search for a term in the search box, the concordance view shows every time the word appears in the corpus of television spots.

Zooming in on keywords in context is a useful way to search for patterns. For example, which words appear near words such as ‘men’, ‘women’, and ‘children’? Having looked at keywords in context, you can then generate a list of words appearing most frequently in the company of the keywords, so-called collocates. In contemporary digital humanities, fine-grained analytical methods have been established that allow scholars to add additional layers of complexity to textual collections, for instance through sentiment analysis or named entity recognition. While such approaches go beyond the scope of this particular study, which focuses on a broader historical development decade-by-decade, the dataset and the generated text lists are accessible at the National Library of Sweden’s KB-Lab and in that sense, they are available for other scholars to experiment with.

Although automatic speech recognition can be utilised to facilitate exciting ways to explore audiovisual media, we need to continuously ask critical questions about the methods and the tools we use. In relation to the concept of distant reading, scholars have zoomed in on theoretical pitfalls as well as possible methodological

biases. For example, some have argued that the reduction of cultural objects to metrics and statistics can be seen as a 'mask of objectivity'.^[37] Indeed, a quantitative, distant reading approach to television studies entails certain weaknesses. For example, such an approach does not place the relationship between spoken dialogue and visual style in the foreground. While television has received little attention in previous digital humanities scholarship, there are interesting examples where visual style in television is interrogated using the concept of 'distant viewing'.^[38] However, as Joanna Byszuk argues, the combination of audio and visual analysis remains rare: 'Comprehensive studies examining all multimodal aspects of films and TV series still belong to the future.'^[39]

Another methodological caveat is that the utilisation of automatic speech recognition and corpus analysis in film and media studies requires institutional support. In this context, the National Library of Sweden's computer scientists developed and trained the speech-to-text model applied to *Anslagstavlan*. Another aspect to take into consideration is the fact that the process requires time-consuming manual cleaning, and in this particular case all work with the preparation of audiovisual media, the automatic speech recognition, the cleaning of the data, and the corpus analysis on the resulting text lists had to be conducted within the premises of the National Library of Sweden's audiovisual department in Stockholm due to copyright reasons. Despite this, automatic speech recognition offers exciting avenues for novel research trajectories. For example, as Hannu Salmi suggests in *What is Digital History?*, automatic speech recognition could be used to increase accessibility to audiovisual sources and help enhance the metadata that already exists.^[40] Or perhaps even more importantly, automatic speech recognition offers the opportunity to study culturally influential film and television genres – from newsreels to sponsored films, from public service announcements to news reports – from a distance, thereby offering a glimpse of the discourses formed in and shaped by them.

A distant reading of televised public information

In 1969, a Swedish Government Official Report highlighted a growing need for public information in society. Citizens deserved to know more about their rights and responsibilities and the expansion of the public sector put new demands on

the services offered by government agencies and municipalities.[41] Televised public information delivered through the public service broadcaster Sveriges Radio (SR) was underlined as a particularly important means of mass communication:

TV captures the viewers' attention easily, which compensates for its ephemerality. A merry brochure, distributed widely, has a low attention value simply because it is drowned out by other forms of print advertising, which obscure it.[42]

An internal report within SR noted that government agency interest in televised public information was 'exceptionally great' and that certain topics had been underlined as particularly important, such as questions concerning the labour market, health, welfare, migration, and traffic security.[43]

But what type of information actually came to dominate televised public information and how did the themes of this information change over time? Drawing on the corpus comparison discussed above, it is possible to trace words that are statistically overrepresented in the smaller corpus than in the larger reference corpus, thus shining light on how discourses have shifted over time. Particularly during the 1970s and 1980s, *Anslagstavlan* was accused of preaching from the top down, relying heavily on coercion and the communication of constraints.[44] In other words, practices that behavioural economist Richard Thaler and law scholar Cass Sunstein would label 'paternalism'.[45] During the late 1970s, some of the most overrepresented words are 'swimming', 'right of public access', 'helmet', 'car inspection', 'moped', and 'traffic', of which several relate to specific welfare, health, and safety measures. A similar trend can be noted during the 1980s, when words such as 'housing benefit', 'traffic', 'alcohol', and 'speed' are among the most overrepresented words. From among the key themes television officials identified as particularly important during the launching of *Anslagstavlan*, words relating to the themes of health, welfare, and traffic security are clearly overrepresented in the 1970s and 1980s, whereas words relating to the labour market and migration are absent. Other overrepresented words, such as 'energy' during the 1970s or 'the Swedish Social Insurance Agency' during the 1970s and the 1980s, likely relate to specific large-scale campaigns that took place over multiple years. With regard to energy, the Oil Crisis of the 1970s led to the

formation of the Swedish Energy Conservation Committee, which combined policy measures with a wide range of information activities on the energy economy. Similarly, in 1974, the Swedish parental leave reform gave fathers the same rights as mothers to use parental leave, and this was a topic that the Swedish Social Insurance Agency produced several information spots on.

