University of Groningen

Semantic text analysis and the measurement of ideological developments within fledgling democracies

Popping, Roel; Roberts, Carl W.

Published in:
Social Science Information

DOI:
[10.1177/0539018414553865](https://doi.org/10.1177/0539018414553865)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2015

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Popping, R., & Roberts, C. W. (2015). Semantic text analysis and the measurement of ideological developments within fledgling democracies. *Social Science Information*, 54(1), 23-37.
<https://doi.org/10.1177/0539018414553865>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Semantic text analysis and the measurement of ideological developments within fledgling democracies

Social Science Information

2015, Vol. 54(1) 23–37

© The Author(s) 2014

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/0539018414553865

ssi.sagepub.com **Roel Popping**

Department of Sociology, University of Groningen, The Netherlands

Carl W. Roberts

Department of Sociology and Department of Statistics, Iowa State University, IA, USA

Abstract

This methodological article presents an introduction to the field of clause-based semantic text analysis. The method is introduced and elaborated with regard to the study of ideological developments within fledgling democracies. In such studies modality plays an important role. Democratic societies are maintained in accordance with either a modality of achievement or one of necessity. This is illustrated using editorial texts from Hungary, one of the countries in Central and Eastern Europe where people had to find their democratic way after communism disappeared in 1989.

Keywords

content analysis, democratization, Hungary, ideological development, text analysis

Résumé

Cet article méthodologique présente une introduction au champ de l'analyse sémantique de texte basée sur les clauses. La méthode est présentée et développée dans le cadre d'une étude des développements idéologiques au sein de jeunes démocraties. Dans de telles études, la modalité joue un rôle important. Les sociétés se maintiennent en accord avec une modalité soit de réussite soit de nécessité. Ce qui est illustré sur la base de textes éditoriaux en Hongrie, l'un des pays d'Europe Centrale et de l'Est où les gens ont dû trouver leur propre chemin démocratique après la disparition du communisme en 1989.

Corresponding author:

Roel Popping, Department of Sociology, University of Groningen, Grote Rozenstraat 31, Groningen

NL 9712 TG, Netherlands.

Email: r.popping@rug.nl

Mots-clés

analyse de contenu, analyse de texte, démocratisation, développements idéologiques, Hongrie

Prior to the mid-1970s the dominant social scientific methodologies for the quantitative analysis of texts applied statistics to data comprised of word counts within categories (Berelson, 1952; Holsti, 1969; Krippendorff, 2004). These text analyses were considered a subset of the broader category of methodologies called ‘content analysis’, in which a range of qualitative data (e.g. words, gestures, art forms, etc.) might be quantified prior to statistical analysis. Yet during the last three decades alternative text analysis methodologies have been developed for affording statistical inferences about populations of texts.

Unlike earlier methods, more recent approaches to text analysis involve not only the identification of concepts, but also the encoding of relations among these concepts. There are striking similarities among these researchers’ methods for encoding texts. In particular, they almost universally encode Subject–Verb–Valence–Object (S–V–V–O) relations among words within clauses of the texts they analyze.

The primary distinction among these relational approaches to analyzing texts is between those based on semantic grammars (semantic text analysis) versus those based on networks (network text analysis). Network text analysis has been used to compare mental maps (Carley, 1986, 1988) or to evaluate performative networks among political policy makers (Van Cuilenburg et al., 1988). This article introduces a semantic text analytical approach to the study of political discourse within the democratizing societies of Central and Eastern Europe. We show that, in the political discourse of newspaper editorials, social-democrats tend to use the modal form ‘inevitable’, whereas liberals tend to use the form ‘possible’. This line of reasoning can be used if one wants to investigate the direction in which a new democracy is developing. Accordingly, we cease to consider network analysis at this point. The following section provides a brief introduction to semantic text analysis. We next describe our research project and the applicability of semantic text analysis to it.

Varieties of semantic grammars

Whereas thematic text analysis concentrates on word or phrase counts, a semantic text analysis examines theme relations. An example of thematic analysis is found in Laver and Garry’s (2000) investigation of differences between political parties. They analyzed party manifestos within which ‘desirable state policies’ are portrayed. In their study the investigators classify words and phrases used in the manifestos into numerous complex categories of state policies. For example, their category labeled *Economy/+State+/-Budget/Spending/Health* is applied to statements in which a desire is indicated for an increase of public spending in the field of health. Alternative categories were developed for reduced spending on health, increased spending on housing, younger eligibility for social security, etc. To cover the whole field of issues that play a role in the left–right classification, the investigators needed a large number of complex categories.

