Coding Issues in Modality Analysis

Popping, Roel; Roberts, Carl W.

Published in:
Field Methods

DOI:
10.1177/1525822X09333433

Publication date:
2009

Coding Issues in Modality Analysis

Roel Popping and Carl W. Roberts

Field Methods 2009 21: 244 originally published online 6 April 2009
DOI: 10.1177/1525822X09333433

Coding Issues in Modality Analysis

ROEL POPPING
University of Groningen
CARL W. ROBERTS
Iowa State University

Ambiguous phrases are the bane of researchers' attempts to ensure acceptably high interrater agreement in the encoding of texts. When modal usages and their associated rationales are encoded as part of a text analysis, ambiguities arise in characteristic (and thus identifiable) ways. This article illustrates the typical sources of disagreement among coders involved in encoding data during a modality analysis and provides concrete strategies for improving interrater agreement.

Keywords: *agreement; content analysis; text analysis; modality analysis*

INTRODUCTION

For more than 60 years, the dominant social scientific methodologies for the quantitative analysis of texts have applied statistics to data consisting of word counts within thematic categories (Berelson 1952; Holsti 1969; Krippendorff 2004). These text analyses are part of a broader class of methodologies called "content analysis"—a generic term for statistical analyses of qualitative data (e.g., words, gestures, art forms, etc.). During the last three decades, a variety of alternative text analysis methodologies have been developed for affording statistical inferences about populations of texts. These more recent approaches involve not only the identification of concepts but also the encoding of relations among them. For example, researchers may encode Subject-Verb-Valence-Object (S–V–V–O) relations among words within clauses of the texts they analyze. Such "ordered tuples" are referred to as instances of a semantic grammar when each is a distinct unit of analysis having a fixed set of semantic relations among its ordered parts (Markoff, Shapiro, and Weitman 1974; Roberts 1989, 1997; Shapiro and Markoff 1998).

Field Methods, Vol. 21, No. 3, August 2009 244–264
DOI: 10.1177/1525822X09333433

Modality analysis (Roberts et al. 2008) is a semantic text analysis methodology for investigating how peoples' intentions are discursively constructed. It is based on the theoretical premise that people motivate themselves and others by persuading them that their actions are possible, impossible, inevitable, or contingent. This persuasion is accomplished through the use of modal auxiliary verbs (Roberts 2008). For example, in light of a statement that one is "not able" or "compelled not" to go, one would stay (i.e., not go) if the statement persuaded one that "going is impossible." Other modal forms convey contingency (noncompulsion or ability not to), inevitability (compulsion or nonability not to), and possibility (ability or noncompulsion not to). The technique may be applied in any study of rhetorical references to peoples' intentions, whether made in the mass media or in small-group interactions.

Beyond conveying one of these four modal forms, every modal statement is (at least in principle) linked to a rationale. That is, assertions of possibility, impossibility, and so on are always open to question (e.g., "Why must I stay?"). Moreover, our experience has been that when no rationale for a modal statement is given, its use is more ritualistic than substantive: "May you live to see your children's children" or "You, too, can be president." Thus, we recommend that every modality analysis employ a coding scheme to classify the various rationales that might be used in accounting for the modal statements in one's text population.

To date, modality analyses have been applied solely to populations of newspaper editorials and/or letters to the editor. When sampled from a newspaper with broad national readership, samples of these texts may (arguably) be generalized to corresponding populations of public discourse at a specific time or on a particular topic. For example, Roberts, Popping, and Pan (forthcoming) apply a modality analysis to a sample of Hungary's largest newspaper's editorials published between 1990 and 1997—a paper in which evidence is provided for a post-1989 shift in Hungarians' politically based (or rationalized) rhetoric from more individualist references to Hungarians' newfound opportunities (or possibilities) to more mutualist references to their responsibilities (or necessities). (Beyond consulting the original study for more detail on the study's objectives and methodology, the reader is directed to Roberts et al. [2008] for an extensive discussion of modality analysis and its application.) In this article, we focus on issues of interrater agreement in coders' assignments of modal forms and rationales to words and phrases within these editorials. Our objective is to better acquaint investigators who want to use this method with decisions required during the coding process.