Figure 1. The 20 most overrepresented terms in *Anslagstavlan* per decade.

Time period	Terms (by falling statistical overrepresentation)
1978–1979	swimming, september, letter, car inspection, right of public access, helmet, property, brochure, moped, swedish, voting card, save, law, vote, the swedish social insurance agency, energy, year, post, meeting, parental leave
1980–1989	housing benefit, brochure, car, year, read, post, land, speak, alcohol, the swedish social insurance agency, speed, crowns, municipality, message, remember, february, employer, income, calm, stop
1990–1999	call, blood, eu, information, weeks, ouch, krastapolis, need, blood donor, countries, april, help, radio, easy to read, conditions, television, the swedish social insurance agency, telephone, books, children
2000–2009	information, www, cancer education, find, discrimination, internet, vhs, life, toys, train, children, welcome, save, ask, packaging, trial, afraid, website, the country board, benefit
2010–2019	violence, fish, around, knowledge, find, women's safety hotline, questions, call, need, precondition, risk, sun, women, damages, number, special pedagogy, student, reach, search, toxins

Fig. 2: The 20 most overrepresented terms in *Anslagstavlan* per decade.

Other overrepresented words hint at large-scale political and social policy shifts in Swedish society. During the 1990s, for example, one of the most overrepresented words is the EU (European Union), which Sweden joined in 1995 following an election preceded, and followed, by intense information campaigns from the government authorities on topics related to travel, health, and the environment. Looking at key words in context, it becomes apparent that many of the spots on *Anslagstavlan* about the EU focused on advertising different information activities, such as the EU information hotline, the EU programme office, and widely distributed brochures with information about the EU. Interestingly, during the 2000s, ‘discrimination’ emerges as an overrepresented word. A Swedish government

agency dedicated to equal rights (Jämställdhetsombudsmannen) had been established decades previously, during the 1980s. The agency monitored the Equal Rights Law (Jämställdhetslagen), which came into effect in 1991 and prohibited discrimination based on gender. Between 2000 and 2007, lawyer and public intellectual Claes Borgström headed the agency. It is noteworthy that during Borgström's leadership, Jämställdhetsombudsmannen was actually criticised for being too active in the shaping of public opinion, for example by a member of parliament from the Liberal Party[46] and by the right-wing think tank Timbro.[47] Around this time, a government report addressed the issue of government agencies overstepping their authority in the shaping of public opinion, highlighting this as a potential democracy problem.[48] Zooming in on the key words in context, it becomes apparent that the public information spots related to this theme centred on a wide variety of topics such as discrimination due to sexuality, disability, or pregnancy. Similarly, a large-scale marketing campaign for a state-supported women's safety hotline contributed to making words such as 'violence', 'women's safety', and 'the women's safety hotline' some of the most overrepresented in the 2010s.

In the last two decades, a greater emphasis seems to be placed on health communication than in previous years. For example, we see a range of words related to information spots urging citizens to make better decisions – on specific topics such as 'fish' (which you should eat and which to avoid), 'sun' (the use of sunscreen and how much time to spend in the sun), and 'toxins in the environment' (how to avoid them). Thaler and Sunstein famously argue that creative *nudges* and a well-balanced choice architecture are more effective than commands, requirements, and prohibitions in influencing human behaviour.[49] In the *Anslagstavlan* spots on health communication from the 2000s and onwards, little emphasis is placed on prohibitions. Instead, these spots highlight moderation and the making of informed decisions. For example, using an app in which you can set an alarm that warns you when you have, according to official guidelines, spent too much time in the sun. Zooming in on keywords in context, it is noticeable that the campaigns of behaviour change are directed at the individual, for example by using a personal address such as 'check how long *you* can be in the sun' or 'if *you* only use enough sunscreen'.

Greater emphasis on individualisation has simultaneously been noted as a trend in Sweden from the 1990s onwards, with society undergoing numerous neo-liberal economic reforms.[50] The overrepresentation of words related to health choices could be interpreted as indicative of this broader trend and the adherence of government agencies to more liberal ideals in public information. Meanwhile, the most frequently used verbs that convey requests (rather than overrepresented verbs) also highlights this tendency. For example, during the 1970s and 1980s, some of the most frequently used verbs are related to citizens' responsibilities, such as 'save', 'vote', and 'stop', whereas in the later decades recurring used verbs like 'find', 'search', and 'ask' do not have the same imperative tone. Similarly, the use of the word 'must' (*måste*), a modal verb used to indicate that it is important or necessary for something to happen, was more frequent during the 1970s and 1980s than during later decades, mirroring a transformation of the tone in government agency communication.