In a semantic text analysis, aspects of categories (e.g. state policies) are differentiated in accordance with the fields of a semantic grammar. For example, in their study of the *cahiers de doléances* of 1789, Shapiro and Markoff (1998) classified grievances in accordance with a Verb–Object semantic grammar within which members of a wide range of French organizations expressed their appeals that King Louis XVI might take a variety of actions (e.g. abolish, improve, modify, re-establish, etc.) regarding a variety of objects (e.g. the government, religion, the economy, etc.). Thus, instead of developing categories that comprise all the possible characteristics of one's phenomena of interest (here, public grievances just prior to the French Revolution), one links categories in accordance with fixed semantic relations (those of verb to object, for instance). Koopmans and Erbe (2004) have done comparable research, only this time focusing on claims addressed to institutions or entire states within the European Union.

Most applications of semantic text analysis involve coding clauses according to the S–V–O triplet. Generally the unit encoded in accordance with such a semantic grammar is the grammatical clause (i.e. the sentence or part of a sentence that explicitly or implicitly contains an inflected verb, an optional subject and/or object, plus all modifiers related to this verb, subject, and object). Yet such a semantic grammar might easily be expanded to include four syntactic components:

- Agent: the initiator of an activity
- Position: the position regarding the agent's activity
- Action: the activity under consideration
- Object: the target of this activity

Because positions can only be taken by intentional agents, the agent cannot be any arbitrary subject. Instead it can only be a person or institution (comprising persons) and not a 'thing'. The position commonly involves the use of a modal auxiliary verb. For example, an editorial that states 'Our politicians ought to cooperate more', would be encoded as politicians (agent) ought (position) to cooperate (action) with politicians (object). Other examples are Poles (agency) can (position) improve (action) their standard of living (object), and gasoline taxes (object) must (position) be lowered (action). (Note that in passive voice this last sentence specifies no agency, because its author did not explicitly implicate the government as the raiser of taxes.)

Table 1 illustrates the type of data matrix produced when texts are semantically encoded. Note that the numbers in the cells of this matrix correspond to specific concepts. For example, 'Politicians ought to cooperate more' could be encoded under the first ID using 12 as the code for 'politician', 7 for 'ought', and 34 for 'cooperate'. Such semantically encoded data can be used to make inferences about the conditions under which texts' authors take specific positions on others' intentions. Accordingly, if a government's totalitarian leadership were to initiate a public-relations campaign to 'democratize' its image in the US press, systematically encoded data could be used to test whether a significant increase had occurred from before to after the campaign in US news stories' depictions of the regime (agency) attempting (position) to heed (action) its citizens (object).

Table 1. A data matrix for a semantic text analysis.

ID-number	Agent	Position	Action	Object
1	12	07	34	12
2	08	03	22	88
3	14	03	34	70
4	10	01	35	70
5	16	11	29	52

The move from a thematic to a semantic text analysis expands the types of questions a researcher can answer. Referring to propaganda techniques in making this point, Roberts (1989: 169) notes that, in a thematic analysis, a possible research question would be: ‘What themes are mentioned in propaganda that are not mentioned in other communications?’ Using the semantic approach the question can be extended to: ‘What syntactic strategies are used by political leaders when their policies fail (or succeed)?’ Unlike the former question, the latter asks about concrete relations among concepts used within different social contexts.

Modality analysis

Generally speaking, modal auxiliary verbs are verbs that are used with (usually, the infinitive form of) another verb to express possibility, inevitability, impossibility, or contingency. In each modal usage there are two verbs associated with the subject, namely a modal auxiliary verb (e.g. want, hope, ought, refuse) and a main verb in infinitive form (typically an action). These usages are not intended to convey facts or to describe events; they are used to communicate something about the likelihood of the agent–action–object link. A semantic text analysis that investigates such uses of modal auxiliary verbs is called a ‘modality analysis’ (Roberts et al., 2010).

A modal clause is recognizable whenever it conveys intentionality in a way that can be transformed (in a manner agreeable to a native speaker) to a form that includes a modal auxiliary verb. For example, English speakers frequently use ‘have to’ instead of ‘must’ when indicating compulsion despite one’s intentions. On the other hand, modal auxiliary verbs are sometimes used in ways entirely unrelated to modality. Probably the best example here is the tendency among English speakers to convey future possibility rather than permission in their usage of the modal ‘may’. In modality analysis, the coder’s challenge is to capture how a text’s author understands others’ motivations (thereby getting into the mind of someone who is getting into someone else’s mind, as it were).¹ Because modal auxiliary verbs convey intentionality, they can be used to learn about people’s motivations, their characterizations of each other, and thus about ideological shifts between and within their societies.

