

**Ideology and language:
Examining an archival method for cross-cultural research**

Erlend Sand Bruer



**Master's Thesis in Psychology
Department of Psychology, University of Oslo
June 2010**

Acknowledgments

Through working on this thesis as a member of the Oslo Ideology Project, I have had the pleasure of cooperating closely with many dedicated and generous psychologists.

First of all, I would like to thank my co-supervisors, the dynamic duo that the Ideology Project centers around: Hilde Eileen Nafstad and Rolv Mikkel Blakar. Thank you for making me feel like a valuable member of the Oslo Ideology Project. Thank you for valuable and professional feedback throughout the writing process. You have an uncanny knack for knowing exactly what to say in order to motivate your students and push them forward.

I have gotten invaluable feedback on methodology from my main supervisor, Dag Erik Eilertsen. It has been a pleasure to discuss method problems with someone who is not afraid to be critical and challenge his students. His command of research methods is unparalleled and inspiring.

Special thanks are deserved by Petra Filkukova, who has done an amazing amount of work in order to create the Czech dataset. In addition, she has spent hours discussing Czech ideology with me by mail or face-to-face. Thanks are also deserved by Albert Botchway, Salman Turken, Caroline Syverstad and Monira Hosainey, who have been kind enough to allow me to use the datasets they have created as basis for this thesis.

Kim Rand-Hendriksen, who laid the practical foundations of the archive method as it is presently used, has been great help in clarifying the methods used by the project.

Erik Carlquist, who has been associated with the Oslo Ideology Project almost from the beginning, has taken time out from a busy schedule to root out my worst abuses of the English language, as well as given valuable feedback about parts I needed clarify for the poor reader wishing to understand the esoteric world of archive methods.

My good friend Torgeir Hoffmann has suffered long discussions on statistics and has offered input and criticism in generous dosage. My dear sister Andrea Miranda Sand Bruer has been kind enough to help with proofreading.

Finally, I would like to thank all the members of the Ideology Project for creating an interesting, fresh and exciting academic environment. I have had a wonderful time working on this thesis, and I hope to be able to assist the project further in the future.

Abstract

The purpose of this thesis is to examine the sensitivity of the archival method of the Oslo Ideology Project for cross-cultural studies of ideologies and ideological developments through language. Moreover, a theoretical framework for this method is outlined. The thesis is part of the Oslo Ideology Project, and the archival method examined was created by and is continually developed within this project.

The substantial motivation for cross-cultural usage of this method is the ongoing effects of globalization on local ideologies. The thesis discusses and analyzes globalization's influence on the balance between individualistic and communal values in societies around the world. A theoretical framework for the method, grounded in discursive psychology and philosophy of language, is outlined. The method is presented in detail and supplemented with statistical tools for examining cross-cultural differences in ideological developments. Theoretical expectations about what the method should capture of ideological variation in three societies; the Czech Republic, Norway and the US, is empirically examined. In conclusion, this archive method is able to capture and describe general variations in ideology as well as providing a means for testing concrete hypotheses about specific ideological developments.

Table of contents

Introduction.....	1
Theory.....	2
Pan-cultural values: Individualism and communality.....	2
The influence of globalization on fundamental values	2
The concept of ideology	3
Ideology and discourse.....	5
Language as a system of usage	6
Words are not neutral	7
Methodology	10
Selection and translation of search words	12
Choosing of archives.....	14
The available datasets	15
Descriptions of the ideologies in Norway, the US and the Czech Republic	18
Adjustments of word frequencies	20
Levels of usage	24
Principal Component Analysis	25
Analyses and results	26
Analysis of the Norwegian dataset.....	28
Analysis of the US dataset	32
Analysis of the Czech data	36
Comparisons between datasets	39
Discussion: Limitations, future directions and concluding remarks.....	44
References.....	48
Web addresses:	53

Introduction

All societies must somehow negotiate the balance between the individual and society (Brewer, 2004; Vetlesen, 2004). This balance is continually redefined, changing the ideological systems that draw the boundaries between the rights and freedom of the individual and the individual's considerations about and duties to society. The process of globalization is spreading certain ideologies all over the world (Steger, 2005a). Over the past decades, neoliberalism (Gledhill, 2004; Harvey, 2005; Leitner, Peck & Shepard, 2007), consumerism (Bauman, 2000; Kasser & Soule, 2003) and radical individualism (Bourdieu, 1998) have been central dimensions of the ideologies spread by globalization, strongly influencing how local ideologies or value systems define their balance between individual and society.

This thesis is concerned with how one can capture and describe the development of ideological discourses of individualism versus communality in different societies, using language archives as empirical data. The focus will be on methodology. The aim is to present improvements to the archive method of using language corpuses for describing and comparing ideological changes and developmental trends; in particular, the aim is to improve and elaborate the theoretical-methodological basis for describing and comparing ideological discourses across different societies. For some years the archive methodology has been developed and refined within the Oslo Ideology Project, organized by Hilde Eileen Nafstad and Rolv Mikkjel Blakar. The project analyzes ideological changes reflected in the language of public media discourse (see e.g., Nafstad, 2002; Nafstad & Blakar, 2002/2006, 2009a; Nafstad, Blakar, Botchway & Rand-Hendriksen, 2009b; Nafstad, Blakar, Carlquist, Phelps & Rand-Hendriksen, 2007a, 2009c; Nafstad, Blakar & Rand-Hendriksen, 2009d; Nafstad, Carlquist & Blakar, 2007b; Rand-Hendriksen, 2008). As a member of this research group I have gotten access to the various datasets collected by the Oslo Ideology Project. This thesis presents new ways of analyzing as well as comparing development trends in archives of language from three different societies: Norway, US and the Czech Republic. The Czech and US materials were collected by Petra Filkukova and Caroline Syverstad, respectively.

Theory

Pan-cultural values: Individualism and communality

The value dimensions of individualism and communality are two of very few pan-cultural value dimensions with which all cultures must concern themselves (Markus & Kitayama, 1991; Triandis, 1990, 1995). Thus it can be argued that every society develops its own ideology connected to individualistic and communal values. Some societies are more individualistic; they value and emphasize the autonomous person with personal freedom (Nelson & Prillettensky, 2005; Chiu & Hong, 2006). Other societies are generally more oriented toward communal values, emphasizing family, community and country (Nelson & Prillettensky, 2005, Chiu & Hong, 2006). Concerning the universality of the individualism-communality dimension, Wierzbicka (2009) takes a perspective based on universalities in language use and argues that all societies must be able to answer the question of whether one should prioritize the community or oneself. Thus all languages should be able to reflect the value dimension of individualism versus communality. It should be pointed out that even though the predominant ideological orientation within a society will be toward either individualistic or communal values, there will always be a mixture of ideologies in each society, varying locally in degrees of individualism and communality (Triandis, 1995, Hermans & Kempen, 1998).

As Bauman (2000) concludes, however, all over the world communal values are disappearing, being replaced by individualistic values. As value systems change in this direction, the way people interact with each other in their everyday lives also changes (Chrysochoou, pp. 122-125). As Scheff (1990) argues, the social bonds that are fundamental parts of people's lives are becoming increasingly weakened. A central methodological issue is therefore how to describe and compare changes in individualism and communal values in different societies.

The influence of globalization on fundamental values

Globalization can be defined as “the closer integration of the countries and peoples of the world which has been brought about by the enormous reduction of costs of transportation and communication, and the breaking down of artificial barriers to the flows of goods, services, capital, knowledge, and (to a lesser extent) people across borders.” (Stiglitz, 2002, p. 9). Globalization has been seen as being driven by ideology in its own right, the ideology of ‘globalism’ (Steger, 2005a).

There are many different ways to approach the globalization processes in terms of social science research (Sklair, 1999). Sklair proposes four general sources of research questions in studies of globalization: One may examine the different roles different actors play in the world-system, one may examine what the effects of the homogenization of culture associated with globalization are, one may analyze what the possible consequences of a truly global society are, or one may study the forces of globalization as located in the mechanisms of global capitalism (Sklair, 1999, p. 149).

The present thesis is addressing Sklair's (1999) issue of the homogenization of culture associated with globalization. More precisely, the aim is to develop a methodology for describing and comparing ideological changes in an era of strong globalization. Concretely, we seek to investigate and map the ideologies of individualism and communality, how they change and whether they develop in similar or different ways in different societies. Psychologically, this is an important issue to study as the ideological situation in a society will have profound implications for how the individual views his or her society, his or her place in it and duties and rights as citizens (Sampson, 1989; Arnett, 2002; Finkel & Moghaddam, 2002).

This thesis will make use of changes in the language of public media discourse as an empirical barometer of ideological change. In this endeavor, we have to take into account that even though the changes seem to be going in the same general direction all over the world, towards a stronger individualism, different societies may change in different ways. All societies have their own economic, political, social and cultural history which makes them implement the globalizing ideologies in their own unique manners (Fairclough, 2006; Nafstad et al., 2009b).

The concept of ideology

The concept of ideology originates in 18th-century France, and has since then taken on many different meanings (Thompson, 1990, Eagleton, 1991). In his introduction to ideology, Eagleton (1991) lists sixteen different definitions in his opening chapter. A distinction is often drawn between two types of definitions of ideology (Thompson, 1990). Some see ideology as something false, something other than reality – classically the Marxist notion of the “false

consciousness” that masks the “real” material basis for the ideas that define society. Others have a more neutral conception of ideology, regarding all systems of thought as ideological. With the first type of definition, there is the problem of “piercing the veil”; what will be the final answer regarding which ideas are “real” and which are “just ideology”? With the latter type of definition, the immediate difficulty of separating ideology from culture arises; if all systems of thought are ideological, does it make sense to speak of ideology at all? Eagleton (1991) argues that conceiving of ideology as something “neutral” takes focus away from the dynamics of power which underlie various thought-systems; many things tacitly taken for granted in different ideologies are “flagrantly false” (Eagleton, 1991, p. 222). Thompson (1990) solves this dilemma by moving beyond the question of what material interests underpin ideology, while still attempting to keep the critical aspect of the ideology concept. He defines ideology as “meaning in the service of power”. This definition is useful for the present analysis, as we are studying changes in ideological meaning structures through studying changes in language use.

Turning to social psychological concepts of ideology, Nafstad et al. (2007a) locate ideology at the macro level of society; the societal level, viewing it as something that “encompasses and permeates the other levels” (p. 314). Ideologies are seen as “fundamental vehicles of power as they serve to control the positive or negative ways of how individuals and groups adapt to and master their environments.” (p. 314). This approach to ideology is closely related to the definition of macrosystem found in Bronfenbrenner (1977, 1979). Here macrosystem is defined thus: “The *macrosystem* refers to consistencies in the form and content of lower-order systems (micro- meso-, and exo-) that exist, or could exist, at the level of the subculture or the culture as a whole, along with any belief systems or ideology underlying such consistencies.” (Bronfenbrenner, 1979, p. 26). Within Bronfenbrenner’s ecological model, ideology and the macrostructure of society are strongly linked to the microstructures of everyday life.

In the present thesis I conceive of ideology as meaning structures at the macro level that shape and form how individuals make sense of their everyday lives. Ideology analysis then deals with the meaning structures that shape evaluations of everyday life.

Ideology and discourse

As this thesis uses language archives for describing and comparing ideological changes and developmental trends, I will draw in part on discursive psychology. Discursive psychology is a method that favors close and rigorous analysis of language over experiments, surveys or interview work (Potter, Edwards & Wetherell, 1993; Potter, 1997; Potter & Edwards, 2001). Potter and Wetherell (1987) present the idea that the proper way to view language in use is through the concept of discourse. Use of language is not determined by a subject's limitless creative construction of words and sentences, but language use is firmly grounded in systems of utterances (Foucault, 1972/2007). Moreover, discourses have certain relations to each other, and the issue of discursive psychology is not to try to identify the "true" utterances reflecting what "really happened" (Curt, 1994). The aim is to identify how certain states of affairs are constructed to represent something rather than something else (Billig, 1987; Miller, 1997).

Ideological values in a society are inseparably tied to the discourses in that society. The Oslo Ideology Project has related ideology to discourse in the following way: "Correspondingly, ideology can be understood as a system of intercommunication, a common language or public discourse within a society, which provides shared categories of thought, values and explanations about the social and material world" (Nafstad et al., 2007b, p. 330). In order to understand and describe ideology in a society, one has to understand the patterns of discourse.

The concept of discourse, however, is complicated. It has been used for different purposes within different traditions (Curt, 1994). Definitions of discourse vary, ranging from "a set of statements which construct an object" to "a set of statements the production conditions of which can be defined" (Íñiguez, 1997, p. 148). I will use the concept of discourse as simply referring to 'any statement or set of statements produced by a language user'. I would like to locate the concept of discourse in the local, practical utterances made by people as they go about in their daily lives.

Íñiguez (1997), moreover, discusses the relationship between discourse and social structure. She holds that "discourse is language as social practice determined by social properties of social systems" (p. 153). Language, discourse and social systems are thus mutually constitutive – there would be no language without discourse (without sets of statements), and

there would be no social systems without the means to communicate them. Ideology could then be seen not as ontologically separate from discourse (not as some external “interest of privileged classes” forcing itself upon the discourse), but as an integrated part of discourse.

How is ideology reflected in discourse? It seems fair to assume that most people most of the time do not talk explicitly about the macrosystems in which they live, but about everyday, “normal” things. Why would these conversations still reflect the ideological structures they take place within? The answer, I suggest, lies in the nature of language. The close relationship between discourse and ideology described above presupposes therefore a certain understanding of language.

Language as a system of usage

Many researchers conceive of language as more or less neutrally capturing and depicting the external world. These researchers conceive of spoken and written language as a cognitive vehicle for organizing thoughts (e.g. Pinker, 2002; Murphy, 1991). Accordingly, words refer to objects or concepts the individual needs in order to describe the world “as it is”. Humans are given concepts through perception, and language simply connects these concepts to words (Pinker, 2002, pp. 208-209). A further example is formal semantics, where “meaning” is defined as the relation between words and the world, and it is thought to be possible to identify definitively all objects a given word refers to (e.g. all objects named “chair”) (Murphy, 1991).

This model of language has been challenged by various philosophers and researchers (Bakhtin, 1968; Gergen, 1985, Shotter, 1992). Can language really be understood as such a simple system of referential words? Ludwig Wittgenstein is one of the key philosophers in criticizing conceptions of language as a more or less neutral way of capturing the external world (Sluga, 1999; Candish & Wrisley, 2008). Whereas Wittgenstein’s earlier work, *Tractatus Logico-Philosophicus* (Wittgenstein, 1921/2004), analyzes the nature of language and knowledge, especially relating to what logic or science can accomplish, it is his later work, *Philosophical Investigations* (Wittgenstein, 1953/2001) that has been particularly influential within psychology and the social sciences.