IDENTIFYING MODAL USAGES AND RATIONALES

Each time a modal auxiliary verb is used, there are two verbs associated with the verb's subject, namely the modal auxiliary verb (e.g., "can," "must," "ought," "refuse") and a main verb in infinitive form. These usages are not intended to convey facts or to describe events but to communicate something about the likelihood of the subject-verb-object link. As illustrated above, the fourfold character of these likelihoods results as the modal auxiliary verb, the main verb, both, or neither is negated—a pattern referred to by modal logicians as the "Square of Oppositions" (Horn 1989; Van der Auwera 1996).

Encoding of data for a modality analysis requires interrater agreement at each of the following four steps:

- Identification of a modal clause
- Identification of the modal clause's rationale
- Classification of the modal clause's form as possible, impossible, inevitable, or contingent
- Classification of the rationale according to a fixed coding schema

In this section, we focus on the first two of these steps by first explaining our criteria for identifying modal clauses/rationales and then describing the disagreements that arose in our coders' applications of these criteria.

Identification Criteria

Differences in coders' identifications of modal clauses may result from the character of the language at hand. For example, English speakers frequently use "have to" instead of "must" when indicating compulsion despite one's intentions, and they commonly use "may" to convey future possibility instead of permission. Moreover, modal auxiliary verbs are sometimes used in *descriptive or diagnostic* ways entirely unrelated to modality. Thus, neither of the following excerpts contains a modal clause:

> . . . anybody who has just a little sense of technology knows that the old, Rotary-type phone switchboards have moving parts too. If these parts wear down or work inaccurately then we *can* hear the background noises. (May 26, 1993, p. 10; note: This and all subsequent editorial citations are from the Hungarian newspaper, *Népszabadság*.)

> An ethical point of view is usually justified in case of all human actions, because its subjects, humans, *cannot* be separated from their actions. (April 1, 1994, p. 14)

The first quotation merely describes the physical causes of one's hearing of background noises; the latter provides the author's expert "diagnosis" of

why an ethical view is justified. In contrast, "*a modal clause always conveys a judgment by its source about the motivations of the modal clause's semantic subject*. . . . In modality analysis, the coder's challenge is to learn more about how the author understands others' motivations (i.e., to get into the mind of someone who is getting into someone else's mind, as it were)" (Roberts et al. 2008).

One does not have a modal clause if the modal auxiliary verb's *semantic subject is not a person* (or a metonym for a person[s]). Thus, one would not encode the following as a modal clause, because the subject of "can" is "economy":

> Why are we pretending as if we have a market economy, which *can* be described by similar numbers like the German [economy]? (August 13, 1996, p. 3)

Yet the following would be encoded as a modal clause, despite the fact that its grammatical subject is not a person:

> Hungarian society is just learning the rules of democracy, and developing the rules of the game, which *must* be adopted by politicians. (October 15, 1993, p. 15)

This is because its use of "must" is in passive voice, and it has people (namely, politicians) as its semantic subject.

As it turns out, this last quotation was not encoded after all, because it did not pass the second step in the encoding process. The editorial in which it appeared contained *no rationale* for "politicians' compulsion to adopt democratic rules" in its first or last three paragraphs. More generally, every encodable instance of modal usage must meet the following criteria:

- The semantic subject of the modal usage must be a member of the population (e.g., citizen and resident of the country) about which one is drawing inferences.
- Modal usages are not encoded when they are direct quotations of others' words.
- A rationale for the modal usage must be explicitly given within one's sampled texts. (In our Hungarian research, only the first and last three paragraphs of each editorial were examined, in part to keep volumes of text comparable among editorials.)

According to the first criterion, modal auxiliary verb use should not be encoded either when its semantic subject is a nonperson or when they are persons who are expatriated or are citizens of other countries. Thus, the

following would not be an encodable modal usage within a study of Hungarian public discourse:

> In Belgium patients receive a small booklet, in which all their prescribed medicines are registered. This way they *can* avoid the prescription of needless or incompatible medicines, especially when the patient visits many physicians. (January 27, 1997, p. 11)

According to the second criterion, the following modal auxiliary verb occurrences would not be encodable, given that they appear *within a direct quotation* from another source:

> The rent increase is going to be compensated to the needy people, as the mayor said, "The value of subsidies will be differentiated, because some people *have to* cover the increased costs fully from their own resources, while others *can* expect greater support." (November 23, 1992, p. 10)

According to the third criterion, appeals to the reader's common sense are insufficient as rationales associated with modal usages:

> Naturally, one association *cannot* take such a responsibility alone. (February 1, 1994, p. 14)

> [E]ven if they *had to* refer to him for some reason, his name was simply mentioned as "Géza Something" in the political columns of the newspapers. (March 29, 1996, p. 11)

To be encodable, a modal usage must be linked to a rationale that indicates more than a general assertion that something is "naturally" impossible or that there were "some reasons" why references to someone were inevitable. Each encodable modal usage calls for an *explicitly stated rationale* within its associated text.