The semantic grammar used in a modality analysis always has two parts at its core. There is a modal form (indicating possibility, impossibility, inevitability, or contingency) and an associated rationale (e.g. expressed in liberal or social-democratic terms). Let us begin with a discussion of the four modal forms.

The modal form of ‘possible’ conveys to the reader that the modal’s predicate is an option for the modal’s subject. The modal form of ‘impossible’ conveys that the modal’s predicate is not an option for the modal’s subject. The form of ‘inevitable’ conveys that the subject–predicate link is imminent. Finally, the modal form of ‘contingent’ conveys the non-imminence of the subject–predicate link. These modals can be used to understand human motives during interactions, and to distinguish subtle nuances in discourse. For example, a sentence like ‘I can do it’ is coded as ‘possible’, implying capability (and, most likely, a sense of self-confidence); and a sentence like ‘I am not able to do it’ is coded as ‘impossible’, indicating incapability (and a potential sense of futility). A sentence like ‘she must be a good person’ is coded as ‘inevitable’, expressing certainty; and a sentence like ‘she is able not to do it’ is coded as ‘contingent’, indicating alternatives (and potential avoidance). Double negative expressions such as ‘not can not’ and ‘not must not’ are extremely rare in vernacular text.

Illustrations of the four modal forms are as follows:

Possible	a	you are <i>able</i> to
	b	you are <i>not required not</i> to
	c	you are <i>permitted</i> to
	d	you are <i>not obliged not</i> to
Impossible	a	you are <i>not able</i> to
	b	you are <i>required not</i> to
	c	you are <i>not permitted</i> to
	d	you <i>should not</i>
Inevitable	a	you are <i>not able not</i> to
	b	you are <i>required</i> to
	c	you are <i>not permitted not</i> to
	d	you <i>should</i>
Contingent	a	you are <i>able not</i> to
	b	you are <i>not required</i> to
	c	you are <i>permitted not</i> to
	d	you are <i>not obligated</i> to

Through modal usage a text’s source (i.e. its author or speaker) socially constructs what constitutes the possible, the impossible, the inevitable, and the contingent regarding the agent–action–object link. In addition, it is always reasonable for the source of a modal clause to be queried as to the rationale (or explanation of ‘why’) the agent is able, required, permitted, etc., regarding the clause’s predicate. For example, a politician might follow his statement, ‘We had to impose austere economic measures’, with the rationale, ‘otherwise our economy would have stagnated’.²

Directions of democracy

Archival documents from the early 1990s afford a rich source of information on how Central and Eastern European nations transformed from totalitarian societies into democracies. In their texts on how democracy might develop in Central and Eastern

Europe, both Ost (1993) and Comisso (1997) consider three socio-political orientations, namely (neo-)liberal, social-democratic, and populist/nationalist. Whereas Ost's typology is predictive and primarily based on economic factors, Comisso's emphasizes cultural-ideological factors. Nonetheless, the two typologies are in many ways mutually supplemental (see Touraine, 1991, for a more nuanced discussion of various definitions of democracy somewhat along the lines presented here).

By liberalism or neo-liberalism, both Ost and Comisso refer to the European sense of the term, indicating trust in the marketplace rather than in one's government. The configuration of the market economy should take place as soon as possible. The once-planned economy is to be privatized. The establishment of a good foundation for investments is needed. As a special form of the liberal type, capitalism was the form most citizens desired immediately after 1989.

Comisso adds political ideas in her typology of the liberals. According to the liberals, the protection and augmentation of individual freedom is the fundament of a democracy. Liberals in Central-East European countries wished for transformations (especially capitalist reforms) to be implemented much faster than may have been realistic. In their view the changes proceeded too slowly.

Berlin's (1969) concepts of negative and positive liberty help clarify the dual nature of liberalism's emphasis on individual freedom. Negative liberty denotes the absence of obstacles, barriers, or constraints. Positive liberty refers to the possibility (or fact) of acting in such a way as to take control of one's life and realize one's fundamental purposes. Thus positive liberty requires negative liberty. In liberal terms, if one favors (positive) individual liberty one should place strong limitations on the (liberty-negative) activities of the state. Counter-arguments to Berlin's view come from a broad range of non-liberal orientations. In brief, these critics argue that state intervention is often required for the successful pursuit of liberty, generally understood as individual or collective self-realization/self-determination.