The central theme in Wittgenstein’s philosophy of language is the notion of word meaning as word use. Wittgenstein (1953/2001) criticizes the traditional account of language, starting

with Augustine's notion of language: Words signify objects, and children learn the names of objects by observing adults' naming of the objects. Wittgenstein (1953/2001) identifies several problems with this account. How is one to know which object is being referred to just by observing another human, index finger stretched out, making some seemingly random noise? What is the human trying to communicate? What is it in the outstretched index finger that has or gives meaning? Wittgenstein's (1953/2001) position is that one does not learn the names of objects when learning a language. One acquires or learns the rules of a language-game in which each word has its use. Wittgenstein thus draws attention to the grammar of language use. We learn a set of grammatical rules that govern when it is appropriate to use a word.

Wittgenstein (1953/2001) acknowledged and formulated the strong interrelations between meaning and word practice or use. Some social psychologists have also turned from cognitive analyses to linguistic analyses (Potter & Wetherell, 1987). They focus on how language is used, rather than what sorts of cognitive processes underlie our usage of words.

Words are not neutral

So far I have argued that language acquires meaning through use and that the patterns of language use conceptualized as discourses have a strong relation to ideology. I will now proceed to discussing the role of the single word. How can analysis of usage of single words provide knowledge about the ideology in a given society? What is the role of the single word? Blakar (1973/1996) argues that words are not neutral. Single words can never fully describe any object, rather the words draw attention to certain aspects of an object, obscuring or omitting other aspects (Rommetveit, 1968). We choose, for example, whether to call a given person a "terrorist" or a "freedom fighter", or to call someone "eccentric" or a "mental patient". None of these labels give a total or "accurate" description of the person in focus.

The methodology of the Oslo Ideology Project is based on analyses of the single word; in the present thesis by means of measuring changes in the usage of single key words. These are words that are especially chosen because they are expected to contain some specific evaluative description of the world – in the present thesis, words that describe the world in communal or individualistic terms. How can these describe ideology? Two answers can now be given: First, when describing the world, single words are chosen from a multitude of

possible words. One can describe an issue or a situation using many different, almost synonymous words, but the choice of words is never neutral. If it is demonstrated that in a given society there is, for example, a marked change in the frequency of words expressing individualistic rather than communal values, this indicates a change in how the social world is described and thereby defined. Changes regarding which words are used can, then, reflect changes in ideology. Second, to borrow a metaphor from Scheff (1990, p. 10); each word is a hologram – each word contains in it the entire language system. Each word acquires its meaning through the system of usage of which it is part. The word can thus never be completely isolated from its originating system of usage. Especially for words that strongly reflect certain dimensions of ideology, the system of usage must be changed in order for it to be possible that the words change in usage (Blakar, 1973/2006). Thus, it should be possible to use an analysis of changes in usage frequencies of chosen key words to analyze changes in words' systems of usage – systems of usage that are assumed to reflect ideologies.

Assuming that words carry with them ideological valence, what reasons do we have for expecting that single words can be compared across cultures? The present thesis will compare the usage of words we expect will reflect individualistic and communal ideologies in three different societies; US, Norway and the Czech Republic. To do this, we have to take as our point of departure that such fundamental ideological dimensions as individualism versus communal values are in some way reflected and mirrored in all these different languages, and in the usage of single words. As Wierzbicka (2009) argues, a certain minimum of basic words are needed in all cultures and all languages. A functioning human language requires a certain number of such basic concepts (Wierzbicka, 2009, p. 262). As previously argued, all cultures must define a balance between individualistic and communal values. Included in Wierzbicka's basic concepts are some words that reflect this central divide, such as "I", "people", "part" or "other", words that serve to mark boundaries between the self and the group (Harkins & Wierzbicka, 2001). The exact function of these words may vary across societies, and there is also evidence that some words central to the individualism-communality divide, such as the word "I", are not universal (Kondo, 1987). Therefore, we must be careful when choosing words to represent individualistic and communal values across languages. The basic concepts found by Wierzbicka support the assumption that the individualistic and communal values are central in all cultures and that they may be analyzed through careful study of single words.

As previously mentioned, Nafstad et al. (2009a, 2009b, 2007b) argue that all cultures have to, in one way or the other, continually negotiate the balance between the interests of the individual and society. Moreover, they argue that words that represent and tap central ongoing processes and negotiations about social equality, social justice, social responsibility, civic engagement, consumerism, etc. are central in capturing the balance between individualism and communal values in different societies (Nafstad et al., 2009b).

To sum up: It is argued that the structure and use of language has some relation to the ideological situation of a given society. Moreover, language is a system of practical usage that people actively use to understand and define the world. Ideology acts as a set of presuppositions that guide language-users in forming and evaluating the world. Which words a language-user utilizes in order to define the world is not arbitrary: Words are not neutral, but express some evaluation of the world. We expect that the single word should carry with it the predominant ideologies within which it is embedded.

The present thesis, therefore, uses a word count strategy approach for studying the values expressed in natural language use. Word count strategies have previously been used to study topics closely related to ideology and values, such as emotional valence and situational influences on language (Pennebaker, Mehl & Niederhoffer, 2003). The strategy of counting word frequencies may miss out on some details of language use, but it gives a good overview of general trends. As media language is central in distributing ideological content (van Dijk, 1998; Thompson, 1990), this thesis uses national newspapers from three different societies as archives for studying natural language use. Ideological debates occurring in societies with free press should, therefore, be reflected in the newspapers of those societies.

To conclude: Developments in word frequencies are assumed to reflect ideological developments in society, and frequencies of word usage are thus assumed to vary along with changes in ideology. This covariance could behave in at least two ways:

- Words that reflect the dominant ideology might be the most used, as people normally view society through the lenses of the dominant ideology.
- Words that reflect ‘counter-ideologies’, that is, words that oppose the dominant ideologies, could increase as dominant ideologies may be challenged and discussed. Thus the frequency of words reflecting for example individualistic values might

increase when there is actually an increase of communal values, as people who oppose the predominant individualistic ideology use individualistic words in opposing the dominant ideology.

My hypothesis is that word frequencies of ideological words should increase as the ideological dimension reflected by that word grows more dominant, and word frequencies should decrease when the dimension they reflect grows weaker.

This prediction about increases and decreases has to some extent been supported by previous research in Norway (Nafstad et al., 2007a; Rand-Hendriksen, 2008). The Oslo Ideology Project has examined electronic newspaper archives from the Czech Republic, Ghana, Iran, Norway, Turkey and the US. In order to examine whether the archive method captures cross-cultural differences in ideology, we have chosen to use the electronic newspaper archives from three of these countries: the Czech Republic, Norway, and the US. We have chosen these countries mainly because they present the best of the available electronic archives (for details, see pp. 15-18). We know that the US is a heavily individualistic society (Cullen, 2003). We also know that Norway is a society with strong emphasis on egalitarian values (Nafstad et al., 2007a, 2009b). Moreover, we know that the Czech Republic has recently had a lot of public debate about values after the fall of the iron curtain (Filukova, personal communication, May 10, 2010). Thus, these societies are relevant to use in an examination of our methodology, as we know that they are different with regard to ongoing ideological discourses about individualistic and communal values. If valid, the archive method of the Oslo Ideology Project should be able to describe these ideological differences quantitatively.

Methodology

The remainder of this thesis will concentrate on examining and improving the archival method used by the Oslo Ideology Project. This is a method that should enable us to analyze ideological developments cross-culturally through changes in word frequencies.

The challenge is to find a method that is close enough to the language data to capture the actual ideological dimension, but still able to detect ideological patterns in large-scale language use. While it may be instructive and valid to use a small-scale method, for example discourse analysis, to analyze the ideological dimensions in a text, it is impossible to analyze

the ideologies of entire societies exclusively through small-scale methods. Moreover, as Potter and Wetherell (1987) point out, discourse analysis must be grounded in a good understanding of the society within which the text appears. How is one to obtain such understanding, not of one, but of several societies? Survey studies have been used to map the attitudes of larger populations in the social sciences and psychology, and survey research has undoubtedly been useful in ideology research (Rand-Hendriksen, 2008). However, survey studies are expensive and time-consuming. The Oslo Ideology Project proposes that an archival method which examines natural language use in mass media can be useful in the study of ideology in society on a larger scale.

In essence, the method consists in using theoretical and conceptual analysis to identify a number of words that carry ideological significance. Then different archives are searched for numbers of occurrences of these words at different times. The developments of usage frequencies of these words are then mapped out. It is expected that different developmental patterns will emerge for different kinds of ideological situations. More precisely, the various steps of the method are as follows:

- 1) First, we identify what areas of ideology we aim to investigate. In the present thesis, it is the ideology systems of individualistic and communal values which are investigated.
- 2) Based on pilot searches and theory, we then choose search words that mirror the chosen areas of ideology. In this thesis we have selected twelve search words to represent individualism and communality, ranging from very frequently used words like 'I' or 'we' to more specifically value-laden words like 'solidarity'.
- 3) We then identify electronic media archives from different countries where we search the chosen words. Within the Oslo Ideology project six electronic newspaper archives from six different countries are currently included. Three of them are analyzed in depth in this thesis.
- 4) Datasets are created by searching the archives for usage frequencies of the chosen words per year. In this thesis, this frequency will be the number of newspaper articles within which a given word occurs. The datasets consist of matrices where each word is represented by a listed frequency for each year.

5) We then calculate proportional levels of usage for the words to compare how often the different words are being used in the various societies. The developmental trends over time for the words are also graphically mapped out.

6) Different trends are expected to emerge from different datasets/societies. These trends will then be interpreted based on theory, knowledge of the different societies and development of one dataset relative to the other datasets.

To sum up: It is expected that different societies will show different levels as well as different developmental patterns with regard to communality and individualism. This should be revealed in different patterns for the search words in the different datasets/societies.

Selection and translation of search words

In the present thesis we are interested in the ideology dimension of individualism-communality. Therefore, we need to select search words that capture the world evaluatively either in an individualistic or communal direction. However, we should try to avoid words specific to particular cultures. Based on reviews of the psychological and social science literature (Brewer, 2004; Finkel & Moghaddam, 2002; Nafstad, 2004; Nafstad & Blakar, 2009a, Jørgensen & Nafstad, 2004) and the Oslo Ideology Project's own empirical studies over several years (Nafstad, 2002; Nafstad & Blakar, 2002/2006, 2009a; Nafstad et al., 2006, 2007a, 2007b, 2009b, 2009c, 2009d) a set of twelve search words capturing individualistic versus communal values has been identified within the Oslo Ideology Project. The search words of this set have been validated as useful markers of ideological development within the Norwegian dataset (Rand-Hendriksen, 2008). The set of identified search words is listed below. As the words were first identified in studies of ideological shifts in Norway, I will provide the original Norwegian forms first and then the chosen English translations:

rettighet*¹ (right²)

plikt* (duty)

ansvar* (responsibility)

felles* (common)

solidari* (solidarity)

¹ * Means that the word was searched truncated.

² Reanalysis of the US dataset has led to the discovery that the single form 'right' must have been used as search word, rather than the plural form 'rights', as was intended. This is unfortunate, as the word 'right' is a synonym also referring, among other things, to the opposite of left.

samhold* (cohesion)
brukere (users)
borgere (citizens)
jeg OR meg (I OR me)
vi OR oss (we OR us)
likhet* (equality)
rettferdighet* (justice)

I would like to point out that some of the words have been searched in different tenses from the listed words to avoid multiple meanings. For example, ‘user’ in Norwegian (‘bruker’) means the same as ‘to use’ – however, searching for ‘users’ (‘brukere’) eliminates this double meaning. I will briefly comment on the list of search words:

‘I or me’ versus ‘us or us’: A fundamental choice when defining the world is to identify the actor and language user in a given situation as either an individual person (I/me) or as an integrated part of a group or community (we/us).

‘Right’ versus ‘responsibility’ and ‘duty’: ‘Right’ captures what the individual is entitled to receive from the group, community or society, whereas ‘duty’ and ‘responsibility’ capture the individual’s obligations towards the group, community and society.

‘Common’ and ‘cohesion’: These words are referring explicitly to communality and communal values.

‘Equality’ and ‘Justice’: Both these search words refer to fundamental values by which any society can be evaluated. Higher degree of realization in a society of ‘equality’ and ‘justice’ respectively means stronger communal values in that society.

‘Solidarity’: This search word refers explicitly to communal values. Solidarity actions are by definition not individualistic.

‘Citizen’ and ‘User’: These two words capture opposite ends of the individual-society relation: A ‘citizen’ is someone who takes responsibility and cares for society whereas the ‘user’ is someone at the end of a chain of services who is entitled to certain goods from society and community.

A critical methodological problem is that the list of search words have to be translated to create as equivalent lists as possible in the other languages to be investigated. Of the societies investigated so far within the Oslo Ideology project translation problems have been

most apparent in the Iranian, Czech and Turkish datasets. English is the written language in Ghana.

All translations from Norwegian/English for analyses in Turkey, Iran and the Czech Republic have been conducted by people who are native speakers of the actual language to be analyzed and who are also trained in the methodology used by the Oslo Ideology Project. Some of the translations have also been validated by professional experts on the actual language. While a more extensive validation of the translations may be desirable, the translations should be adequate for my purposes. Having selected and translated the search words, the next step is to identify relevant archives for analysis in the various societies.

Choosing of archives

The archival method is an old method, used in psychology and the social sciences for purposes ranging from statistical analysis of connections between uniform color and aggression to qualitative analysis of letters (Hoyle, Harris & Judd, 2002). By ‘archive’ it is in this thesis meant selected electronic databases of newspapers. ‘Dataset’ means the sets of usage frequencies that have been obtained by searching the archives. My analyses focus on the datasets. However, a note on the archives as such is necessary to make clear the rationales for selection of these specific archives, as well as to identify some possible methodological limitations that may follow from using them. When deciding which archives to use in my analysis, a number of properties were important:

Electronically searchable: A technical prerequisite is that the newspaper archives to be used must have electronic search systems that allow identification of articles containing a specific search word on annual basis to allow us to map out developmental trends over time.

Representativeness: Since our aim is to use data from natural language use to gain information about the underlying ideologies in various settings or societies the archive represents, it is important to know to what extent the archive is representative. As Shaughnessy, Zechmeister & Zechmeister (2009, pp. 388-390) discuss, it can be very difficult to define a proper population of newspaper archives and then draw a representative sample from this archive. Especially considering the present goal of cross-cultural comparisons, the task of mapping out all newspapers in different countries would be enormous. When using the archival method, representativeness is hard to reach and it may even be hard to specify what a

representative sample of natural language use would actually mean (Bauer & Aarts, 2000). In the present study, representativeness is understood as how well the archive reflects the general debates in its home society.