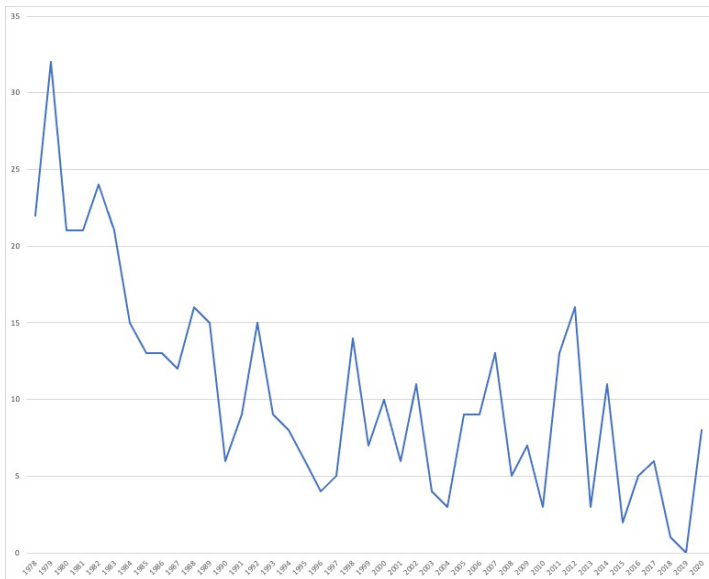


Fig. 3: Word frequency of the modal verb must (*måste*).

Zooming in on themes and stylistic development

By studying keywords in context and collocates, issues related to gender in televised public information also come to the fore. First of all, there is an obvious difference in terms of how frequently different groups are addressed in *Anslagstavlan*. Whereas the word ‘men’ features 45 times, ‘women’ features 205 times. Zooming in on the keywords in context, it is noteworthy that topics such as compulsory military service, the Home Guard, civil defence, and emergency service feature frequently in relation to the word ‘men’. By contrast, the public information mentioning ‘women’ centres on topics such as discrimination, men’s violence against women, pregnancy, and equal rights. While these results are hardly surprising, it is nevertheless interesting to note how strong of a component gender is in public information. While themes related to ‘men’ are associated with the defence of the nation, conveying typical masculine attributes such as power and authority,[51] women are framed primarily in relation to victimhood and maternity. Notably, children are the focus of more spots on *Anslagstavlan* than both men and women – featured 317 times in total. Here, the concordances show that these segments tend to focus on words such as ‘accidents’, ‘dangers’, ‘risks’, ‘harm’, ‘life-threatening injuries’, and ‘support’. A closer analysis of gender dynamics, however, would also require looking more closely at pronouns (such as he/him/his or she/her/hers).

Interestingly, references to other media are among the most overrepresented words in each decade. For example, the word ‘brochures’ occurs frequently during the 1970s and 1980s; ‘call’, ‘radio’, ‘television’, and ‘telephone’ are overrepresented during the 1990s; and ‘www’, ‘internet’, and ‘website’ are overrepresented words after the turn of the millennium. This underlines that *Anslagstavlan* was part of a media ensemble including both old and new media. While the report ‘Expanded Public Information’ includes ambitious plans for the use of ‘new media’, such as television, and optimistic visions about the impact of televised public information, the overrepresentation of these terms indicates that televised information did not replace older media forms such as telephone counselling, brochures, or paper forms, but rather complemented and drew attention to them. A

distinctive shift seems to take place in the early 2000s, when references to the internet become more frequent. Meanwhile, 'call' remains one of the most overrepresented words in the 2010s as well, highlighting that the larger media system of public information comprises a wide range of media forms.

So far, the analysis has focused mostly on the main themes in Swedish televised public information. But how has the style and tone of *Anslagstavlan* changed over time? As advertising historian Einar Korpus notes, the demarcation between consumer information, public information, and advertising became less clear in the 1970s, as facts, figures, and statistics dominated both commercial and public information newspaper advertisements.[52] One way to gauge this stylistic development is by focusing on the vocabulary and the lexical variety of the corpus. To do so, emphasis can be placed on the type-token ratio, which measures the vocabulary variation within a written text. The type-token ratio is calculated by dividing the total number of types by the total number of tokens (*100). In this case, there is a measurable change toward lesser lexical variety from the 1970s and after. The move toward lesser lexical variety thus confirms previous findings on government agency communications during the 1970s and 1980s, which indicates that archaic words became less frequent, sentences became shorter, and the overall language style became more colloquial.[53] In the *Anslagstavlan* corpus, the word 'requisition' (*rekvirera*) appears infrequently in the 1970s and 1980s, but then disappears completely during the 1990s when the more casual 'order' (*beställa*) becomes dominant, an example that hints at this development. The fact that the lexical variety was greatest in the 1970s and has been consistently simplified since then testifies to the consistent work of government agencies on this matter. In 1983, the government appointed a committee to investigate language use in the cabinet office and government agencies, which resulted in the government report 'Plain Talk: A Basis of Good Public Service'.[54] Another influential step was when the so-called 'plain talk reform' (*klarspråksreformen*) became law in 2009, mandating that public talk in Swedish should be simple and understandable (2009: 600).