For citizens with a populist/nationalist orientation changes occur too fast, and they may react by advocating an exchange of democracy for a strong, centralized government that will lessen the speed and thus the pain of reform. Such public reactions are often embedded in appeals to national pride, to ethnic rights, to xenophobic concerns, etc. Mann (2005) argues that widespread expressions of this sort are likely to foreshadow atrocities such as ethnic cleansing as the will of an ethnic majority takes precedence over the democratic process itself.

The social-democratic type shares with the liberals the idea that large-scale industry is of highest importance for the economy (Ost, 1993). However, social costs should remain minimal, making their position not only an economic one but a moral one as well. According to the social-democrats, the workers, who paid the costs for so long and who enabled a revolution, should have the opportunity to benefit now from the gains. They should not once again be sacrificed to the interests of future generations.

This approach strongly resembles Comisso's second, egalitarian type. She also sees a resemblance between liberal and social-democratic types, both of which start from the individual. However, the difference with the liberal type is that the egalitarians do not endorse maximizing individual freedom. The egalitarian prefers an equal vote for

everybody in the political process, even it entails state interference in the economy. This is contrary to the liberal democratic view, according to which the free-market sector should be unconstrained. For both Comisso and Ost, social-democratic and liberal orientations have reverse loci of trust, respectively with government and the market. Only with the former is a powerful government legitimate in ensuring equitable (re)distribution of the nation's resources among its citizens.

Rather than conceiving of liberal, populist, social-democratic, and totalitarian as discrete orientations, Potůček (1999: 65) places liberal and totalitarian doctrines (communism, fascism, and so on) at opposite ends of a continuum. He differentiates points along this continuum in terms of support for values that would be expected of persons located at various points between the extremes. Potůček's argument is that people's value orientations along this continuum are important, because they determine possible 'choices of society'. However, at the right end of his continuum, an awkwardly unexplained 'leap' exists from social-democrat to totalitarian – a leap that need not be made in studies of discourse within democratizing (i.e. no longer totalitarian) societies.

Within democratic (or democratizing) national contexts, a social-democratic versus liberal dialectic (with populism comprising an ever-present subversive avenue from democracy to totalitarianism) may afford more analytic purchase. Based on a review of the democratization literature, Roberts et al. (2009) note that it is typically via an alliance between the masses and the bourgeoisie that authoritarian regimes are ousted from power. Subsequent democratic efforts generally consist of struggles between parties that represent the respectively social-justice (or social-democratic) and market-justice (or liberal) interests of these two previously allied groups. Their subsequent modality analysis of post-1989 Hungarian editorials is based on the premise that legislative decisions generally call for choices between the policies advocated by representatives from each of these camps, because implementation of one of the former allies' interests will commonly prove counter to the other ally's interests:

- By redistributing wealth to meet constituents' needs, the government inevitably does so by limiting reinvestment in the marketplace.
- By allowing complete discretion over constituents' profits, the government will have no resources with which to assist its neediest constituents.³

And so, we now arrive at the familiar left–right political divide that characterizes party divisions within most contemporary democracies. On the left are those who justify government activities in terms of how well they manage a social system within which constituents' hardships are minimized; on the right are those who justify such activities in terms of how well they ensure a stable economic field within which constituents can accumulate wealth. Beyond a consensus on the general benefits of a healthy national economy, at issue is whether the state's role is to minimize its citizens' risks or to maximize their opportunities. We now turn to a consideration of how modality analysis might be used to capture appeals to these two distinct ways of justifying political action.

Data

The data for this study are taken from a representative sample of editorials that appeared in the Hungarian newspaper, *Népszabadság*, between August 1990 (the month when Árpád Göncz became Hungary's first post-communism elected president) and June 1997.⁴ *Népszabadság* was selected because this newspaper is the least biased, yet consistent public forum for political and social discourse available on Hungary during the period of study. Our sample, consisting of 470 combinations of modal form and rationale, is described in detail in Roberts et al. (2009).

From this data set, 47 combinations were selected that illustrate liberal or social-democratic views in line with positions taken by Ost and Comisso. Two independently operating coders selected these combinations, taking into account the editorial context within which each was written. Details on the coding instructions and related issues are presented in Popping and Roberts (2009). Based on the 47 combinations that were assigned by at least one of the coders to either the liberal or social-democratic orientation, intercoder agreement following Scott (1955; Popping, 2010) is .88 ($z = 31.16$, $p < .001$). This subset is less than ideal, given that the texts were not sampled for the purpose of this study. Nonetheless, the subset is useful for illustrating how liberal and social-democratic rhetoric can be differentiated.