Completeness: One should know whether the archive actually contains all articles published in the actual newspaper during the timespan investigated. It might be possible, for example, that some kinds of articles have been left out when digitalizing earlier editions or that some articles are not saved for whatever reason. This could lead to biases in the archives.

Timespan: Since the methodology we use primarily make sense when analyzing shifts and changes over time, the time span of the archive is essential. Some of the archives used by the Oslo Ideology Project only go back a few years, while others cover a time span of 25 years or more. To capture ideology developments over time, one usually wants to have archives with as long timespans as possible. One major advantage newspaper archives have over other possible archives, such as internet blogs, is that newspaper archives potentially go further back in time.

The available datasets

Six datasets have so far been created by the Oslo Ideology Project, all from different archives in different societies.

Norwegian dataset: (Nafstad et al, 2007a, 2009a). The Norwegian dataset is based on searches in the electronic archives “Retriever”, an online archive where all articles from many of Norway’s newspapers are made accessible. The archives run back to 1984. The data used in this thesis are from the newspaper “Aftenposten” which is at present (May 2010) Norway’s largest newspaper. Aftenposten contains a very substantial number of articles, about 70 000 on average each year. The Norwegian dataset is especially easy to search, as one of the members of the Oslo Ideology Project, Kim Rand-Hendriksen, has designed a web interface that automatically extracts word frequencies per year, produces a graphic representation of frequencies of usage, and also calculates frequencies of searched words adjusted relative to the total number of articles each year. The Norwegian dataset is also by far the most thoroughly controlled and validated. Moreover, word development in the newspaper Aftenposten has been shown to correlate closely with word development in other Norwegian newspapers: for all the search words in the present dataset the development correlates significantly (alpha of at least .05) with the development in five other Norwegian newspapers

taken together. This means that the trends observed in the Norwegian dataset are very unlikely to be caused by idiosyncrasies in *Aftenposten*. Also, the developmental trends of our twelve search words have been shown to depart markedly from the developmental pattern of the 10 000 most frequently used words in Norwegian (Rand-Hendriksen, 2008), an observation that suggests that these trends have not been caused by random fluctuations in the Norwegian language.

US dataset: (Syverstad, 2009) The US dataset is based on searches in the electronic archives of the New York Times, with data available from 1984. The New York Times has an enormous archive, consisting of, on average, 92 000 articles annually. To ensure representativity only search words for which the development in the New York Times correlate with the development in Washington Post for the years 1987-2008 with significance at least at an alpha of .05 will be included in the analyses. This means that word searches have undergone some representativeness tests showing that the observed trends of development are not solely due to particular properties of the New York Times.

Ghanian dataset: (Nafstad et al., 2009b) A searchable electronic news archive containing selected articles from various Ghanaian newspapers is available in the form of the news site “Ghanaweb”. This news archive is an edited archive, and does not cover complete newspaper output. Still, this archive includes articles from the main Ghanaian newspapers and covers a broad range of issues: from politics and economics to culture, religion and sport. The Ghanian archive goes back to 1995. However, for 1995 and in particular for 1996, very few articles are included (411 and 90, respectively). Consequently, it has been decided within the Oslo Ideology Project to use 1997 as the starting year. It is unfortunate that there is no alternative electronic news archive available in Ghana to check the representativity of Ghanaweb. Nor do we know the editorial policy by which articles are selected from the various newspapers to be included in Ghanaweb.

*Czech dataset*³: The data in the Czech dataset was obtained by searching the electronic archives of the Czech Republic’s second largest newspaper, “Mladá fronta DNES”.

³ The Czech material will be presented internationally as an integrated part of the comparative studies of the Oslo Ideology Project at the 5th European Conference of Positive Psychology in Copenhagen, June 2010 (Nafstad et al., in press).

Historically, this paper used to be the mouthpiece of the Czech Communist Party's Socialist Youth, but its journalists were among the first to challenge censorship and cover the Czech Velvet Revolution. After the revolution, the word "dnes" (Czech for "today") was added to the title of the newspaper, signifying end of state ownership and censorship of the paper. The second largest newspaper was chosen instead of the largest Czech newspaper, "Blesk", because Blesk was regarded as too tabloid compared to the other available archives. Comparison between archives could have been difficult if the selected newspapers varied greatly in journalistic style. The words used in the search were translated into Czech by Petra Filkukova, a Czech Ph. D-student associated with the Oslo Ideology Project. Great care was taken in searching for translated words; as the Czech language has seven grammatical cases, the searches were conducted by adding hits from all cases of the words (using the Boolean OR function while searching). For the sake of simplicity, only English versions of the words are presented in this thesis. The archives were accessed through a third party, Newton Media, as the public search function of the Mladá fronta DNES web page was inadequate; it only went back to 1999 and it was not possible to limit the search to only the newspaper – searching included two magazines sometimes published as a supplement to the paper. The Mladá fronta DNES archive is the largest one included in this thesis, consisting of on average about 191 000 articles annually. Data were available from 1996, but due to great fluctuations in number of articles in the first years (1996-8) only word developments from 1999 on are included in the final dataset. We did not have resources to check the representativity of this archive against other archives, as was done in the Norwegian and the US datasets. However, the sheer number of articles included in the archive, along with its position as second largest newspaper in the Czech Republic, should ensure a certain degree of representativity of the public discourse in the Czech Republic.

Iranian dataset: (Hosainey, in preparation). The Iranian dataset was obtained from the electronic archives of the newspaper "Iran", the official government newspaper published by the Islamic Republic News Agency. The archives were searched through a publicly available news search function on the web page of the newspaper. The archives are of limited size, averaging only about 5500 articles annually. The timespan is also the shortest of the available datasets of the Oslo Ideology project, with only the years 2002 through 2008 available. Translation was done by a master's student currently writing her thesis, Monira Hosainey. Translation was checked with a professor of Iranian studies at the University of Oslo. There

may be some problems with translation or the choice of words, as several of the words have zero frequency in some years.

Turkish dataset: (Türken et al., in preparation.) The Turkish dataset was obtained by using the public search function of the web version of the newspaper “Hürriyet”. The archive contains an average of about 88 000 articles each year, and data was obtained from 1998-2008.

Translation was done by Salman Turken, a Ph. D student associated with the Oslo Ideology Project. Translation was checked with a professor of Turkish studies at the University of Oslo. However, the collection of the Turkish data was not entirely completed when work on this thesis started.

As previously mentioned, I have chosen the datasets from Norway, the US and the Czech Republic for an examination of the validity of the archival method for cross-cultural studies. There are three reasons for choosing these specific datasets: First and foremost, as has been shown, the quality of these datasets is by far the best in these three societies. Second, as previously discussed, we assume that these three countries represent markedly different ideological situations. If the archival method does not succeed in capturing the essence of these differences, the method will not prove useful and sensitive in cross-cultural research. Third, our knowledge about the ideology situations in Iran, Turkey and Ghana is more limited than our knowledge about the situation in Norway, US and the European Czech Republic.

Descriptions of the ideologies in Norway, the US and the Czech Republic

The three chosen countries have had very different histories and ideological developments. In order to establish whether the archival method is adequate for capturing the different ideological patterns, we need to know something about what kinds of ideologies have dominated and currently dominate these three societies. As the present thesis only aims at testing general expectations about the analyses of the datasets, the descriptions of the different societies is brief. More detailed examinations of the datasets would require greater knowledge of the examined societies.

Norway: Norway is a country where communal values such as egalitarianism and social equality traditionally have been highly valued (Nafstad et al, 2009b). It is one of the

Scandinavian welfare states, and has experienced very strong economic growth the past decades, mainly due to rich natural resources in the form of oil. The political situation in Norway has been stable for the period we have available data (1984-2008).

During the past decades, the predominant ideology has been changing towards a more neo-liberalist orientation (Carlquist, Nafstad & Blakar, 2007; Nafstad et al., 2007a). Most recently (from 2006 to present), however, the Oslo Ideology Project has reported somewhat more counter-ideological tendencies which may indicate that neoliberalist influence has reached a top or turning point (Nafstad et al., 2009b, 2007a).

US: The United States of America is a country where personal autonomy and individual rights and freedom are national ideals (Cullen, 2003). Indeed, the nation was founded on these principles, and even the Declaration of Independence states that upholding individual rights is the prime function of government (Declaration of Independence, 1776). The early European settlers who fled to America were seeking to escape societal bonds, especially bonds regulating religion, economy and social position (Jenkins, 1997). These ideas still characterize the USA today, as it is a country where free-market liberalism and individualism are strong ideologies (Cullen, 2003).

Politically, the nation has been stable. There has, however, been a change in ideology following the 9/11 attacks in 2001, and the resulting ‘war on terror’ (Steger, 2005b). These changes have pushed dominant US ideology towards even stronger neoliberalism (Steger, 2005b).

*Czech Republic*⁴: Czechoslovakia was created in 1918, following the collapse of the Austro-Hungarian Empire. From 1939-1945, the country was occupied by Nazi Germany. In 1948, the Communist Party took control over Czechoslovakia and the country became part of the Eastern Bloc. In 1989, Czechoslovakia changed from a communist regime to liberal democracy, during the so-called “Velvet Revolution”. The Czech Republic was created in 1993 when Czechoslovakia peacefully split into two independent countries; the Czech Republic and Slovakia (Agnew, 2008). According to Švejnar & Hvižd’ala (2008), Czech

⁴ The section on Czech history and ideology is written with much help from Petra Filkukova, a Czech Ph.D student and member of the Oslo Ideology Project. The cited sources Agnew (2008) and Švejnar & Hvižd’ala (2008) are written in Czech.

ideology has been deeply marked by 41 years of communist rule, leading to a widespread attitude where individuals are unwilling to take responsibility for decisions, political or other. Responsibility was, during communist rule, seen as a collective matter.

The first Parliament election in the new Czech Republic resulted in a government consisting of liberal conservatives. Liberal conservatives ruled until 1998, when a social democratic government took power. In 2006, however, a liberal conservative coalition once again took power. There is a growing public dissatisfaction with the inability of Czech politicians to deal with corruption, and several commentators have called the present societal system “wild capitalism” (Filukova, personal communication, May 10, 2010).

The state ideology in the Czech Republic has undergone profound changes the past years, from communist regime to liberal conservatism. The country is struggling to deal with these transitions, and there is growing public disillusionment with the present government.

Adjustments of word frequencies

As previously discussed, systematic quantitative mapping of the usage of particular words over time is conducted to identify and investigate ideologies and ideological changes in the societies the various archives represent. The datasets analyzed in the present thesis were created by searching newspaper archives for number of articles within a given year containing a particular search word. Criteria for selection of search words have been presented above (p.12). The electronic searches returned the number of articles within which each search word was used. ‘Word frequency’ in the present thesis means the number of articles containing a given word. This may be somewhat imprecise, as search words may be used many times in a single article. Considering that the electronic archives in the three chosen societies consist of hundreds of thousands of articles, however, it is assumed that number of articles containing a word gives a reasonable representation of how often that word is used.

The Oslo Ideology Project has chosen to use the calendar year as unit of analysis, registering how many articles contain a given word in a given year. Naturally, it is possible to use shorter time-units to conduct a more fine-grained investigation of the archives. Using calendar years gives few data points and means that one requires several years of data to be able to conduct a meaningful analysis. However, this timespan is appropriate given that the types of ideologies

(such as individualistic versus communal values) investigated by the Oslo Ideology Project are expected to represent fundamental parts of the meaning structures in societies. Thus, these ideologies are expected to change over years rather than months or weeks.

Using the searched word frequencies directly in our analyses, however, is problematic. The total number of articles in a newspaper varies from year to year. Changes in the frequency of a given word across different years may be caused partially by changes in number of articles. It is therefore difficult to conclude from the “raw” data frequencies directly whether changes in usage frequencies of words reflect ideological changes, or simply reflect increases or decreases in total numbers of articles published per year. Thus, the raw frequencies have to be adjusted to compensate for the yearly changes in number of articles in the archive. Solutions to this problem adopted by the Oslo Ideology Project have been presented by Rand-Hendriksen, (2008). To ensure completeness in my presentation of the methodology, an outline of these methodological solutions is given here.

In order to obtain an estimate of number of articles in an archive or in a specific year, the archive is searched for extremely common words or elements (Rand-Hendriksen, 2008). If the search engine allows it, one may search for blank spaces (“ ”), which all articles necessarily contain. Almost all articles will also include some very common words, such as “a”, “an” or “the” (or appropriate translations of such words). Given such estimates of the annual total number of articles, one can adjust the raw frequencies so that they will be comparable across different years. There are potentially different ways to do this. One obvious way could be to use a measure similar to a price index, and set a certain year as the base year against which all frequencies are compared. One could then use the ratio of word frequency to total number of articles in the archive as an index adjustment. However, it turns out that some of the words searched have a frequency of usage of zero some years. Should a word have zero frequency in the base year, an index would not be defined for that particular search word. To avoid this, raw frequencies of word usage have in the Oslo Ideology Project been adjusted using the average number of articles all years as “base year”. This adjusted frequency is used as the basis for all calculations in the present thesis.

$$A_i = \frac{S_i * \bar{N}}{N_i}$$

A_i = Adjusted frequency for word A in year i

S_i = Number of articles containing word A, year i ('raw frequency')

\bar{N} = Average number of articles in the newspaper archive across all years

N_i = Number of articles in newspaper archive year i

Adjusted frequency: The adjusted frequency is obtained by multiplying the raw frequency for each year with the average number of articles in the archive all years divided by the number of articles in the archive that year. If the number of articles does not fluctuate much, the average number of articles divided by number of articles for a specific year will be a number close to 1. The adjusted frequencies will then be close to the raw scores. If, on the other hand, the number of articles in the archive varies a lot from year to year, the adjusted frequencies will differ more from the raw scores.

The adjusted frequency allows us to compensate for changes in the number of articles and thus to compare frequencies across different years. The adjusted frequency is used as basis for two further transformations of the raw frequency; relative frequency and proportional frequency. It is important to point out that these transformations only are linear transformations of the adjusted frequencies – they do not change the developmental profiles of the words. Only the scale of presentation is changed. The relative and proportional frequencies are only used to ease comparison when comparing different words and across different datasets, respectively.