In this context, it should also be noted that the move toward colloquial speech coincided with a greater emphasis on visual communication in advertising.[55] This was also the case on *Anslagstavlan*, where the omnipresent voice-of-God narration

became less of a dominant trope as the producers of televised public information drew inspiration from the advertising world. In the 1990s, the public service monopoly was broken up with the introduction of commercial broadcasting, with the satellite channel TV3 beginning to broadcast in autumn 1986 (from the United Kingdom) and the first commercial terrestrial television channel TV4 in 1992.[56] Already in the early days of commercial television, government agencies such as the Swedish Tax Agency and the Swedish Social Insurance Agency bought advertisement spots which were significantly shorter than the spots on *Anslagstavlan* (usually 20-30 seconds as opposed to 60 seconds), rapidly cut, and to a high degree inspired by contemporary television advertisement visual aesthetics. Rather than lengthy presentations of facts and figures – presented by a narrator – the main message started to come in the form of short slogans, like the Swedish Tax Agency’s ‘Relocation notification is free’ or the Swedish Social Insurance Agency’s ‘Everyone is insured with us’. This aesthetic development surely also had an impact on the development of the lexical variety in Swedish public information.

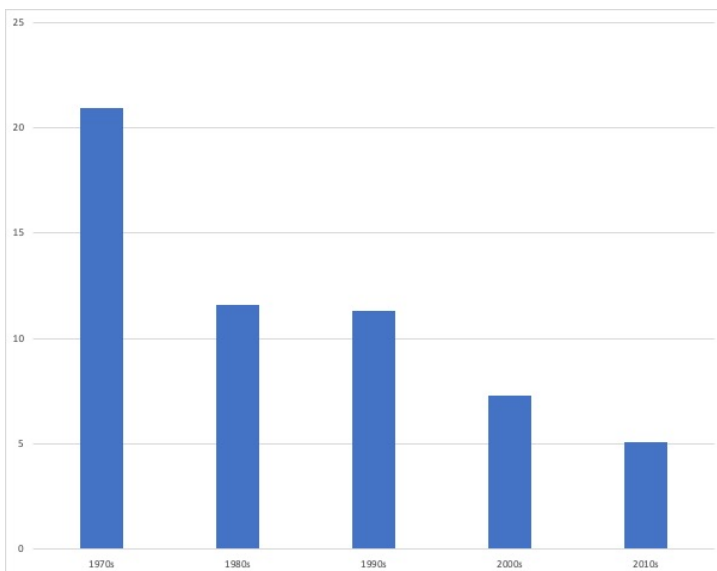


Fig. 4: Type-token ratio in *Anslagstavlan*.

Conclusions

In the past years, the visual turn within digital humanities research has highlighted that the field is text-heavy while advocating for new ways of approaching audio, visual, and audiovisual media – from Manovich's notion of *cultural analytics* to Arnold and Tilton's concept of *distant viewing*. In light of this, working with automatic speech recognition to turn decades of television programmes into text might seem counterintuitive. By necessity, such an approach becomes text-centred, rather than focused on visual characteristics, aesthetic form, or processes of mediation. Nevertheless, automatic speech recognition for small languages such as Swedish provides researchers with new ways of analysing audiovisual media. Previous research on government agencies communication efforts often comprises case studies of particularly iconic and influential campaigns. As Moretti has argued, working with the concept of *distant reading* allows you to shift focus to themes, tropes, genres, and systems. When performing a distant reading of a large collection of televised public information spots on Swedish television, several measurable trends and themes emerge. For example, internal reports within the public service broadcaster Sveriges Radio underlined questions concerning the labour market, health, welfare, migration, and traffic security as particularly key themes. However, the most overrepresented words during the late 1970s and 1980s were primarily related to specific health and safety measures, while themes relating to the labour market and migration were underrepresented. This result is surprising given the fact that the information problems related to these topics were used to justify the launching of *Anslagstavlan* in the first place.

As this study shows, some themes appear regularly – such as topics related to traffic safety ('car', 'speed', 'damages', etc.) – which is due to the seasonality of public information, where messages about the importance of winter tires appear frequently in the winter months. Other themes and tropes have come and gone. For example, the Oil Crisis of the 1970s and the early 1980s led to a massive information campaign by the Swedish Energy Conservation Committee, something which the overrepresentation of the word 'energy' in the corpus reflects. Other overrepresented words hint at the political and social policy shifts in Swedish so-

ciety, which required large-scale information campaigns. For example, the progressive parental leave reform makes the Swedish Social Insurance Agency stand out as the most frequent producer of televised public information spots during the 1980s, and Sweden's EU membership and the campaigns surrounding it makes this topic stand out during the 1990s. While the themes notable during the 2000s and 2010s are varied, the issues of discrimination, equality, and men's violence against women stand out.