Measuring democracy-related ideological statements

Each modal statement-plus-rationale can in its most extended form be rendered according to the following template (or semantic grammar):

There is an R (rationale) why Z (action-object) is {possible / impossible / inevitable / contingent} for Y (agent) according to X (source).

Now the task is to identify R, X, Y, and Z. Here X is the author or speaker (presumably an adherent to a certain ideology), Y is the agent (to whom X believes the ideology applies), and Z is the predicate (the activity X is socially constructing as attributable to Y). Rationales (R) account for predicates' possibility, impossibility, etc., in terms of political, economic, cultural, etc., ideological factors.

The speaker, X, and agent, Y, can be coded as liberal, socialist, and so on. Predicates, Z, can be coded in various degrees of detail. Yet it is the rationales that link the former three in accordance with ideological justifications. Roberts et al. (2009) distinguish among rationales' political, economic, cultural, and security domains (see also Ben-Sira, 1977, 1979; Semetko & Valkenburg, 2000). They also argue that social-democratic orientations indicate necessity (e.g. what must responsibly be done). That is, predicates (i.e. expected or planned courses of events) are depicted as community members' (*inevitable*) responsibilities. Accordingly, an illustration of a social-democratic orientation is as follows:

The government (lit: state) *must* support the weak [the poor people]. (10 September 1996)⁵

Coded, it reads as: ‘The government (agent) must (position = inevitable) support (action) the weak (object).’ Supporting the poor is inevitable because of community members’ responsibility (here, via the government).

Liberal orientations typically convey possibility in the form of ability (e.g. of opportunities that can be pursued). Here a statement’s predicate is depicted as possible. If a predicate is *impossible*, it is because of the individual’s limitations. An illustration of a liberal view is the following:

Strong, young guys *could* work as porters (exclusively for tips only), they can, however, keep their numbered hats as souvenirs. (3 April 1991)

The first clause in the sentence is read as: ‘Strong, young guys (agent) could (position = possible) work as porters (action).’ The predicate in this illustration is depicted as possible. The rationale refers to the subject’s bodily resources.

Beyond characteristically social-democratic mentions of inevitability and liberal mentions of possibility, all other modal forms can be used in conveying social-democratic and liberal ideological orientations within public discourse. That is, we are not assuming that social-democratic modal rhetoric will always convey ‘inevitability’ and that liberal modal rhetoric will always convey ‘possibility’. Following Roberts (2008), all four modal forms are used in accordance with each political orientation. (See Roberts & Wang, 2010, for further illustrations along these lines in a comparative analysis of US and Japanese editorials or Roberts & Liu, 2014, for US and Canadian editorials.)

Social-democratic modal forms

We first consider the four social-democratic modal forms.

(1a) *must* – indicates what the subject is responsible to do. The meaning conveyed is that of meeting citizens’ needs. This type, illustrated in a previous example, is illustrated again in the following:

We *need to* declare that welfare and cultural assets of the companies are common assets of the companies’ employees and all waged/salaried employees, and companies are only managers of those assets. (18 February 1991)

To be encoded as: ‘We (agent = Hungarians) need to (position = inevitable) declare (action) common ownership of companies’ welfare and cultural assets (object).’ Here again a specific predicate is inevitable because of community members’ responsibilities. This form with ‘must’ is the most frequent form among modal statements expressing a social-democratic view.

(1b) *not must* – indicates what the subject is not [to do]. Instances may convey that an agent has ‘done enough’ or is her- or himself needy. Although in the following sentence

(a conditional variant of) ‘can’ instead of ‘must’ is used in conveying contingency, a social-democratic instance of this form is illustrated here:

[Financial] experts have recommended for a long time that a national fund [similar to the National Deposit Security Fund] is needed to protect small investors, so small investors *could not* lose their wealth, which was accumulated with years of work, because of a bad decision in a few minutes. (16 February 1996)

Encoded as: ‘Small investors (agent) are able not to (position = contingent) lose (action) their wealth (object).’ The predicate here is one of contingency because small investors should not be subject (as the author subsequently argues) to risks of which they are unaware. Since the argument is for minimizing citizens’ risks and not for submission to market forces, this is a social-democratic modal statement. Such expressions of contingency are exceedingly uncommon among all types of modal statements.