Relative frequency: When one wishes to compare the development of different words in the same or in different archives, the adjusted frequency may be difficult to interpret directly. Some words will show dramatic developments, while others may change slowly in frequency. Then it may be useful to calculate relative frequencies for different words in order to better compare them to each other. There are different ways to measure relative changes within a set of numbers. One obvious way is to standardize the numbers by subtracting the mean and dividing by the standard deviation. This yields a standardized Z-score that tells us how much a number varies relative to the mean and standard deviation of its home set. The Oslo Ideology Project (Rand-Hendriksen, 2008) has previously used a relative frequency to get a similar measure:

$$R_i = \frac{A_i}{\bar{A}}$$

R_i = Relative frequency for word R, year i

A_i = Adjusted frequency for word A, year i

\bar{A} = Average adjusted frequency for word A, all years

The relative frequency is obtained by simply dividing the adjusted annual frequency for a word by the average adjusted frequency for that word across all years. This gives a small number that usually varies between 0.5 and 2, exceeding 2 only when the actual search word fluctuates immensely (i.e., when a word increases in one year so much that it reaches more than twice the average frequency of the usage of that word). The standardized Z-score can be a negative number and often will vary somewhat even when calculated from words that do not vary very much, and it has a slightly different interpretation; while relative frequency tells us how much a word varies relative to the average adjusted frequency of usage, the standardized score tells us how many standard deviations the word is from the average. When represented graphically, the shape of the graph will be identical to the adjusted frequency itself. As with the adjusted and relative frequencies, only the scale of presentation varies.

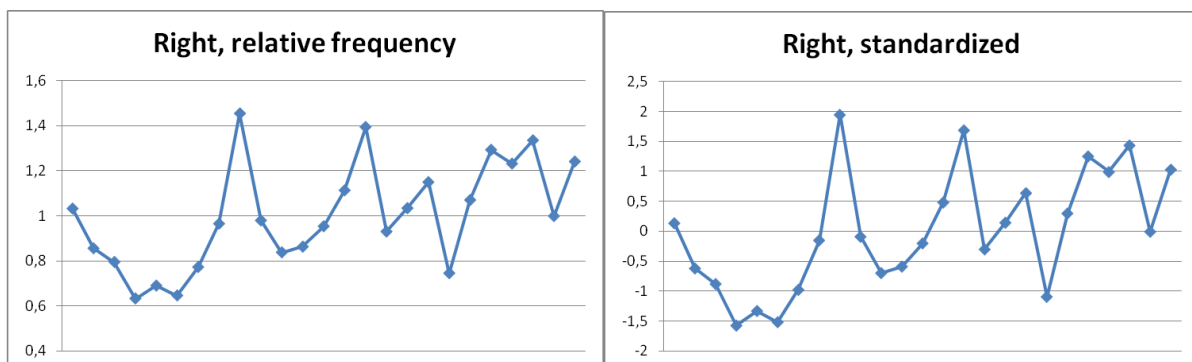


Figure 1: *Standardized and relative frequencies, 'right', Norwegian dataset.*

Proportional frequency: When comparing words across different datasets, it may be problematic to interpret the adjusted frequencies directly. The adjusted frequencies of a word do not in themselves say anything about proportion of use of that word. A word being used in 100 articles in an archive consisting of only 100 articles could potentially have the same adjusted frequency as a word being used in 100 articles in an archive consisting of 10 000 articles. The interpretation of these scores should, however, be radically different due the different proportion of usage of the word. In order to scale the adjusted frequency so that it is

more easily comparable with adjusted frequencies from other datasets, proportional frequency is used.

$$P_i = R_{Pi} \frac{\bar{A}_P}{\bar{N}} * 100$$

P_i = Proportional frequency for word P, year i

R_{Pi} = Relative frequency for word P, year i

\bar{A}_P = Average adjusted frequency for word P, all years

\bar{N} = Average number of articles in archive all years

This is simply an adjustment to the relative frequency where it is multiplied by the average proportion of use. I have multiplied the proportion by 100 to create a percentage score. Again, proportional frequencies will have the same developmental profiles as the adjusted frequencies. Proportional frequencies tell us how many percent of the articles contain the given word. Proportional frequencies thus vary between 0 and 100.

Levels of usage

In addition to studying the development of word frequencies, it is interesting to consider the levels of usage for the different words. This means looking at how frequently the words are used, relative to the number of articles in the archive.

Two datasets may show the same frequencies of use for the same word a given year and the same change in frequency over several years. If these datasets represent archives with different numbers of articles, however, our interpretations of the word developments should be different. A search word appearing in 100 in out of 1000 articles should be interpreted differently from a search word appearing in 100 out of 10 000 articles. In this thesis, the measure of proportional frequency is used to show how many percent of the articles in an archive have occurrences of a given search word.

As previously discussed (pp. 9-10) it is assumed that word frequencies will vary along with variations in ideology. It is furthermore predicted that this variation will take form of higher word frequencies occurring for words that reflect the dominant ideologies. This means that we expect higher levels of use for words that reflect the dominant ideologies. These words should

appear high on the rank order when sorting words by usage levels and societies with a strong slant towards one dimension of ideology should have higher usage levels for words reflecting that dimension than other societies.

It is important to be careful when making direct comparisons of the same search words across different datasets. Different translations will result in slightly different meanings for the search words (Harkins & Wierzbicka, 2001). One should therefore not only look at the proportion of use for one word across different datasets. The relative size of differences between usage levels of different words within the same database, as well as the rank order, should be taken into account when interpreting usage levels.

Level of usage is presented in this thesis by averaging the proportional frequency from 1998-2008. I have chosen to average over a decade in order to get a more stable measure of a society's usage level. The usage levels are meant to show more stable traits of the datasets. One could have used a longer span, but this would make comparison difficult because we only have available Czech data for these 11 years.

Principal Component Analysis

Principal component analysis (PCA) is a method for data reduction. PCA is similar to factor analysis, but where factor analysis aims at explaining covariance among observed variables by unobserved underlying phenomena, PCA merely involves a linear transformation of the data, reducing an observed correlation matrix to fewer components (DeVellis, 2003, Pedhazur & Schmelkin, 1991). In the present thesis, PCA is used to identify similarity in development trends. PCA has not been used in previous analyses by the Oslo Ideology Project, but it is a valuable tool when working with large amounts of data. In the present thesis I examine only 12 words in three societies, an amount of data that could be handled without PCA.

The PCA was performed on word-by-word correlation matrices and correlations were computed from word (column) by year (row) data matrices. The correlation coefficients may thereby be interpreted as measures of similarity in developmental profiles. If all words show the same development across time, one component would explain most of the variance, and theoretically, PCA should return one component for each distinctive development trend in the dataset. Words that follow similar developments across time will have similar component

loadings. Words that follow opposite developments will have similar component loadings, but with opposite signs.

In practice, however, PCA is very sensitive to small fluctuations in developmental trends, and gives little information about the specific profiles. Profiles may deviate from each other due to non-linearity or because similar trends may have started at different points in time. Thus, in order to draw conclusions about the trends in the dataset we will need to visually inspect the individual developmental profiles. I will use PCA primarily to identify developmental trends that are systematically different and to identify words that represent these main trends. Characteristic developmental profiles will be presented graphically.

Analyses and results

It is expected that the different datasets will reflect the different kinds of ideologies in the three selected societies described above. I will examine the presented methodology by testing theoretical expectations against the datasets from the chosen three societies, looking at development within each dataset as well as comparisons between the datasets. This thesis concerns itself with developing a methodology, not about analyzing and mapping out ideologies in depth. Thus, the examination of data aims at checking whether our method captures the essence of the ideological developments in each society. Finally, I will discuss whether the method has captured the societies' different ideological developments in an adequate manner, using our knowledge about the ideological situation in the three different countries.

Basic analysis of development over time

In order to familiarize the reader with the datasets, let us first take an overview by looking at some basic properties of the word developments. The greatest strength in using these kinds of simple presentations of the data is the closeness between the data and the presentations. When interpreting these kinds of datasets, one can risk removing oneself too far from the natural language they are based on. When interpreting more advanced statistical representations of the data, then, we need to be careful. Simple analyses like listing how many percent increase or decrease there has been in usage frequency of a word, on the other hand, can be interpreted more directly.

Table 1: *Percent change in adjusted frequencies*

	Norway	US	Norway	US	Czech
Timespan	1984-2008	1984-2008	1998-2008	1998-2008	1998-2008
I/me	47.7%	49.6%	-4.1%	10.3%	6.6%
Right	20.2%	72.0%	-11.0%	11.4%	-25.9%
Duty	-39.7%	20.9%	-24.5%	-8.9%	-16.4%
Responsibility	0.7%	11.1%	-18.4%	1.2%	-22.9%
Common	-25.6%	14.0%	-24.5%	0.1%	-22.1%
Solidarity	-62.3%	-43.6%	-23.0%	33.4%	-57.5%
Cohesion	-41.9%	87.6%	-22.6%	-23.0%	-59.4%
User	68.4%	180.1%	17.8%	20.0%	5.8%
Citizen	-42.5%	28.9%	-34.2%	18.3%	0.5%
We/us	10.4%	56.1%	-5.8%	17.3%	-9.3%
Equality	-25.7%	27.3%	-27.5%	36.3%	-1.0%
Justice	-32.6%	18.6%	-22.9%	-19.2%	-15.1%

Table 1 shows the percent increase or decrease of the 12 words for the entire timespan covered by the three datasets. The percentages were calculated by subtracting the adjusted frequencies the first available year from the adjusted frequencies the last available year and dividing by the adjusted frequencies the first available year. The numbers thus represent percentage change relative to the first available year in the dataset. Comparison between countries is difficult when the timespans are different, so I have included both the entire timespan for the US and Norway datasets and percent change for the past 11 years. This is so that the changes can be compared with the Czech dataset.

From looking at table 1, we observe some immediate trends. For example, we see that the US dataset shows an increase of all words but one from 1984 to 2008 and that all but three words in the Czech dataset show a decrease from 1998 to 2008. Table 1 gives an easy and quick overview, and we will return to this table later in the analyses.

Table 2 lists the correlation between the adjusted word frequencies and year (1984, 1985, 1986 and so on). As with the percent change, I have calculated correlations for the Norwegian and US data both for the entire timespan and for 1998-2008. Asterisks denote correlations that are statistically significant at an alpha level of .05.

We have calculated correlations between word frequency and year, and the correlations are thereby indices of linear change in word frequencies across time. High correlations imply linearity. On the other hand, low correlations are not necessarily indicating “no systematic trend”. Word frequencies may show other kinds of systematic development trends: as non-linear or cyclic trends. Trends that are systematic but non-linear will also produce low correlations. From table 2 we see, however, that the majority of word frequencies follow a more or less linear development.

Table 2: *Correlation between adjusted frequencies and time*

	Norway	US	Norway	US	Czech
Timespan	1984-2008	1984-2008	1998-2008	1998-2008	1998-2008
I/me	.788*	.975*	.104	.894*	.779*
Right	.558*	.976*	.174	.811*	-.859*
Duty	-.885*	.787'	-.707*	.298	-.683
Responsibility	-.462*	.749*	-.632*	.611*	-.896*
Common	-.733*	.636*	-.885*	.120	-.756*
Solidarity	-.875*	-.581*	-.845*	.029	-.946*
Cohesion	-.893*	.529*	-.337	-.319	-.837*
User	.830*	.888*	.199	-.017	-.620
Citizen	-.882*	.597*	-.727*	.851*	-.834*
We/us	-.240	.987*	-.205	.982	-.354
Equality	-.805*	.738*	-.844*	.600*	-.249
Justice	-.723*	.719*	-.653*	-.576*	-.720*

* Statistically significant at alpha level of .05

Analysis of the Norwegian dataset

Norway is a country where egalitarianism and social equality have traditionally been strong values, but these are changing and being replaced by more individualistic values (cf. Nafstad et al., 2007a, 2009b, 2009c). The Norwegian dataset is based on language use in the largest newspaper in the country, Aftenposten. Based on the given descriptions of the ideology in Norway, it is expected that words reflecting communal values in the dataset have had high levels of use. However, communal words are expected to show a trend of decrease, while individualistic words are expected to increase in use.

I have performed a principal components analysis using the adjusted frequencies of the words in the Norwegian dataset. The solution is shown in Table 3 below. The Kaiser criterion for extraction was used, and two components were extracted. The first component accounted for 42.7 % of the variance, the second component 42.4 %.

Table 3: *PCA of Norwegian data, rotated solution sorted by loadings on first component*

	Component	
	1	2
Equality	0.955	0.225
Citizen	0.942	-0.065
Common	0.935	0.294
Solidarity	0.931	-0.149
Duty	0.930	-0.072
Cohesion	0.870	-0.289
Justice	0.853	0.088
Responsibility	0.735	0.636
We/us	0.598	0.738
I/me	-0.501	0.833
Right	-0.529	0.407
User	-0.822	0.394

From looking at table 3, we see that individualistic words ('I/me', 'right', 'user') and communal words have opposite loadings on component 1 from the rest of the words. This indicates that the individualistic words have had opposite development trends from the other, communal, words. To see what these trends are, one needs to look at graphical mappings of the words. I have chosen to show graphs of 'duty', 'responsibility', 'cohesion', and 'I/me', as the component loadings show that these are representative of the main trends.