Meanwhile, in terms of style and genre, it is noteworthy that there is a measurable change toward lesser lexical variety from the 1970s and onwards. During the early days of televised public information, facts, figures, and statistics dominated in public information. As previous scholarship points out, the 1970s and 1980s saw a general information trend with government agencies adopting a more colloquial language style. A government report in 1985 and the later implementation of the language law in 2009 further underlined the importance of simplified language, a tendency reflected in the distant reading of *Anslagstavlan*. Changes in the media landscape – most importantly the introduction of commercial television – likely also spurred a greater emphasis on visual communication and easy slogans in public information.

Over time, the genre of public information also undergoes changes. For example, during the 1970s and 1980s, the most overrepresented verbs in the corpus clearly relate to citizens' responsibilities, whereas in later decades the most overrepresented verbs do not have an imperative tone. This can be interpreted as indicative of a broader societal trend toward individualisation and less paternalistic ideals in public information. Or alternatively put, in the wake of neoliberal privatisation reforms, the public sector increasingly addresses the citizen as a customer. Meanwhile, it should be noted that despite government agency adoption of a more colloquial tone and informal language style, the state could still take paternalistic measures and formulate repressive information campaigns, something which the heavily criticised Swedish AIDS-awareness campaigns during the 1980s particularly testifies to.[57]

Automatic speech recognition and corpus analysis should be viewed as a useful tool to explore and gain an overview of a larger corpus of audiovisual media. Most importantly, the method of distant reading allows us to notice patterns and trends that are not visible by close reading alone. Meanwhile, the benefits of mixed-methods research, combining quantitative and qualitative perspectives, have been increasingly highlighted in recent years.[58] As underlined in the analysis, the process of distant reading entails both zooming out and zooming in, for example when examining key words in context related to questions concerning gender. Drawing on this, the distant reading carried out in this study should be understood as an exploration and as basic research that could help with the formulation of new and unforeseen research questions in the future.

Author

Emil Stjernholm is an Assistant Professor in Media and Communication Studies at the Department of Communication and Media, Lund University and a visiting researcher at the Department of Communication at the University of Copenhagen. Currently, he is working on two digital humanities research projects: *Televising Information: Audiovisual Communication of Swedish Government Agencies* (Swedish Research Council, 2020–2023) and *Modern Times 1936* (Riksbankens Jubileumsfond, 2022–2025, <http://modernatider1936.se/en>). His areas of research include digital history, propaganda studies, and television history. Together with Fredrik Norén and Claire C. Thomson, he has recently edited the volume *Nordic Media Histories of Propaganda and Persuasion* (Palgrave Macmillan, 2022). He has previously published articles in journals such as *Media History*, *Historical Journal of Film, Radio and Television*, *Studies in European Cinema*, and *VIEW Journal of European Television History and Culture*.

Acknowledgements

This work was supported by the Swedish Research Council (ref 4.3-2019-06414).

ORCID:

Emil Stjernholm <https://orcid.org/0000-0001-9871-5191>

References

- Acland, C.R. and Hoyt, E (eds). *The Arclight guidebook to media history and the digital humanities*. Sussex: REFRAME Books, 2016.
- Acland, C.R. and Wasson, H (eds). *Useful cinema*. Durham: Duke University Press, 2011.
- Acketoft, T. 'A parliamentary question: The politization of JämO', 14 October 2005: https://www.riksdagen.se/sv/dokument-lagar/dokument/interpellation/jamos-politiserade-roll_GT1043.
- Anthony, L. *AntConc* (Version 3.5.9). Tokyo: Waseda University, 2020: <http://www.laurenceanthony.net> (accessed 22 October 2021).
- Arnold, T. and Tilton, L. 'Distant viewing toolkit: A python package for the analysis of visual culture', *Journal of Open Source Software*, Vol. 5, No. 45, 20 January 2020.
- Arnold, T., Tilton, L., and Berke, A. 'Visual Style in Two Network Era Sitcoms', *Journal of Cultural Analytics*, Vol. 4, No. 2, 2019: 1-28.
- Ascari, M. 'The Dangers of Distant Reading: Reassessing Moretti's Approach to Literary Genres', *Genre*, Vol. 47, No. 1, 2014: 1-19.
- Baldwin, P. *Disease and democracy: The industrialized world faces AIDS*. Berkeley: University of California Press, 2005.
- Bednarek, M. *Language and television series: A linguistic approach to TV dialogues*. Cambridge: Cambridge University Press, 2018.
- Bednarek, M. *The language of fictional television: Drama and identity*. New York: Continuum, 2010.
- Berggren, H. and Trädgårdh, L. *Är svensken människa? Den svenska individualismens historia (Is the Swede a Human? The History of Swedish Individualism)*. Stockholm: Norstedts, 2014.
- Bordwell, D. *Narration in the fiction film*. London: Methuen, 1985.
- _____. *On the history of film style*. Cambridge: Harvard University Press, 1997.
- Butler, J. 'Statistical Analysis of Television Style: What Can Numbers Tell Us about TV Editing?', *Cinema Journal*, Vol. 54, No. 1, Fall 2014: 25-44.
- Byszuk, J. 'The Voices of Doctor Who – How Stylometry Can be Useful in Revealing New Information About TV Series', *DHQ: Digital Humanities Quarterly*, Vol. 14, No. 4, 2020: 1-23.
- Connell, R.W. *Masculinities*. Cambridge: Polity Press, 2005.