(1c) *must not* – indicates what the agent is responsible not to do. The meaning conveyed here is that the subject would be unqualified or reckless to proceed. In the social-democratic texts one example of ‘must not’ was found. However, since this example does not refer to the situation in Hungary, it (like the previous illustration) was not included in the Roberts et al. (2009) analysis and thus is in addition to the 47 combinations in our primary data set.

In Holland, they have introduced a law, according to which the maximum medicine prices *must not* be higher than the average price [lit.: value] in the neighboring countries, so the average price in another [country] may be the highest price in Holland. (27 January 1997)

This is encoded as: ‘The Dutch (agent) must not (position = impossible) sell (action) medicine that is more expensive than the average price in other countries (object).’

(1d) *not must not* – indicates that the agent is being irresponsible or lazy. The intended meaning is one of unmet citizens’ needs. (Consider, for example, the statements, ‘It’s untrue that you must no longer continue. Get back to work!’) Although no specific instance of ‘not must not’ occurs in the texts, its logical equivalent (namely, ‘can’) does occur within social-democratic endorsements of possible actions to be taken as alternatives to neglecting one’s social responsibilities.

Of course, it would not only be unjust but also a political and moral mistake, for anybody to blame the current government for unemployment. However, after three years we *can* ask [the government] about the size of unemployment and the weakness of policies developed for its ‘handling’. (9 July 1993)

The encoding here would be, ‘We (agent = Hungarians) are not compelled not (position = possible) to ask (action) the government about its policies (object). Although one must not blame the government, it is untrue that one must not hold the government to account.’

Liberal modal forms

We also distinguish four types of modal forms for the liberal view.

(2a) *can* – indicates an opportunity available to the agent. In addition to our previous illustration, a second illustration of this liberal view (and of negative liberty) is as follows:

Austria, and many international firms see this [expo] as a huge business opportunity, and we *can* share the success and the profits too. (10 December 1990)

Rendering: ‘We Hungarians (agent) *can* (position = possible) share (action) the success and profits (object).’

(2b) *not can* – indicates a handicap for the subject. The intention of such statements is to convey the agent’s lack of genetic or material resources. This is illustrated by:

The voters *cannot* recognize yet, that the freedom of press is their freedom as well. (14 March 1994)

This is encoded as: ‘The voters (agent) *not can* (position = impossible) recognize (action) that the freedom of press is their freedom as well (object).’ Implied here is that the agent needs time to learn this truth about press freedom.

(2c) *able not* – indicates contingency or subject discretion. The rationale is the agent’s availability of alternative opportunities. (For example, ‘I am capable of not doing it because I have other alternatives.’) No example of this type was found in the texts.

(2d) *not able not* – indicates agent coercion. The implication here is that the rules of the marketplace (and the enforcement of these rules) are inevitable constraints for all agents. Note the equivalence in modal logic between ‘must’ and ‘not able not’.

The [agricultural] sector increasingly needs to turn toward foreign markets due to the shrinking domestic market; and in this process we *must* fight with international protectionism. (10 August 1990)

Encoding would be as follows: ‘Hungarians (agent) are *not able not* (position = inevitable) to fight against (action) international protectionism (object).’ Since protectionism distorts the marketplace, we must fight against it.

Unanticipated combinations of modal form and rationale

In analyzing modal use in other studies of such editorials, we have repeatedly found that references to possibility and inevitability are much more frequent than references to impossibility or contingency. The findings in Table 2 are no exception to this. The table contains the combinations of modal form and political orientation as found in the Hungarian data. For these data $\chi^2 = 23.86$ (d.f. 6; $p < .001$), so the distribution of the frequencies over the cells given the marginals is far from random. Roberts et al. (2009) suggest that the modal form ‘possible’ is most prevalent among liberal modal statements and the ‘inevitable’ form is most frequent among social-democratic ones. As indicated in

Table 2. Modal forms identified as liberal and social-democratic.

	Liberal	Social-democrat	Neither	Total
Possible	14	5	234	253
Impossible	3	7	112	122
Inevitable	4	13	69	86
Contingent	0	1	8	9
Total	21	26	423	470

illustrations within the previous two subsections, one does find liberal statements indicating inevitability and social-democratic statements indicating possibility in Hungary’s post-1989 public discourse. However, they constitute a minority of the social-democratic and liberal modal statements in this discourse.