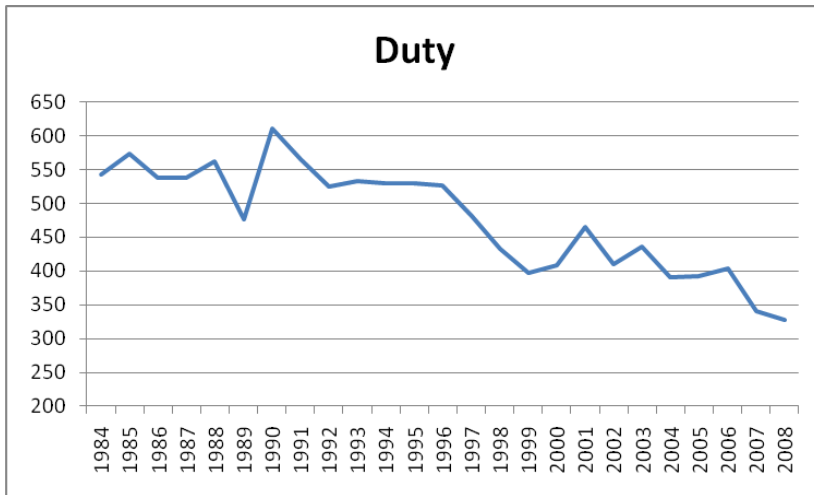


Figure 2: Graph of adjusted frequencies, 'duty', Norwegian dataset

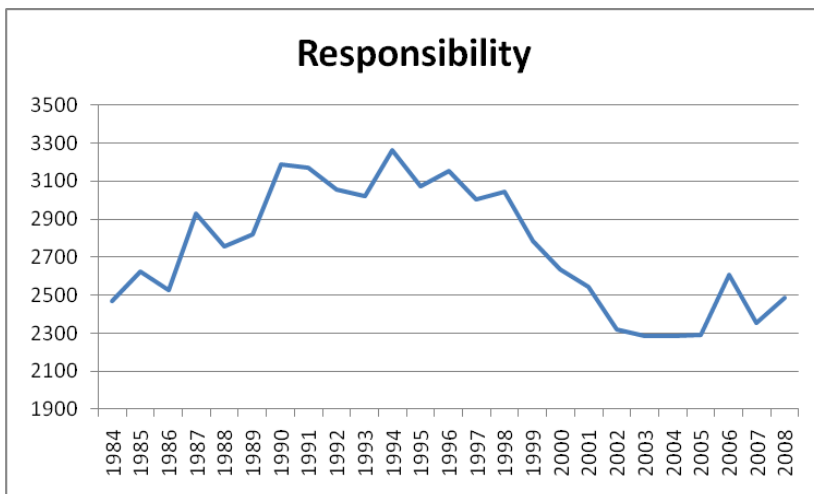


Figure 3: Graph of adjusted frequencies, 'responsibility', Norwegian dataset

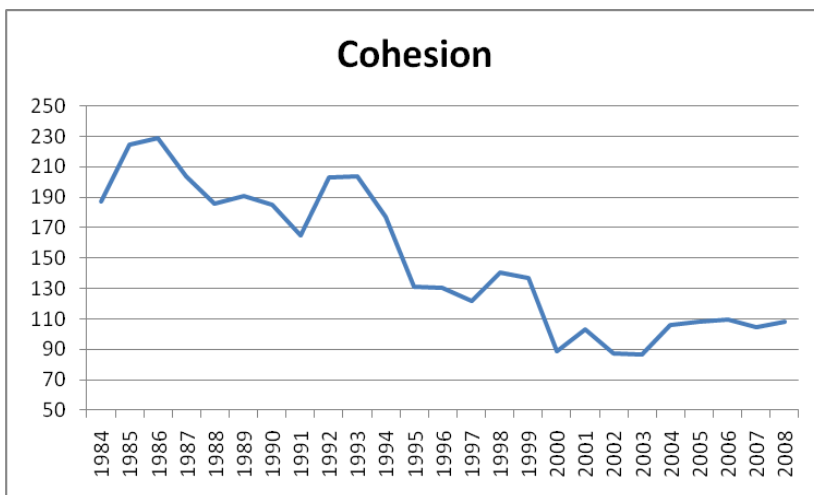


Figure 4: Graph of adjusted frequencies, 'cohesion', Norwegian dataset

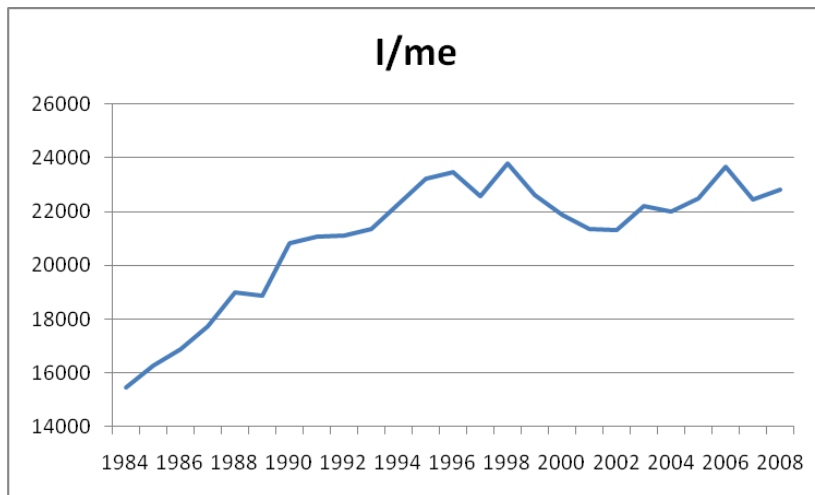


Figure 5: *Graph of adjusted frequencies, 'I/me', Norwegian dataset*

From the graphs we identify the two trends as one trend of declining frequencies and one of increasing frequencies. The trends appear rather stable, as there are no sudden peaks in the graphs.

The PCA returned two components, which might indicate that there are two kinds of development trends. We have seen that one of the trends is a trend of almost linear increase or decrease, as evident from figures 2 ('duty') and 5 ('me'). However, as previously discussed, PCA is somewhat sensitive, and small differences in development trends may lead to the PCA extracting more components to account for the differences. Figure 3 ('responsibility') shows that there are at least some deviations from linear increase or decrease in the Norwegian dataset. However, from the PCA, graphs and the table of correlations (table 2) taken together, we may conclude that the Norwegian dataset shows mainly stable, linear trends of increase or decrease.

Table 4 lists the average percent of articles using each word the past 11 years. Looking at these levels of use we see that communal words have a higher usage level than individualistic words. From the PCA and graphs we see that there is a trend of decrease for communal words and a trend of increase for individualistic words. The only two communal words to show increase are 'responsibility' and 'we/us', though these increases are small. From looking at figure 3, the graph of 'responsibility' we see that this increase also appears more unstable than the increase of the individualistic words.

Table 4: 11-year average proportional frequencies, Norwegian dataset

Level of usage	
We/us	39.04
I/me	28.06
Responsibility	3.10
Common	2.49
Citizens	2.01
Equality	1.12
Duty	0.50
Users	0.45
Solidarity	0.38
Justice	0.23
Cohesion	0.13
Right	0.07

In summary, communal words have a high level of use in the Norwegian dataset, but are decreasing. Words reflecting individualistic values are on the rise. This fits well with what we know about ideology in Norway: a long tradition of egalitarianism shifting towards a more individualistic value set. Our method reflects both the general level of values in Norway and the direction of change.

Analysis of the US dataset

The US is a strongly individualistic nation, valuing personal autonomy and freedom. It is also thought to be the home of some of the driving forces behind globalization, for example through mass media (van Dijk, 1998) or financial dominance (Stiglitz, 2002). The dataset is gathered from the New York Times, which is one of the largest newspapers in the US. We expect from the data that words reflecting individualistic values will have a high level of use and that the data should not show great fluctuations. The US ideology should register on our analysis as an ideologically stable, highly individualistic society.

When running a PCA of the US dataset using SPSS, two components are extracted. The first component accounts for 40.6% of the total variance and the second component accounts for 36.0%.

Table 5: PCA of US data, rotated solution sorted by loadings on first component

	Component	
	1	2
Justice	0.837	0.305
Duty	0.771	0.406
Social cohesion	0.747	0.137
I	0.685	0.706
Right	0.678	0.714
We	0.678	0.713
Users	0.607	0.668
Responsibility	0.519	0.539
Equality	0.386	0.688
Common	0.364	0.717
Citizens	-0.134	0.929
Solidarity	-0.837	-0.113

From table 5, we see that most of the words load more than .3 on both component 1 and 2, with many words showing loadings of .5 and more. The only exceptions are ‘solidarity’, ‘cohesion’ and ‘citizens’. This indicates that most words follow a similar development trend. I have chosen to display graphs for the words ‘I’, ‘right’, ‘citizens’, ‘solidarity’ and ‘justice’.

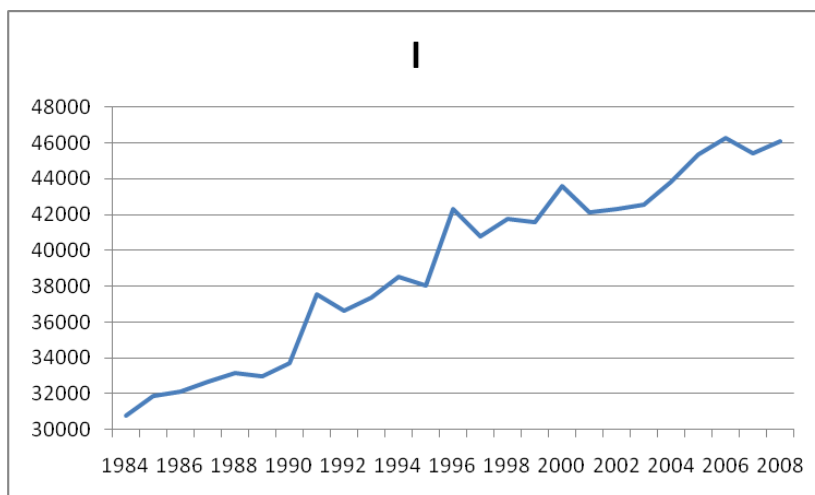


Figure 6: Graph of adjusted frequencies, ‘I’, US dataset

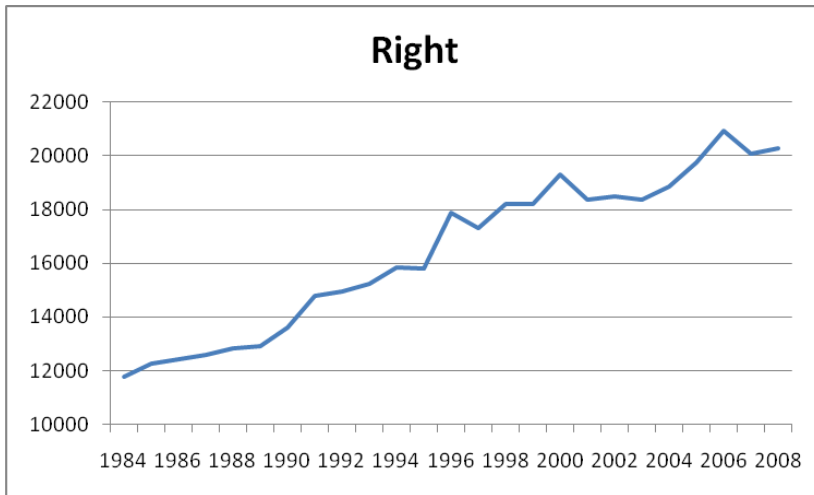


Figure 7: Graph of adjusted frequencies, 'right', US dataset

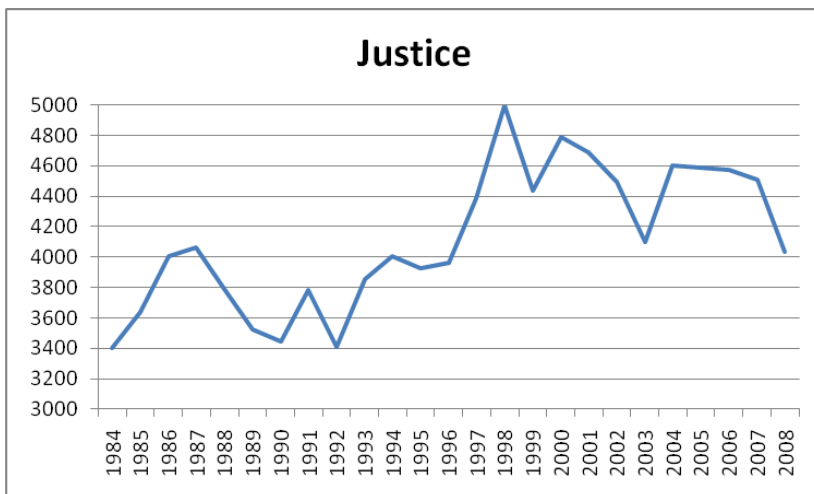


Figure 8: Graph of adjusted frequencies, 'justice', US dataset

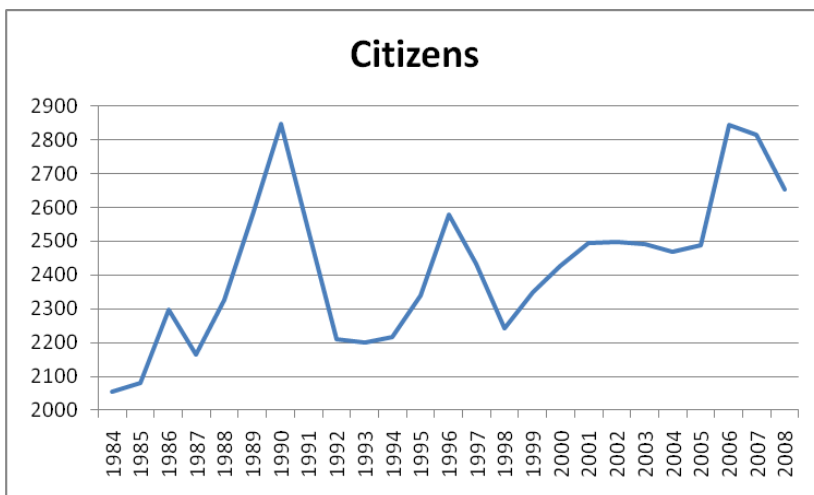


Figure 9: Graph of adjusted frequencies, 'citizens', US dataset

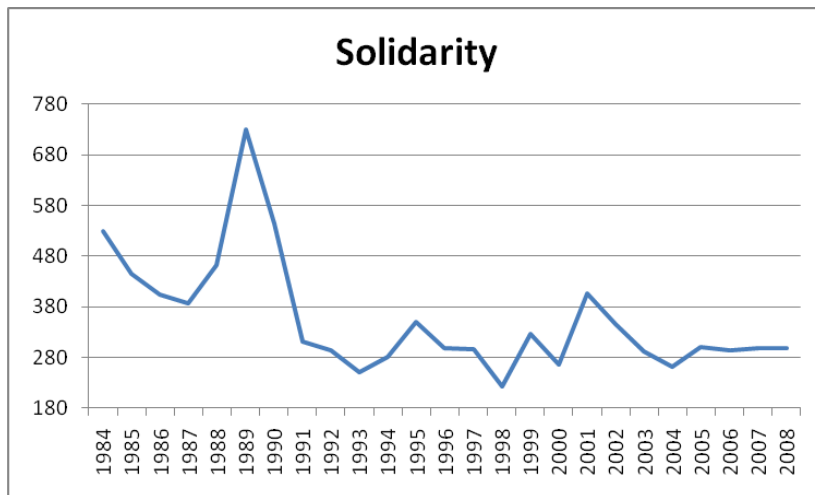


Figure 10: *Graph of adjusted frequencies, 'solidarity', US dataset*

From looking at the PCA results in table 5 and the graphs, we see that most words follow a trend similar to 'I'. From the graphs we can conclude that this is a trend of increase. The PCA and graphs together with the correlations in table 2 tell us that most of the words from the US dataset show a relatively stable and mostly linear increase. The only word to show decrease is 'solidarity'.

Table 6: *11-year average proportional frequencies, US dataset*

Level of usage	
I	49.21
We	43.76
Right	21.59
Common	5.86
Justice	5.02
Responsibility	3.05
Citizens	2.86
Duty	1.92
Users	1.74
Equality	0.38
Solidarity	0.35
Social cohesion	0.02

Table 6 lists the average percent usage of words for the past 11 years. 'I' has the greatest proportion, being used in almost half of all articles and being used markedly more than 'we'. The high proportion of 'right' is harder to interpret, given that the word can have several

meanings. The very low proportion of ‘social cohesion’ means it is also somewhat hard to interpret. As expected in a strongly individualistic society, ‘equality’ and ‘solidarity’ are low in level of usage.