Disagreements in Identification

Of the 485 modal usages identified by either of two coders in our sample of Hungarian editorials, only twenty-nine were coded by one but not the other. The vast majority of their differences (twenty or 69%) resulted because what one coder saw as indicative of the subject's intentionality, the other interpreted as descriptive or diagnostic. The following excerpt is provided to convey the difficulty involved in drawing this distinction:

> What did he get from the mailman? A check that could be cashed in a bank in Graz, for which he *was able to* buy a gambling automat or playing machine that was written off there [in Austria] already. (July 17, 1995, p. 11)

Is the author merely describing the increase in someone's purchasing capacity by virtue of having received a check in the mail, or is this a reference to the person's intention to buy something that only became possible once the check had arrived? The latter would suggest its being a modal usage; the former would not. Or how about a reference to "a telephone company, which *could* build on the existing network" (October 13, 1992, p. 3)? Is this a description of a type of phone company or a reference to the phone company as an agent for which "building on the existing network" is a possibility? At times, even native speakers may disagree about whether an author has used a modal auxiliary verb to impute intentionality to the verb's subject.

Beyond this, three of the coders' differences were whether a rationale was provided, and three were on whether an auxiliary verb's subject was a person. (In the latter case, one might reasonably ask whether the previous quotation's author anthropomorphizes the "phone company" as an agent with a capacity to "build.") In two additional cases, one coder failed to recognize that the modal auxiliary verb appeared within a direct quotation.

The final difference in coders' identifications of modal usages was over a modal auxiliary verb that appeared in a question:

> It is worthwhile to think about the issue: *could* we . . . regulate so much the different levels of interest-harmonization? (April 19, 1991, p. 6)

In this case, we have a *serious question*, which should not be encoded, as the author does not take a position on whether it is possible for Hungarians to regulate. On the other hand, coders should encode modal usages within *rhetorical questions*:

> I think, the new leaders of the National Metal and Mineral Mines . . . inherited the more than 700 million forints burden of HAF [a used battery processing plant]. . . . [T]he authorities who permitted [this] do not provide any meaningful financial support, saying that this is a [private] company's investment. *Can* we talk about only a company's interest, even when the country becomes a big waste disposal field? . . . [T]he people of Gyöngyösoroszi . . . were not involved in the planning of the investment, they could only hear references to foreign plants, guarantees, experts. That is why they may have doubts today. (May 16, 1991, p. 12)

What distinguishes a serious question from a rhetorical one is that its author presumes that her or his audience knows the author's position on the question as it is stated. Here, the author's intent is clearly that Hungarians

are *able not* to refer merely to companies' interests when public welfare is at risk. In contrast, serious questions leave the author's intent masked. Thus, the coding rule here is that one encodes modal usages within rhetorical questions but not within serious ones.

A Note on Interrater Agreement

After identifying one's units of analysis (here, modal + rationale tuples) and developing a coding schema, one may evaluate how consistently coders apply the latter to the former in practice. The following two sections focus on discrepancies in two coders' applications of modal-form and rationale categories in the 456 modal + rationale tuples that they coded in common. Here we use kappa (Scott 1955; Popping 2000: 132-37) as our measure of interrater agreement. (Calculations were obtained using the computer program AGREE 7.3 [February 2008] [Popping 1984].) The measure uses the marginal distribution over all coders as the basis for computing deviations from expected agreement and is appropriate when comparing a large number of classifications among trained coders.