Let us consider a few cases inconsistent with our expectations, most notably modal statements in which economic reasons are given for inevitability and when welfare-related reasons are given for impossibility:

If we want to harvest to obtain a bigger share from the goods of the world economy, then we also *have to* do something. (10 December 1990)

This modal statement is from an editorial written by a trained economist, who follows a liberal line of thought, and therefore was jointly identified by our coders as expressing a liberal view. However, note that when its context is not taken into account, the statement might also have been made by a social-democrat (i.e. if the ‘harvesting’ were for our common wellbeing rather than for improved opportunities to individual entrepreneurs). Such ambiguity is greatest in situations where the ‘opposite’ of our theoretical expectations is found.

The modal form ‘impossible’ is used in combination with statements from either a liberal or a social-democratic perspective. These statements typically have the modal form ‘not can’. For the liberals, impossibility typically conveys the subject’s personal limitations (see illustration 2b above). The following typifies one of seven social-democratic statements of impossibility:

[The government] *cannot* leave local governments alone, the local authorities are unable to solve these problems due to the lack of money [lit.: materials]. (23 November 1992)

This statement might be considered as another way of expressing a parallel social-democratic modal statement that we would expect to find:

[The government] *is obliged* to support local governments, the local authorities are unable to solve these problems due to the lack of money [lit.: materials].

The reasoning is that ‘leaving local governments alone’ is logically equivalent to ‘not getting involved with (supporting?) local governments’, and that ‘not able (i.e. cannot)’ is logically equivalent to ‘must not’ – each conveying impossibility. Taking these together

we get the former statement transformed into ‘The government must not not get involved with local governments’, which is equivalent to ‘The government must (or “is obliged to”) get involved with local governments.’ This reasoning is based on the rules of modal and first-order logics.

This transformation is applicable to all seven modal statements of impossibility from Table 2 that have a social-democratic orientation, justifying our moving them from the row for ‘impossible’ to the row for the modal form of ‘inevitable’. The resulting χ^2 for the data in the revised table now becomes 60.94 (d.f. 6, $p < .001$), thereby strengthening our argument that social-democratic public discourse is characterized by references to modal statements intended to convey inevitability.

The data used in the study on Hungarian public discourse were not collected with the goal of finding modal statements expressing liberal or social-democratic orientations. For that reason only 10% of all modal statements in the data set are useful for the present study. But if data had been collected with the goal of sampling modal rhetoric characteristic of these two orientations, we hypothesize that a pattern would have been found similar to that in (the revised) Table 2.

Conclusion

This article builds on a technique of semantic text analysis using modal forms. It is argued that in political discourse social-democrats tend to use the modal form ‘inevitable’, whereas liberals tend to use the form ‘possible’. It is shown that statements having a social-democratic rationale indeed generally contain the modal form ‘inevitable’ and that statements having a liberal rationale generally use the modal form ‘possible’.

So in studies of ideological developments in the direction of democracy, the modal form used within the nation’s editorials may be useful in predicting whether the direction of democratization is toward liberalism or social-democracy. Accordingly, the type of modality analysis performed here may be usefully applied to editorials and other archival texts from democratizing countries within which it is not completely clear what the country’s ideological direction might be. It affords an alternative measure of popular political perspectives that may not be in line with voting behaviors – a measure less susceptible to short-term interpersonal and historical influences.

However, this is only a beginning. To draw conclusions about the character of political rhetoric, a larger, more representative sample needs to be used, necessarily including multiple newspapers from more than one new democracy. It should also be clear that the press systems in fledgling democracies are the voices of political positions.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Notes

1. Popping (2001) provided a first attempt at classifying actions based on the use of modal auxiliary verbs by differentiating among ego’s potentiality, ego’s needs, alter’s morality, and alter’s

obligation – classifications indicative of what the speaker says reflexively or about someone else regarding what might or must be done.

2. For more background on finding the modal forms, see Roberts et al. (2010).
3. In a cross-examination using survey data, results in line with the findings in this study were found (Popping, 2013).
4. The editorials have been translated into English.
5. Each of this paper's editorial citations is from the Hungarian newspaper, *Népszabadság*.