In summary, all words apart from one increase in usage during the examined period. The increase is stable and mostly linear. ‘I’ show a higher level of usage than ‘we’. The stability of the increase and high level of usage for individualistic words fit well with what was expected from the US dataset.

Analysis of the Czech data

The Czech Republic is, as previously presented, a relatively new state that has recently changed from communism to liberal democracy. Ruling ideologies have thus undergone profound and rapid alterations, especially regarding the individualism-communality dimension. There is also a growing disillusionment with politicians’ ability to solve the country’s problems. The dataset is gathered from Mladá fronta DNES, one of the largest newspapers in the Czech Republic. From the data, we expect that most of our ideological search words should show a general decline, as the examined dimensions of ideology are becoming less dominant. We also expect that the development trends should be unstable.

Table 7: *PCA of Czech data, rotated solution sorted by loadings on first component*

	Component		
	1	2	3
Right	0.955	-0.172	0.168
Common	0.955	-0.047	0.167
Duty	0.908	-0.051	0.375
We/Us	0.907	0.389	-0.057
Responsibility	0.883	-0.241	0.201
Solidarity	0.849	-0.457	-0.117
Users	0.567	0.249	0.222
Justice	0.467	-0.316	0.506
Equality	0.123	0.093	0.936
I/Me	0.108	0.958	-0.155
Citizens	-0.076	-0.684	0.574
Cohesion	-0.221	0.822	0.045

I have performed a PCA of the Czech data with SPSS. Using Kaiser's criterion, three components are extracted. The first component explains 47.1% of the variance, the second 23.1 %, and the third 13.1 %.

From looking at table 7, we see that the top five words have similar loadings on the first component. However, loadings on the second and third components vary. Looking at the rest of the words, we see that few words load equally on the same factors. This indicates that the words in the Czech dataset show several different development trends.

To examine the development trends, I have chosen to graph 'right', 'users', 'cohesion' and 'equality'. These words show different loadings on the PCA, but because of the high variation in component loadings, the graphed words do not represent the dataset as well as the graphed words in the Norwegian or US dataset.

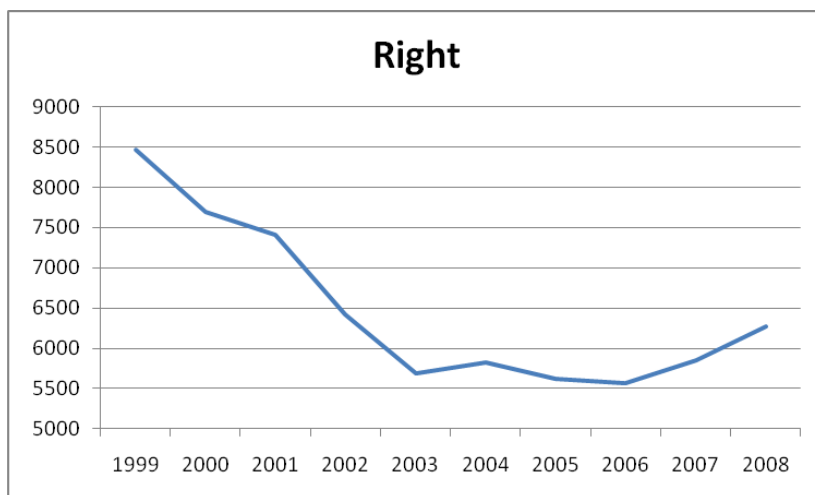


Figure 11: *Graph of adjusted frequencies, 'right', Czech dataset*

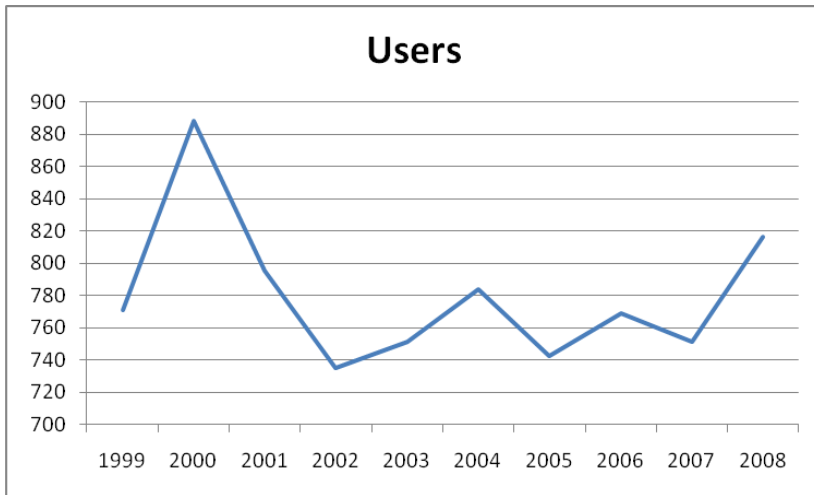


Figure 12: Graph of adjusted frequencies, 'users', Czech dataset

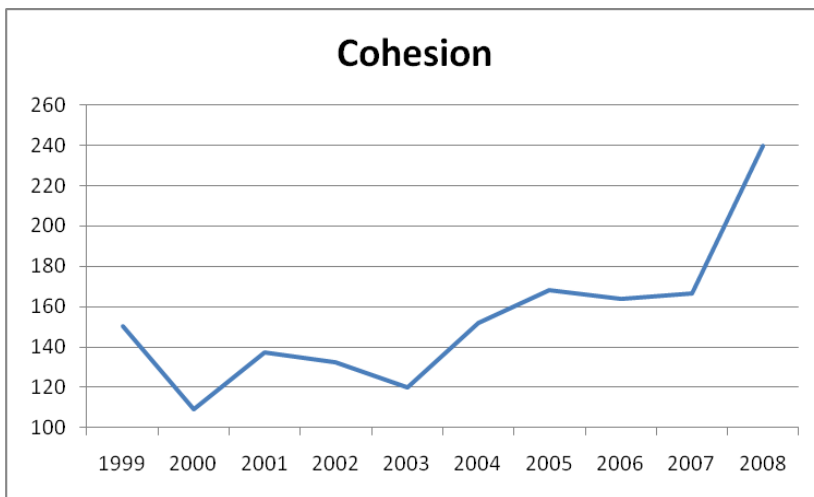


Figure 13: Graph of adjusted frequencies, 'cohesion', Czech dataset

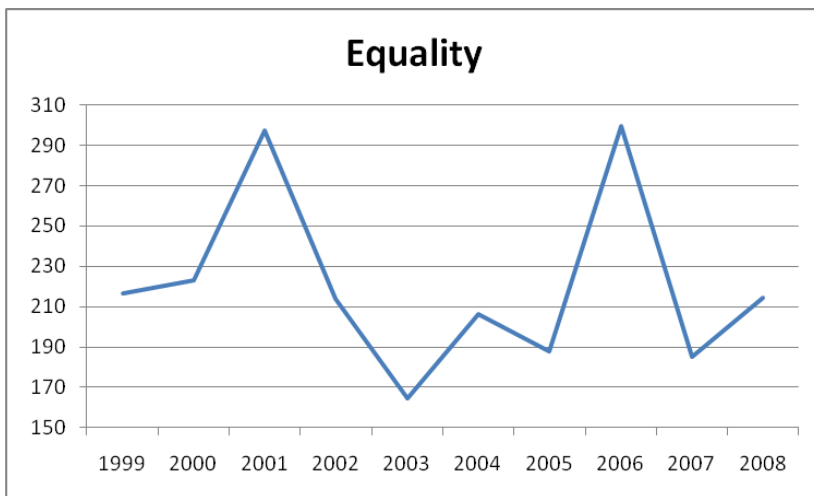


Figure 14: Graph of adjusted frequencies, 'equality', Czech dataset

The PCA and graphs indicate that there is some variation in the trends in this dataset. The graphs show unstable development trends with sudden peaks, especially for words ‘users’ and ‘equality’.

Table 8: 11-year average proportional frequencies, Czech dataset

Level of usage	
We/us	22.71
I/me	16.48
Common	3.33
Right	3.29
Citizens	2.52
Duty	1.61
Responsibility	1.15
Justice	0.59
Users	0.40
Solidarity	0.22
Equality	0.12
Cohesion	0.08

Table 8 lists the 11-year average percent usage of words. From table 1 on page 27 we see that all words show decrease except for ‘users’, ‘citizens’ and ‘ I/me’.

From the analyses so far, we see that the development trends in the Czech dataset are more unstable than the other two datasets. As was expected, there is a downward trend for the majority of the search words. Our method captures the most essential aspects of the ideology development regarding individualistic and communal values in the Czech Republic.

Comparisons between datasets

These analyses of each separate dataset show that our archive method captures the essence of the development of ideology regarding individualistic and communal values in each of the three societies. We now turn to a comparison of the datasets from the three societies. Based on the previous presentations, we would expect that:

The US dataset will demonstrate the highest level of individualistic word usage.

The Norwegian dataset will show the highest level of communal word usage.

The Czech dataset will demonstrate the greatest fluctuations in developments of word usage.

Table 9: Average proportional levels of use, all countries, 1998-2008

	Czech	US	Norway
Citizens	2.52	2.86	2.01
Cohesion	0.08	0.02	0.13
Common	3.33	5.86	2.49
Duty	1.61	1.92	0.50
Equality	0.12	0.38	1.12
I/me	16.48	49.21	28.06
Justice	0.59	5.02	0.23
Responsibility	1.15	3.05	3.10
Right	3.29	21.59	0.07
Solidarity	0.22	0.35	0.38
Users	0.40	1.74	0.45
We/us	22.71	43.76	39.04

As previously discussed, one must take care when comparing levels of usage for single words directly between datasets. There may be translation issues, in which case the words may not be directly comparable, even if they capture the same general ideology dimension.

Comparison also requires that the datasets themselves must be comparable, meaning that they must represent the same kinds of media discourse. It may be easier to draw conclusions from the patterns that emerge when analyzing several words than to compare single words. Bearing that in mind, there are some interesting differences that emerge when we compare the usage levels from the different datasets.

From table 9 we see that the US dataset has the highest levels of usage of individualistic words ‘I/me’, ‘users’ and ‘right’ (though the high proportion of ‘right’ is probably due to synonymous meanings, see footnote on p.12). Moreover, we see that the Norwegian dataset has a high level of use of communal words. Interestingly, however, these levels are only barely above or even below US levels of usage. Usage levels of the communal words ‘cohesion’, ‘equality’, ‘responsibility’ and ‘solidarity’ are highest in the Norwegian dataset.

The present thesis has not suggested any expectations or concrete hypotheses for the specific development trends of separate words. However, some of the words show especially interesting differences in levels of use, and I will in the following briefly comment upon these words. I will also display graphs of the developments in proportional frequencies.

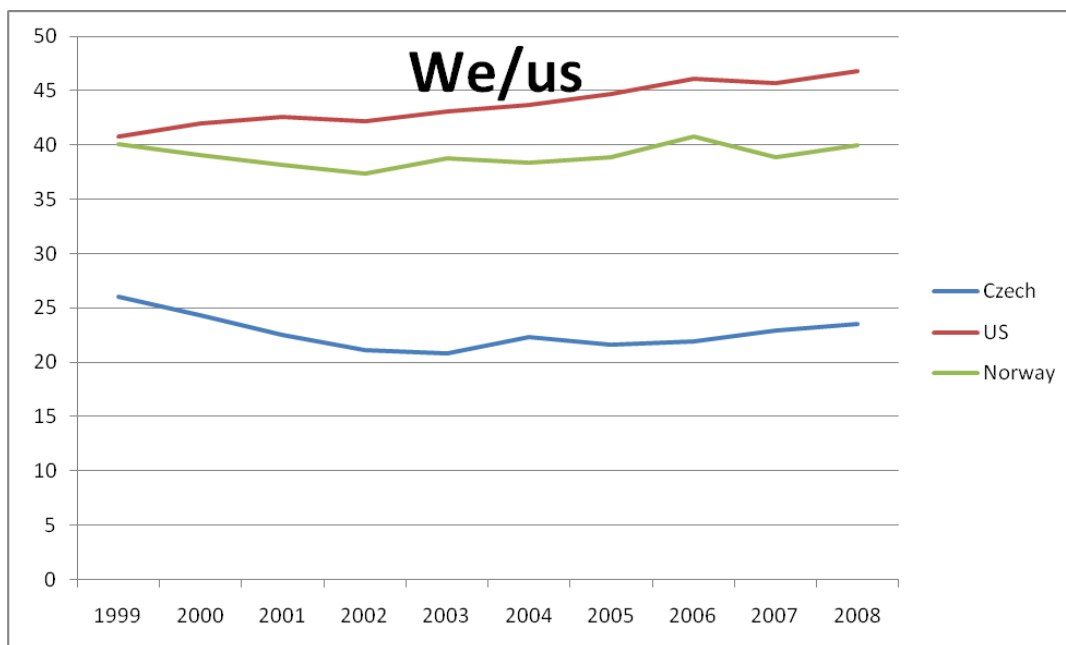


Figure 15: *Proportional frequencies, 'we/us'*

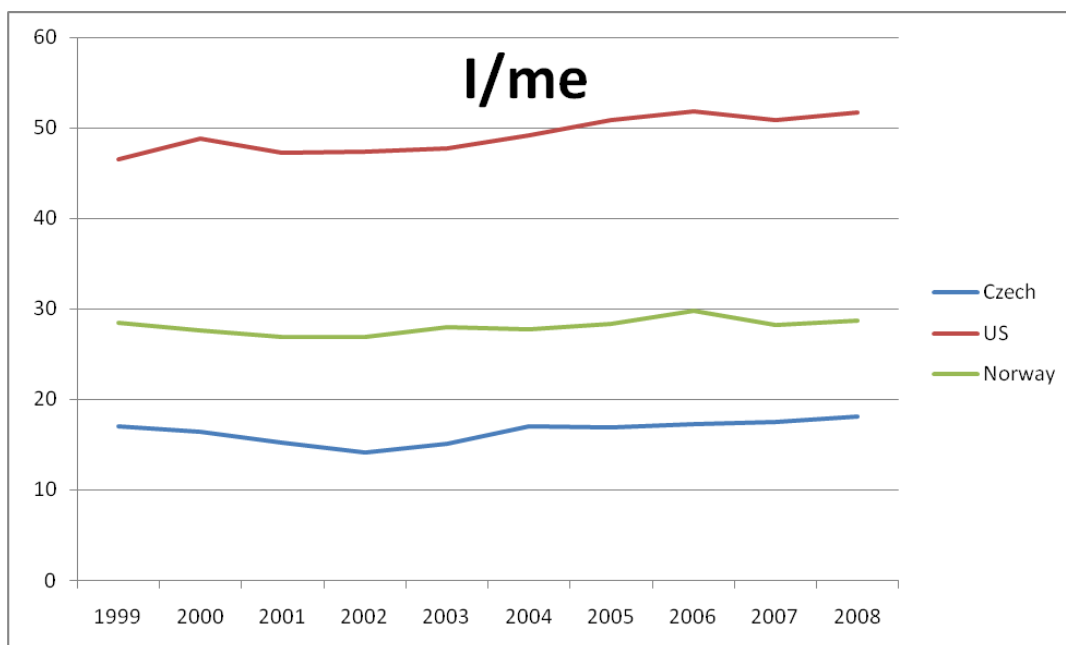


Figure 16: *Proportional frequencies, 'I/me'*

Proportional use of 'we/us' and 'I/me': As expected, the US dataset shows the highest level of usage of 'I/me'. The US dataset also has the highest level of 'we/us'. The Norwegian dataset, however, shows a higher relative difference in usage of 'I/me' and 'we/us'. Looking at the Norwegian dataset, we find that the difference in level between 'we/us' and 'I/me' is 10.98 percent while the difference is 5.95 percent for the US dataset. Interestingly, US levels

of use for ‘we/us’ started close to Norwegian levels but showed a relatively large increase where the Norwegian levels showed a slight decrease. The Czech dataset has a lower level of usage of both ‘I/me’ and ‘we/us’, but this is probably due to personal pronouns being used differently in Czech compared to English and Norwegian, as in Czech one can usually tell the subject of a sentence from the form of the verb. The explicit pronoun is therefore often omitted. We did not have any specific expectations regarding the levels of use for ‘I/me’ and ‘we/us’ for the Czech dataset. The relative differences between usage levels for ‘I/me’ and ‘we/us’ do, however, indicate that the Czech republic is in an intermediate position between the US and Norway, being less individualistic than the US, but more individualistic than Norway.