- Csomas, E. and Young, R. 'Language use in pop culture over three decades: A diachronic keyword analysis of Star Trek dialogues', *International Journal of Corpus Linguistics*, Vol. 26, No. 1, 9 February 2021: 71-94.
- Dahlgren, P. M. Swedish stop words. *GitHub*.
<https://github.com/peterdalle/svensktext> (accessed 21 January 2022)
- Darnton, R. 'An Early Information Society: News and the Media in Eighteenth-Century Paris', *The American Historical Review*, Vol. 105, No. 1, 1 February 2000: 1-35.
- Djerf-Pierre, M. and Weibull, L. *Spegla, granska, tolka: Aktualitetsjournalistik i svensk radio och TV under 1900-talet (Mirror, Investigate, Interpret: News Journalism on Swedish Radio and TV during the 20th Century)*. Stockholm: Prisma, 2001.
- Djerf-Pierre, M. and Ekström, M (eds). *A history of Swedish broadcasting: Communicative ethos, genres and institutional change*. Göteborg: Nordicom, 2013.
- Elhamifar, E. and Naing, Z. 'Unsupervised procedure learning via joint dynamic summarization' in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, Computer Vision Foundation, 2019: 6341-6350.
- Ellis, J. 'Interstitials: How the "Bits in Between" Define the Programmes' in *Ephemeral media: Transitory screen culture from television to YouTube*, edited by P. Grainge. Basingstoke: Palgrave Macmillan, 2011.
- Elsässer, S. *Att skapa en konsument: Råd & Rön och den statliga konsumentupplysningen (Creating a Consumer: Advice and findings and the government consumer report)*. Göteborg: Makadam, 2012.
- Fickers, A., Williams, M., and Snickars, P. 'Audiovisual Data in Digital Humanities', *VIEW. Journal of European Television History and Culture*, 31 October 2018, Vol. 7, No. 14: 1-4.
- Flueckiger, B. 'A Digital Humanities Approach to Film Colors', *The Moving Image*, Vol. 17, No. 2, Fall 2017: 71-94.
- Fournout, O., Beaudouin, V., and Ferrarese, E. 'De l'utopie numérique à la pratique: le cas de l'annotation collaborative de films' (From Digital Utopia to Practice: The case of collaborative video annotation), *Communication & Langages*, Vol. 180, No. 2, 1 June 2014: 95-120.
- Froehlich, H. 'Corpus Analysis with Antconc', *Programming Historian*, 19 June 2015: <https://doi.org/10.46430/phen0043> (18 November 2021).
- Griffith, R. 'The Selling of America: The Advertising Council and American Politics, 1942-1960', *The Business History Review*, Vol. 57, No. 3, 1983: 388-412.