References

- Ben-Sira Z (1977) A facet-theoretical approach to voting behavior. *Quality and Quantity* 11(2): 167–188.
- Ben-Sira Z (1979) Toward a facet theory of sequential order of societal needs. *Quality and Quantity* 13(3): 233–253.
- Berelson B (1952) *Content Analysis in Communication Research*. New York: The Free Press.
- Berlin I (1969) Two concepts of liberty. In: Berlin I, *Four Essays on Liberty*. London: Oxford University Press.
- Carley K (1986) An approach for relating social structure to cognitive structure. *Journal of Mathematical Sociology* 12(2): 137–189.
- Carley K (1988) Formalizing the expert's knowledge. *Sociological Methods and Research* 17(2): 165–232.
- Comisso E (1997) Is the glass half full or half empty? Reflections on five years of competitive politics in Eastern Europe. *Communist and Post-Communist Studies* 30(1): 1–21.
- Holsti OR (1969) *Content Analysis for the Social Sciences and Humanities*. London: Addison-Wesley.
- Koopmans R, Erbe J (2004) Towards a European public sphere? Vertical and horizontal dimensions of Europeanized political communication. *Innovation* 17(2): 97–118.
- Krippendorff K (2004) *Content Analysis: An introduction to its methodology*. Thousand Oaks, CA: Sage.
- Laver M, Garry J (2000) Estimating policy positions from political texts. *American Journal of Political Sciences* 44(3): 619–634.
- Mann M (2005) *The Dark Side of Democracy: Explaining ethnic cleansing*. Cambridge: Cambridge University Press.
- Ost D (1993) The politics of interest in post-communist East Europe. *Theory and Society* 22(4): 453–485.
- Popping R (2001) Modal auxiliaries in text analysis. *Bulletin de Méthodologie Sociologique* 70: 5–22.
- Popping R (2010) Ag09. A computer program for interrater agreement for judgments. *Social Science Computer Review* 28(3): 391–396.
- Popping R (2013) Democratic transformation in Hungary 1990–1997: A cross-examination. *Quality and Quantity* 47(5): 2415–2424.
- Popping R, Roberts CW (2009) Coding issues in modality analysis. *Field Methods* 21(3): 244–264.
- Potůček M (1999) *Not Only the Market: The role of the market, government and civic sector in the development of postcommunist societies*. Budapest: Central European University Press.
- Roberts CW (1989) Other than counting words: A linguistic approach to content analysis. *Social Forces* 68(1): 147–177.
- Roberts CW (2008) *'The' Fifth Modality: On languages that shape our motivations and cultures*. Leiden: Brill.
- Roberts CW, Liu H (2014) On the cultural foundations for universal healthcare: Implications from late 20th century US and Canadian health-related discourse. *Journal of Communication* 64(4): 764–784.

- Roberts CW, Wang Y (2010) Traitor in our midst: Cultural variations in Japanese vs. Oklahoman public discourse on domestic terrorism in the spring of 1995. *Comparative Sociology* 9(4): 463–494.
- Roberts CW, Popping R, Pan Y (2009) Modalities of democratic transformation: Forms of public discourse within Hungary's largest newspaper, 1990–1997. *International Sociology* 24(4): 498–525.
- Roberts CW, Zuell C, Landmann J, Wang Y (2010) Modality analysis: A semantic grammar for imputations of intentionality in texts. *Quality and Quantity* 44(2): 239–257.
- Scott WA (1955) Reliability of content analysis: The case of nominal scale coding. *Public Opinion Quarterly* 19(3): 321–325.
- Semetko H, Valkenburg PM (2000) Framing European politics: A content analysis of press and television news. *Journal of Communication* 50(2): 93–109.
- Shapiro G, Markoff J (1998). *Revolutionary Demands: A content analysis of the cahiers de doléances of 1789*. Redwood City, CA: Stanford University Press.
- Touraine A (1991) What does democracy mean today? *International Social Science Journal* 128: 259–268.
- Van Cuilenburg JJ, Kleinnijenhuis J, De Ridder JA (1988) Artificial intelligence and content analysis. *Quality and Quantity* 22(1): 65–97.

Author biographies

Roel Popping works on historical shifts in public opinion, values, and scientific knowledge, primarily within the context of post-1989 Central and Eastern Europe; and in decision-making in negotiations processes. His book, *Computer-assisted Text Analysis*, was published by Sage in 2000.

Carl W. Roberts's current research on theory and intercultural comparisons is motivated by an interest in how language shapes perception and behavior. He has written extensively on linguistic structure in articles appearing in *Social Forces*, *Sociological Methodology*, *Journal of Communication*, as well as in his edited collection, *Text Analysis for the Social Sciences* (1997). Beyond his work with Roel Popping on political discourse within the newly democratizing countries of Eastern Europe, he is investigating cultural variations evident in sampled texts ranging from medieval documents to contemporary foreign and domestic newspapers.