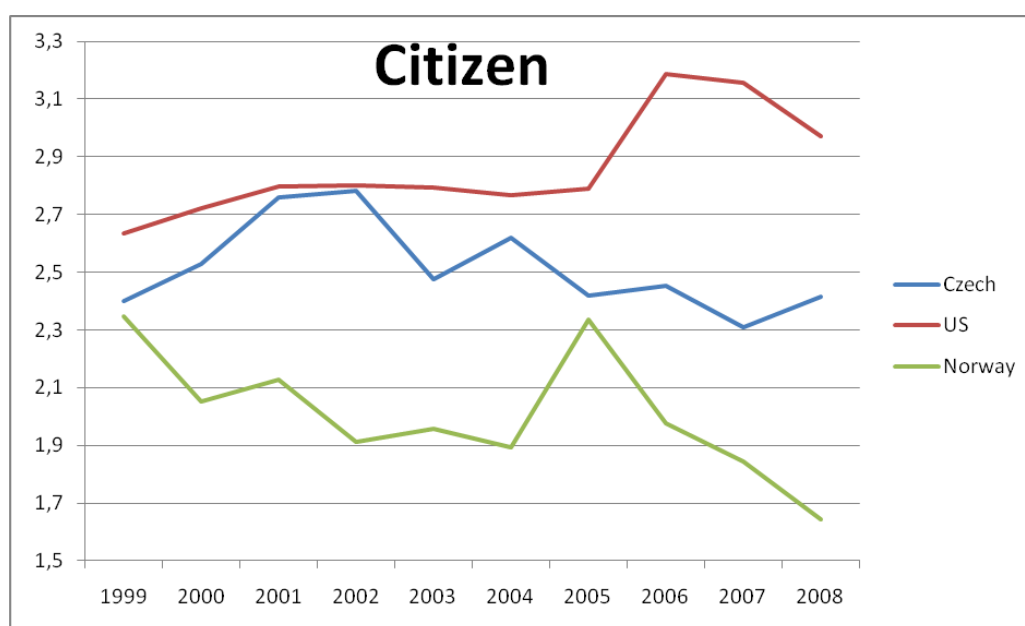


Figure 17: *Graph of proportional frequencies, 'citizen'*

Proportional use of 'citizen': As can be seen from figure 17, usage of ‘citizen’ in the US was on a stable level from 1999-2005, but showed a sharp increase in 2006 and 2007, before dropping again. Conversely, there was a trend of decrease in Norway during the same period. In the Czech Republic, development was relatively stable, with a very small decrease in use (see table 1 on p. 27). What ideological developments could this reflect? Most likely, there has been a debate in the US about citizenship, where the dominant ideology may be changing towards prescribing a more active political participation from US citizens. The gap between the usage levels in Norway and the US has increased in the period, perhaps signaling that

ideologies on citizenship in these countries are becoming gradually more different from each other.

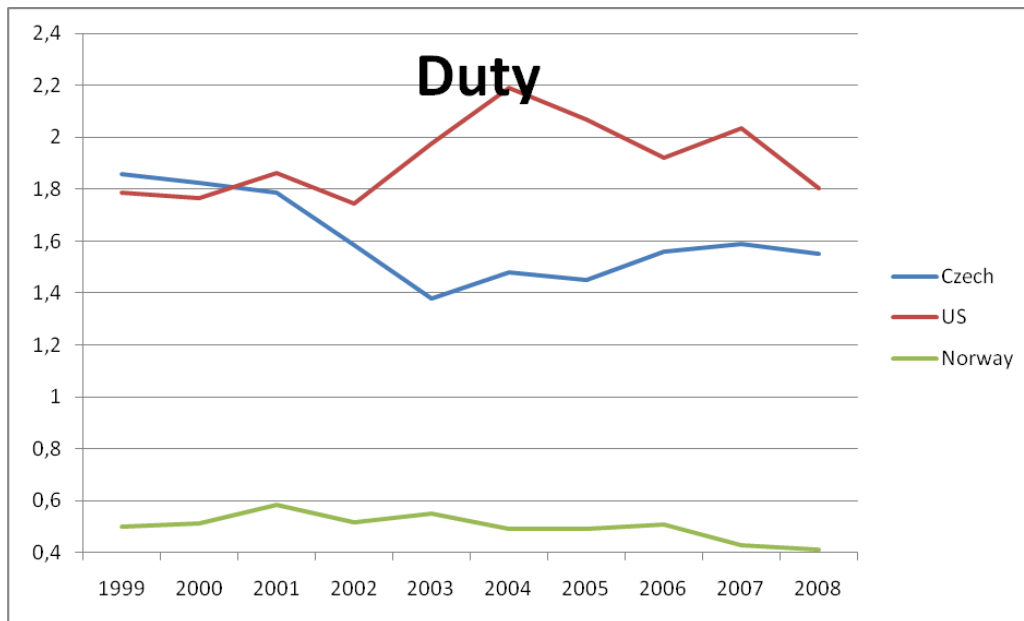


Figure 18: *Graph of proportional frequencies, 'duty'*

Proportional use of 'duty' : Figure 18 illustrates how the Czech and US datasets present very different patterns of usage of 'duty' than the Norwegian dataset. 'Duty' has a much higher level of usage in the US and Czech datasets (3.82 and 3.22 times higher, respectively) than in the Norwegian dataset. This fits well with what we know about the different ideologies. In a welfare state like Norway it would be expected that personal duty is less of an issue, as the state plays a larger role in taking care of the well-being of its citizens. A question to be raised is whether this is due to translation issues – there is a possibility that the different translations of 'duty' are not exactly equivalent. It would, however, be interesting to compare the levels of usage of 'duty' in the presented datasets with usage levels in other countries. Do the US and Czech levels represent a 'normal' level of usage or are they unusually high? Another question that arises from the differences is why Czech society would be almost equally concerned with duty as the US, when we expect Czech ideology to be dominated by disillusionment.

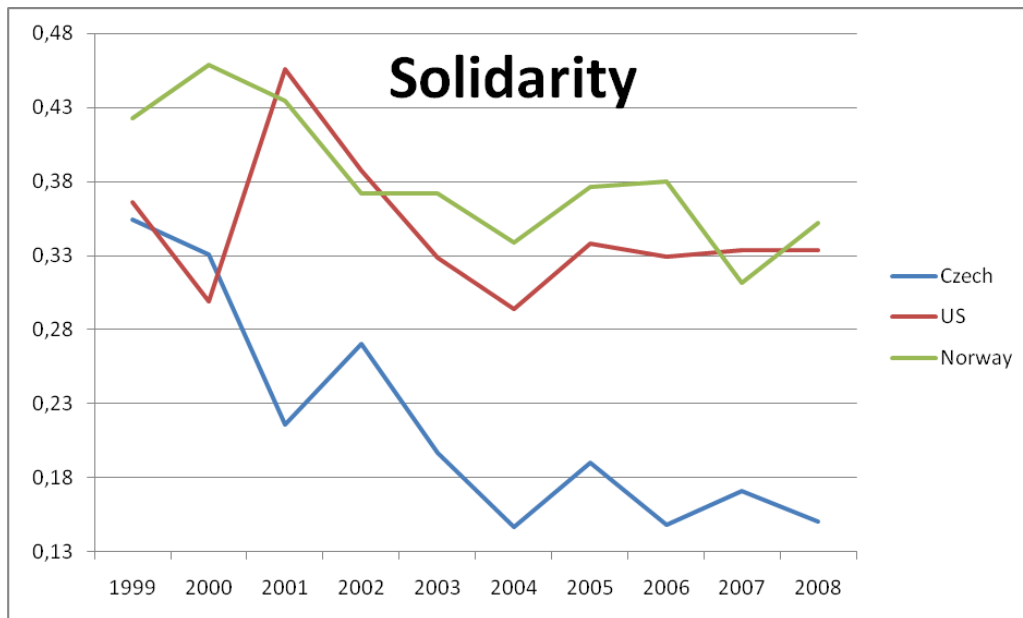


Figure 19: *Graph of proportional frequencies, 'solidarity'*

Proportional use of 'solidarity': Changes in the usage frequency of 'solidarity' show interesting trends when we compare the three datasets, despite this specific word being used in less than half percent of the articles in any of the archives. From figure 19 we see a sharp decrease in the Czech dataset and a smaller decrease in the Norwegian dataset. The US dataset shows a stable level with a peak in 2001. The decrease in the Norwegian dataset was expected, and the relative stability of the decrease fits well with the known trend of a steadily growing individualism. Surprisingly, however, Norwegian levels dropped below US levels in 2007 before rising again. The Czech drop in usage level would be expected both from the rise of "wild capitalism" (see discussion on p. 20) in the country and from the growing sense of disillusionment.

Discussion: Limitations, future directions and concluding remarks

We have seen that the presented archival method captures the essence of ideology development in the three societies. Through analysis of the Norwegian dataset, we were able to quantitatively describe the ongoing shift in ideology toward more individualistic values. The shift showed up on our PCA as the individualistic and communal words having different trends; while communal words either decreased or increased in a non-linear fashion, individualistic words increased in a linear fashion. We were also able to describe the existing ideology of egalitarianism through examination of usage levels for communal and individualistic words. Through analysis of the US dataset, we were able to capture the strong

individualism that characterizes US ideology and describe it quantitatively. The high level of individualism showed up as high usage levels for individualistic words as well as low levels of usage for communal words. As expected, there was also a relatively high degree of stability in the developments. Through analysis of the Czech dataset, we were able to describe quantitatively the turbulent shifts in ideology that has characterized the country the past years. These shifts showed up on our PCA, which showed a greater diversity in developmental trends than the other datasets. Lastly, we conclude that we captured the growing trend of disillusionment by identifying a general trend of decrease in all ideology words as well as low levels of usage for certain words.

When evaluating the presented method, it is important to be certain that the archives used as basis for the datasets are of good quality. Is it possible to evaluate the quality of the archives post-hoc by looking at the analyses? One indication that the archives are of comparable quality is that many of the searched words have comparable levels of use. No search words (with the exception of 'right') show very different usage levels across datasets. Even the search words 'I/me' and 'we/us', who would be expected to have low levels of use in Czech, due to differences in grammar, show about half the level of use of the same words in the Norwegian and US datasets. Comparable levels of use could indicate that the archives reflect the same kinds of public discourse. While we assume that all kinds of media discourse reflect ideology, we do not assume that all kinds of media discourse reflect ideology in the same manner. Had we chosen a purely tabloid newspaper as one of our archives, for example, this might have posed a challenge, because that archive may have reflected public discourse in a different way than the other chosen archives. It may be hard in practice to find good archives for the societies one wants to study, but for the present analysis it seems that the chosen archives are good and give comparable representations of public discourse.

One possible source of error that has not been controlled for is article length. We do not know if the chosen archives differ in length of the articles. This is a potential challenge because archives with longer articles may have greater chance of the chosen search words occurring than archives with shorter articles. Previous studies have examined this effect in detail for the Norwegian dataset (Rand-Hendriksen, 2008, p.30). We do not, unfortunately, have similar knowledge about the other datasets.

Validity of translation is another issue. When comparing words across different cultures, we need to be certain that the words represent the same ideological dimensions. Translations have been conducted by native speakers, thus we can be reasonably sure that they that they are semantically equivalent; that they mean approximately the same. However, the essential property of translations when using the present archive method is that the different translations should capture the same ideology dimension. We have not conducted in-depth analyses of the words to determine whether this holds true. One suggestion could be to choose separate words that reflect the same ideology dimensions from each society, and then perform analyses on these ideologically equivalent clusters of words rather than semantically equivalent single words.

A further critical question is whether analyses of developments of usage frequencies of words, a purely quantitative, almost decontextual, method alone provide valid empirical descriptions of ideologies and ideological shifts. One way to validate the kind of macroanalysis we are conducting is to approach the same research problem from a microanalysis perspective. Analyses of macrostructures such as ideologies should be able to link the larger shifts in society to micro-instances of human interaction (Scheff, 1990, Hermans & Dimaggio, 2007). It would therefore be valuable to supplement the present method with more qualitative approaches. Discursive psychology has been discussed as an important source of theoretical inspiration, and discourse analysis of chosen words as they appear in the media language could help both to validate choice of search words and to aid in interpreting results. In turn, macroanalyses like quantitative descriptions of ideology development could inform microanalyses, as discourse analysis, providing a framework for more fine-grained analyses. However, macroanalyses can identify potentially interesting ideology developments that would be missed in analyses of smaller amounts of text.

Having shown that our archival method captures the essence of ideology development across different societies, the next step would be to fine-tune the method in order to capture and describe ideology development in greater detail. One way to do this would be to increase the number of search words. By using principal components analysis, we would be able to identify clusters of words that show the same developments. These clusters can be used as basis for calculating levels of usage and creating graphs of representative words to obtain a basis for a more large-scale analysis of ideology. In the present thesis, our hypotheses have

been more general, specifying roughly what general trends we expected to see in the datasets. With larger datasets, it would be possible to formally test more specific predictions regarding ideology development. This could be useful for comparing different societies in a more precise, systematic manner.