- Harvard, J and Stadius, P. 'A Communicative Perspective on the Formation of the North: Contexts, Channels and Concepts' in *Communicating the north: Media structures and images in the making of the Nordic region*, edited by J. Harvard and P. Stadius. Farnham: Ashgate, 2013: 1-24.
- Heftberger, A. *Digital humanities and film studies: Visualising Dziga Vertov's work*. New York: Springer, 2018a.
- _____. 'Exploring the Moving Image: The Role of Audiovisual Archives as Partners for Digital Humanities and Cultural Heritage Institutions' in *Digital humanities, libraries, and partnerships: A critical examination of labor, networks, and community*, edited by R. Kear and K. Joranson. Oxford, Chandos Publishing, 2018b: 45-57.
- Hirdman, Y. *Att lägga livet till rätta: Studier i svensk folkhemspolitik* (Putting Life in Order: Studies on Swedish Welfare State Policy). Stockholm: Carlsson, 1990.
- Hoyt, E., Tran, T., Long, D., Hughes, K., and Ponto, K. 'Searching, Mining and Interpreting Media History's Big Data' in *The Routledge companion to media studies and the digital humanities*, edited by J. Sayers. London: Routledge, 2018.
- Jarlbrink, J. and Snickars, P. 'Cultural heritage as digital noise: nineteenth century newspapers in the digital archive', *Journal of Documentation*, Vol. 73, No. 6, 2017: 1228-1243.
- Jönsson, M. and Snickars, P. *Medier och politik: Om arbetarrörelsens mediestrategier under 1900-talet* (Media and politics: On the media strategies of the workers' movement during the 20th century). Lund: Mediehistoriskt arkiv, 2007.
- Kanatova, M., Milyakina, A., Pilipovec, T., Shelya, A., Sobchuk, O., and Tinitis, P. 'Broken Time, Continued Evolution: Anachronies in Contemporary Films', *Stanford Literary Lab*, No. 14, 1 April 2017: 1-20.
- Korpus, E. *Reklamiska: Guldäggsannonser 1975–2007* (Advertising Talk: Golden Egg Ads 1975–2007). Örebro: Örebro universitet, 2008.
- Lindgren, S. 'Introducing Connected Concept Analysis: A Network Approach to Big Text Datasets', *Text & Talk*, 36:3, 2016: 341-362; <https://doi.org/10.1515/text-2016-0016>.
- Lindgren, S. *Digital media and society*. Thousand Oaks: SAGE, 2017.
- Malmsten, M., Haffenden, C., and Börjeson, L. 'Hearing voices at the National Library – a speech corpus and acoustic model for the Swedish language', 2022. *arXiv preprint arXiv:2205.03026*.
- Manovich, L. *Cultural analytics*. Cambridge: MIT Press, 2020.
- Mårtensson, E. 'Den familjära myndigheten: Intimiseringen av det offentliga språket' (The Familiar Authority: The intimisation of the public language) in

Hej, det är från försäkringskassan! Informaliseringen av Sverige (Hi, this is from the Swedish Social Insurance Agency! The Informalisation of Sweden), edited by O. Löfgren. Stockholm: Natur & Kultur, 1988:105-127.

Möller, T. 'Politiseringen av statsförvaltningen' (The Politicisation of the State Administration) in *Makt: om Sveriges demokratiska underskott (Power: On Sweden's Democratic Deficit)*, edited by F. Segerfeldt and M. Rankka. Stockholm: Timbro, 2006.

Noordegraaf, J., Lotze, K., and Boter, J. 'Writing Cinema Histories with Digital Databases: The Case of Cinema Context', *TMG Journal for Media History*, Vol. 21, No. 2, 1 November 2018: 106-126.

Norén, F. '*Framtiden tillhör informatörerna*': *Samhällsinformationens formering i Sverige 1965–1975. (The Future Belongs to the Information Officers': The Formation of Governmental Information in Sweden 1965–1975)*. Umeå: Umeå universitet, 2019.

Norén, F. and Stjernholm, E. *Efterkrigstidens samhällskontakter (Post - War Society Contacts)*. Lund: Mediehistoriskt arkiv, 2019.

Norén, F., Stjernholm, E., and Thomson, C.C. *Nordic media histories of propaganda and persuasion*. Cham: Palgrave Macmillan, 2022.

Ordelman, R., Melgar, L., Van Gorp, J., and Noordegraaf, J. 'Media Suite: Unlocking Audiovisual Archives for Mixed Media Scholarly Research' in *Selected papers from the CLARIN Annual Conference 2018, Pisa, 8-10 October 2018*. Linköping: Linköping University Electronic Press, May 2019, No. 159: 133-143.

Pakarinen, E. 'PM om samhällsinformativa program mot ersättning i SR' (Memoranda on Commissioned Public Information Programmes on SR), Sveriges Radio's Document Archive, CK PS/PRS Handlingar rörande samhällsinformation (Documents concerning social information) F2:1, 28 September 1970.

Palmblad, E. and Eriksson, B.E. (eds). *Kropp och politik: Hälsouppllysning som samhällspegel (Body and Politics: Health Information as a Mirror of Society)*. Stockholm: Carlsson, 2014.

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

Seifarth, S. *Råd i radion: Modernisering, allmänhet och expertis 1939-1968 (Advice on the Radio: Modernization, the Public and the Experts 1939-1968)*. Stockholm: Carlsson, 2007.

Skogerbø, E., Ihlen, Ø., Nørgaard Kristensen, N., and Nord, L (eds). *Power, communication, and politics in the Nordic countries*. Gothenburg: Nordicom, 2021. SOU 1965:20 *Radions och televisionens framtid i Sverige (The Future of Radio and Television in Sweden)*.

Snickars, P. *Datalabb på KB: En förstudie (Data Lab at KB: A Feasibility Study)*. Stockholm: Kungliga Biblioteket, 2018.

SOU 1969:48 *Vidgad samhällsinformation (Expanded Public Information)*.

Syvertsen, T., Enli, G., Mjøs, O.J., and Moe, H (eds). *The media welfare state: Nordic media in the digital era*. Michigan: University of Michigan Press, 2014.