By using the present archive method, it is also possible to analyze different kinds of archives. While mass media/newspaper language plays an especially important role in shaping and communicating ideologies in society, other archives of natural language use also reflect ideology or the meaning structures of society. Internet blogs are one example of emerging and easily accessible archive material. It is possible to search separately for blog posts from specific time periods and for posts published from specific regions. If one is careful in choosing of search words, tailoring the search to contain words that are suitable for the different archives' expressions of ideology, this could be another psychologically interesting way of analyzing ideology development cross-culturally.

The archive method of the Oslo Ideology Project has in the present thesis been supplemented with two new tools for examining ideology development across different societies; proportional levels of use and principal component analysis. The archive method has been tested on three datasets created from three archives representing three different societies. The archive method captures the most important general aspects of the ideology situations in each of the three societies. In addition, the method can be used to reveal in more detail how and in what ways the ideologies reflected in the various datasets from our three chosen societies are different. While it may be problematic to use this word count method as the sole basis for comparative studies of ideology, the archive method provides both tests of specific hypotheses regarding ideology and a basis for creating new hypotheses.

References

- Agnew, Hugh LeCaine (2008). *Češi a země Koruny české*. [Czechs and the lands of the Bohemian Crown]. Praha: Academia.
- Arnett, J. J. (2002). The psychology of globalization. *American Psychologist*, 57, 774-783.
- Bakhtin, M. M. (1952/1986). *Speech genres & other late essays*. Austin: University of Texas Press.
- Bauer, M. W. & Aarts, B. (2000) Corpus construction. In Bauer, M. W. & Gaskell, G. (Eds.) *Qualitative researching with text, image and sound: A practical handbook*. London: SAGE Publications.
- Bauman, Z. (2000). *Liquid Modernity*. Cambridge: Polity Press.
- Billig, M. (1987). *Arguing and thinking: A rhetorical approach to social psychology*. Cambridge: Cambridge University Press.
- Blakar, R. M. (1973/1996). *Språk er makt*. Oslo: Pax.
- Bourdieu, P. (1998). *Acts of resistance. Against the new myths of our time*. Oxford: Blackwell.
- Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist*, 1977, 513-531.
- Bronfenbrenner, U. (1979). *The Ecology of Human Development*. Cambridge, MA: Harvard University Press.
- Brewer, M. B. (2004). Taking the social origins of human nature seriously: Toward a more imperialist social psychology. *Personality and Social Psychology Review*, 8, pp. 107-113.
- Candish, S. & Wrisley, G. (2008). *Private Language*. Retrieved May 26, 2010, from The Stanford Encyclopedia of Philosophy Web site:
<http://plato.stanford.edu/entries/private-language/>
- Carlquist, E., Nafstad, H. E. & Blakar, R. M. (2007). Community psychology in a traditional Scandinavian welfare society: The case of Norway. In S. Reich, M. Riemer, I. Prilleltensky & M. Montero (Eds.). *International community psychology: History and theories*. (pp. 282-298). New York: Springer.
- Chiu, C. & Hong, Y. (2006). *Social Psychology of Culture*. New York: Psychology Press.
- Chrysoschoou, X. (2004). *Cultural diversity: Its social psychology*. Malden: Blackwell.

- Cullen, J. (2003). *The American Dream*. New York: Oxford university press.
- Curt, B. C. (1994). *Textuality and Tectonics: Troubling Social Science*. Buckingham: Open University Press.
- DeVellis, R. F. (2003). *Scale development*. Thousand Oaks: Sage Publications.
- van Dijk, T. A. (1998) *Ideology: A Multidisciplinary Approach*. London: Sage Publications.
- Eagleton, T. (1991). *Ideology*. London: Verso.
- Fairclough, N. (2006). *Language and globalization*. London: Routledge.
- Finkel, N. J. & Moghaddam, F. M. (Eds.). (2002). *The psychology of rights and duties*. Washington: American Psychological Association.
- Foucault, M. (1972/2007). *The archaeology of knowledge*. London: Routledge
- Gergen, K. J. (1985). The social constructionist movement in modern psychology. *American Psychologist*, 40, 266-275.
- Gledhill, J. (2004). Neoliberalism. In D. Nugent & J. Vincent (Eds.), *A companion to the anthropology of politics* (pp. 332-348). Malden: Blackwell.
- Harkins, J. & Wierzbicka, A. (2001). *Emotions in crosslinguistic perspective*. Berlin: Mouton de Gruyter.
- Harvey, D. (2005). *A brief history of neoliberalism*. Oxford: Oxford University Press.
- Hermans, H. J. M. & Dimaggio, G. (2007). Self, identity and globalization in times of uncertainty: A dialogical analysis. *Review of general psychology*, 11, 31-61.
- Hermans, H. J. M. & Kempen, H. J. G. (1998). Moving cultures. The perilous problems of cultural dichotomies in a globalizing society. *American Psychologist*, 53, 1111-1120.
- Hosainey, M. Master's Thesis. In preparation. Department of Psychology, University of Oslo.
- Hoyle, R. H., Harris, M. J. & Judd, C. M. (2002). *Research methods in social relations*. London: Wadsworth.
- Íñiguez, L. (1997). Discourses, Structures and Analysis: What Practices? In Which Contexts? In Ibáñez, T & Íñiguez, L. (Eds.) *Critical Social Psychology*. (pp. 147-155). London: SAGE Publications.
- Jenkins, J. P. (1997). *A history of the United States*. London: Macmillan Press.

- Jørgensen, I. S. & Nafstad, H. E. (2004). Positive psychology: Historical, philosophical and epistemological perspectives. In Linley, A. & Joseph, S. (Eds.) *Positive psychology in practice* (pp. 15-34). New York: Wiley.
- Kanner, A. D. & Soule, R. G. (2003). Globalization, corporate culture, and freedom. In T. Kasser & A. D. Kanner (Eds.). (2003). *Psychology and consumer culture*. (pp.49-67). Washington, DC: American Psychological Association.
- Kondo, D. K. (1987). Creating an Ideal Self: Theories of Selfhood and Pedagogy at a Japanese Ethics. *Ethos*, 15, 241-272
- Leitner, H., Peck, J. & Sheppard, E. S. (Eds.). (2007). *Contesting neoliberalism: Urban frontiers*. New York: The Guilford Press.
- Markus, H. R. & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98, 224-253.
- Murphy, G.L. (1991). The Psychology of Word Meanings. In Schwanenflugel, P.J. (Ed.) *The Psychology of Word Meanings*. Hillsdale, NJ: Lawrence Erlbaum Associates Publishers.
- Miller, J. G. (1997). Theoretical issues in cultural psychology. In Berry, J. W., Poortinga, Y. H. & Pandey, J. (Eds.), *Handbook of cross-cultural psychology*. (pp. 87-119). Boston: Allyn & Bacon.
- Nafstad, H. E. (2002). The neo-liberal ideology and the self-interest paradigm as resistance to change. *Journal of Radical Psychology*, 3, pp. 3-21.
- Nafstad, H. E. & Blakar, R. M. (2002/2006). Ideologier kartlagt gjennom mediespråket: En analysemetode. *Working report from the ideology project*, no 1 (pp. 1-23). Reprinted in R. M. Blakar (Ed.). *Språk er makt*. (pp. 262-290). Oslo: Pax.
- Nafstad, H. E. & Blakar, R. M. (2009a). Ideologier om fellesskapsordninger i en tid som prioriterer neo-liberalistiske verdier og forståelsesrammer. In H. E. Nafstad & R. M. Blakar (Eds.). *Fellesskap og individualisme*. (pp.135-148). Oslo: Gyldendal Akademisk.
- Nafstad, H. E., Blakar, R. M., Botchway, A., Bruer, E. S., Filkukova, P. & Rand-Hendriksen, K. (in press). Cross-cultural studies of communal values: A post-communist East-European republic, a Nordic welfare state and a modern West-African society. Paper to be presented at the 5th European Conference on Psychology, Copenhagen, June 23-26, 2010

- Nafstad, H. E., Blakar, R. M., Botchway, A., Rand-Hendriksen, K. (2009b). Globalization, ideologies and well-being: a study of a West African and a North European society. *The Journal of Positive Psychology, 4*, 305-315.
- Nafstad, H. E., Blakar, R. M., Carlquist, E., Phelps, J. M. & Rand-Hendriksen, K. (2007a). Ideology and power: The influence of current neo-liberalism in society. *Journal of Community and Applied Social Psychology, 17*, 313-327.
- Nafstad, H. E., Blakar, R. M., Carlquist, E., Phelps, J. M. & Rand-Hendriksen, K. (2009c). Globalization, neo-liberalism and community psychology. *American Journal of Community Psychology, 43*, 162-175.
- Nafstad, H. E., Blakar R. M. & Rand-Hendriksen, K. (2009d). The spirit of society and the virtue of gratitude: Shifting societal ideologies of gratitude. In Freire, T. (Ed.). *Understanding positive life. Research and practice on positive psychology.* (pp. 291-312). Lisboa: Climepsi Editores.
- Nafstad, H. E., Carlquist, E. & Blakar, R. M. (2007b). Community and care work in a world of changing ideologies. *Community, Work and Family, 10*, 329-340.
- Nelson, G. & Prilleltensky, I. (2005). (Eds.). *Community Psychology. In Pursuit of Liberation and Well-being.* New York, N. Y.: Palgrave Macmillan.
- Pedhazur, E. J. & Schmelkin, L. P. (1991). *Measurement, design and analysis: An integrated approach.* New Jersey: Lawrence Erlbaum Associates, Inc.
- Pennebaker, J. W., Mehl, M. R. & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology, 54*, 547-77.
- Pinker, S. (2002). *The Blank Slate.* London: Penguin Books Ltd.
- Potter, J. (1997). Discourse and critical psychology. In Ibáñez, T & Íñiguez, L. (Eds.) *Critical Social Psychology.* (pp. 55-67). London: SAGE Publications.
- Potter, J. and Edwards, D. (2001). Discursive Social Psychology. In W.P. Robinson and H. Giles (Eds.) *The New Handbook of Language and Social Psychology.* London: Wiley.
- Potter, J., Edwards, D. & Wetherell, M. (1993). A model of discourse in action. *American behavioral scientist, 36*, 383-401.
- Potter, J. & Wetherell, M. (1987). *Discourse and Social Psychology.* London: Sage Publications.
- Rand-Hendriksen, K. (2008). *Ideological changes measured through changes in language: Development, description and preliminary validation of a new archival*

- method.* (Unpublished master's thesis). Department of Psychology, University of Oslo. Retrieved May 28, 2010, from <http://www.duo.uio.no/sok/work.html?WORKID=72611>
- Rommetveit, R. (1968). *Words, Meanings and Messages*. New York: Academic Press
- Sampson, E. E. (1989). The challenge of social change for psychology: Globalization and psychology's theory of the person. *American psychologist*, *44*, 914-921.
- Scheff, T. (1990). *Microsociology*. Chicago: University of Chicago Press.
- Shaugnessy, J. J., Zechmeister, E. B & Zechmeister, J. S. (2009). *Research methods in psychology*. New York: McGraw-Hill.
- Shotter, J. (1992). Bakhtin and Billig: Monological versus dialogical practices. *American Behavioral Scientist*, *36*, 8-21.
- Sklair, L. (1999) Competing concepts of globalization. *Journal of world-systems research*, *2*, pp. 143-163.
- Sluga, H. (1999). Wittgenstein's Critique of Philosophy. In Sluga, H. & Stern, D. G. (Eds.), *The Cambridge Companion to Wittgenstein*. (pp. 34-58). Cambridge: Cambridge University Press.
- Steger, M. B. (2005a). Ideologies of globalization. *Journal of Political Ideologies*, *10* (1), pp. 11-30.
- Steger, M. B. (2005b). From market globalism to imperial globalism: Ideology and American power after 9/11. *Globalizations*, *2*, pp. 31-46.
- Stiglitz, J. E. (2002). *Globalization and its discontents*. New York: W.W.
- Švejnar, J., & Hvižd'ala, K. (2008). *Kam kráčíš, Česko?* [Where do you go, Czechia?] Praha: Rybka Publishers.
- Syverstad, C. (2009). *Fellesskap og individualisme i globaliseringens tidsalder*. (Unpublished master's thesis). Department of Psychology, University of Oslo. Retrieved May 28, 2010, from <http://www.duo.uio.no/sok/work.html?WORKID=92060>
- Turken, S. Blakar, R. M. & Nafstad, H. E. Article in preparation.
- Thompson, J. B. (1990). *Ideology and Modern Culture*. Oxford: Polity Press.
- Triandis, H. C. (1990). Cross-cultural studies of individualism and collectivism. In Berman, J. (Ed.) *Nebraska symposium on motivation*. (pp. 41-133). Lincoln: University of Nebraska Press.
- Triandis, H. C. (1995). *Individualism and collectivism*. Boulder, CO: Westview Press.

United States Declaration of Independence (1776), retrieved May 11, 2010, from http://en.wikisource.org/wiki/United_States_Declaration_of_Independence

Vetlesen, A. J. (2004). Det frie mennesket? Et sosialfilosofisk blikk på patologiene i opsjonssamfunnet. In Nafstad, H. E. (Ed.). *Det omsorgsfulle mennesket*. (pp. 17-54). Oslo: Gyldendal Akademisk.

Wierzbicka, A. (2009). What makes a good life? A cross-linguistic and cross-cultural perspective, *The Journal of Positive Psychology*, 4, pp. 260-272

Wittgenstein, L. (1921/2004) *Tractatus Logico-Philosophicus*. Retrieved May 5, 2009, from Project Gutenberg Web site: <http://www.gutenberg.org/etext/5740>

Wittgenstein, L. (1953/2001). *Philosophical Investigations*. (3rd ed.). Malden, MA: Blackwell Publishing.

Web addresses:

The following addresses can be used to access the electronic archives of the newspapers described in this thesis:

Ghanaweb: <http://www.ghanaweb.com/GhanaHomePage/search.php>

Iran: <http://www.iran-newspaper.com/> English version: <http://iran-daily.com/> (Unfortunately, the specific search function used to gather the Iranian dataset is no longer available. Searching is still possible, though cumbersome.)

Mladá fronta DNES online: <http://www.idnes.cz/> (this newspaper archive was searched through <http://mediasearch.newtonit.cz/>)

New York Times: <http://query.nytimes.com/search/alternate/query?query=&st=fromcse#to>

Hürriyet: <http://hurarsiv.hurriyet.com.tr/arsiv/>

Retriever (Norwegian newspapers): www.rand-hendriksen.net/atekst

Washington Post: <http://pqasb.pqarchiver.com/washingtonpost/search.html>