Thaler, R.H and Sunstein, C.R. *Nudge: Improving decisions about health, wealth, and happiness*. New Haven: Yale University Press, 2008.

Thomson, C.C. *Short films from a small nation: Danish informational cinema 1935-1965*. Edinburgh: Edinburgh University Press, 2018.

Tsivian, Y. 'Cinematics, part of the humanities' cyberinfrastructure' in *Digital tools in media studies: Analysis and research*, edited by M. Ross, M. Grauer, and B. Freisleben. Transcript Verlag: Bielefeld, 2009: 93-101.

Van Noord, N., Olesen, C., Ordelman, R. and Noordegraaf, J. 'Automatic Annotations and Enrichments for Audiovisual Archives' in *Proceedings of the 13th International Conference on Agents and Artificial Intelligence - Volume 1*, ARTIDIGH, 2021: 633-640.

Vonderau, P. and Hediger, V (eds). *Films that work: Industrial film and the productivity of media*. Amsterdam: Amsterdam University Press, 2009.

Wildy, T. 'From the MoI to the CoI—Publicity and Propaganda in Britain, 1945–1951: The National Health and Insurance Campaigns', *Historical Journal of Film, Radio and Television*, Vol. 6, No. 1, 1986: 3-16.

Notes

- [1] All Swedish-language quotes and word lists in this article have been translated from Swedish to English by the author.
- [2] Jönsson & Snickars 2007; Syvertsen et al 2018; Norén 2019; Skogerbø et al 2021.
- [3] Hirdman 1990; Thomson 2018; Norén et al 2019; Norén et al 2022.
- [4] Fickers & Williams & Snickars 2018.
- [5] Bordwell 1985; Bordwell 1997.
- [6] Tsivian 2009; Fournout et al 2014; Butler 2014; Manovich 2020; Arnold & Tilton 2020.
- [7] Flueckiger 2017; Kanatova et al 2017; Heftberger 2018.
- [8] Noordegraaf et al 2018.
- [9] Snickars 2018; Van Noord et al 2021.
- [10] Darnton 2000.
- [11] SOU 1954:32.
- [12] Elsässer 2012
- [13] Palmblad & Eriksson 2014.
- [14] Seifarth 2007.
- [15] SOU 1965:20, pp. 421-438.

DISTANT READING TELEVISED PUBLIC INFORMATION: THE COMMUNICATION OF SWEDISH GOVERNMENT AGENCIES, 1978-2020

- [16] Djerf-Pierre & Weibull, p. 211.
- [17] Cf. Vonderau & Hediger 2009; Acland & Wasson 2011.
- [18] Ellis 2011, p. 95.
- [19] Cf. Griffith 1983; Wildy 1986.
- [20] Harvard & Stadius 2013.
- [21] SOU 1969:48.
- [22] Djerf-Pierre & Ekström 2013.
- [23] Froehlich 2015.
- [24] Moretti 2000, p. 57
- [25] Hoyt et al 2018, p. 421.
- [26] Bednarek 2010; Bednarek 2018; Csomay & Young 2021.
- [27] Acland & Hoyt 2016, p. 12.
- [28] Lindgren 2017.
- [29] Heftberger 2018b.
- [30] Snickars 2018, p. 19
- [31] Ordelman et al 2019.
- [32] <https://huggingface.co/KBLab/wav2vec2-large-voxrex-swedish>, accessed 22 April 2022. For more on KBlab's development of a model for automatic speech recognition, see Malmsten & Haffenden & Börjeson 2022.
- [33] Elhamifar & Naing 2019, p. 6342
- [34] Jarlbrink & Snickars 2017.
- [35] Dahlgren 2021.
- [36] Anthony 2020.
- [37] Ascari 2014, p. 3.
- [38] Arnold & Tilton & Berke 2019.
- [39] Byszuk 2020, p. 5.
- [40] Salmi 2021, p. 65.
- [41] SOU 1969: 48, p. 9.
- [42] SOU 1969: 48, p. 52.
- [43] Pakarinen 1970, p. 2.
- [44] Eg. *Aftonbladet*, 5 June 1987.
- [45] Thaler & Sunstein 2008.
- [46] Acketoft 2005.
- [47] Möller 2006.
- [48] SOU 2007: 107, p. 35.
- [49] Thaler & Sunstein 2008, p. 60.
- [50] Berggren & Trägårdh 2014.
- [51] Connell 2005.
- [52] Korpus 2008, pp. 34-35.
- [53] Mårtensson 1988, p. 119.
- [54] SOU 1985: 3.
- [55] Korpus 2008, p. 38.
- [56] Hadenius 1998, pp. 290-313; Djerf-Pierre & Weibull 2013, p. 320.
- [57] Baldwin 2005.
- [58] Nicholson 2012.