Beyond documenting acceptable levels of intersubjectivity in one's research, interrater agreement statistics like kappa afford occasions for self-assessment. The following sections reach behind these measures to illuminate precisely where coder disagreements originate—problems to which we offer solutions in hopes of greater agreement in our (and hopefully our readers') future studies. Indeed, given the ever-elusive nature of interrater agreement, intercoder reliability remains an open-ended project to every content analyst (Lombard, Snyder-Duch, and Bracken 2002; Hruschka et al. 2004).

CLASSIFYING MODAL FORMS

In analyzing the Hungarian editorials, our research objectives only called for encoding instances of the modal auxiliary verbs, "can" ("could" or "able to") and "must" ("have to" or "need to"). Instances of these verbs are typically used to convey possibility ("can," "could," "is able"), impossibility ("can't," "cannot," "couldn't," "could not," "is unable"), inevitability ("have to," "has to," "had to," "must," "need to"), or contingency ("mustn't," "must not," "needn't," "need not"). After locating all such character strings in our texts, coders were asked to consider each string and then either to eliminate it from consideration (because it was not a modality usage, as discussed in the previous section) or to classify it into one of the

TABLE 1
Two Coders' Assignments of Modal Auxiliary Verb Instances to Four Modal Forms

		Coder 2					
		Possible	*Impossible*	*Inevitable*	*Contingent*	*Not Coded*	*Total*
Coder 1	Possible	217	24	0	0	15	256
	Impossible	2	120	0	0	2	122
	Inevitable	0	0	84	0	2	86
	Contingent	0	1	0	8	0	9
	Not coded	4	3	3	2	0	12
	Total	223	148	87	10	17	485

four modal forms that the string's author intended to convey with it. Table 1 provides a summary of two coders' agreements and disagreements in making these modal-form classifications for the 485 modal usages that at least one of them coded (i.e., did not eliminate).

If one eliminates the twenty-nine cases already discussed in the previous section (i.e., cases in which only one of the coders assigned a modal form or rationale), the kappa statistic for this table equals .91 ($z = 413.33$). Most noteworthy is that twenty-six of their twenty-seven discrepancies occur when one coder coded an instance as "possible" that the other coded as "impossible." As it turns out, of these twenty-six discrepancies, seventeen (65%) are for the single expression, "can only":

> These appointments created only the minimum conditions . . . to clear up the ruins. . . . The now appointed presidents and vice-presidents *can only* build on their personal credibilities. Experiences show that this [credibility] fades away very quickly. (July 21, 1994, p. 3)

> We *can only* rely on our physicians in an emergency situation. (February 28, 1996, p. 11)

In reviewing these seventeen cases, our judgment is that in most (namely, thirteen) cases, both coders encoded them incorrectly. If one is only able to pursue a single activity, then one finds that activity *inevitable* (i.e., neither possible nor impossible, as assigned by the coders). The phrase, "can just," is also commonly used in conveying inevitability:

> Until then we *can just* hope that yesterday's appointments do not start another round in the four-year long war. (July 21, 1994, p. 3)

Accordingly, the author conveys in this quotation that "we Hungarians" must (inevitably) hope for peace.

So when might "can only" expressions convey possibility or impossibility? The answer lies with the words being modified by "only." If "only" modifies "can," then inevitability is the meaning being conveyed. However, if "only" modifies a different part of the clause containing the modal auxiliary verb "can," the verb's modal form is correctly coded as "possible":

> [An excavation] could prove unambiguously whether the skeleton . . . is that of Petöfi or not. [The excavation's sponsor] argues that even the Academy is not averse to approve the opening of the grave. However, in a letter . . . by [leaders of the Academy] we *can only* read that the Academy is for the freedom of research and it doesn't want to object to the publication of points of view that are different from those of the Academy. (September 27, 1996, p. 15)

> Palotás . . . led the people's movement to victory against the government, or at least he pressured it into a compromise with his hard siege of arguments, that *nobody else* [than Palotás] *was able to* achieve until the last minute, [but Palotás] was also from a governing party, or [at least] got his mandate as a Parliamentary Representative with the support of the MDF [the governing party]. (November 14, 1990, p. 7)

Note that in the first quotation, the author is not conveying the inability not to read (i.e., the inevitability of reading) a letter, but the possibility of reading "only specific content" within it. That is, "only" modifies "that the Academy 'is for freedom' and 'doesn't want to object.'" It does not modify "can." In the latter quotation, replacing "nobody else [than Palotás] was able" with "only Palotás was able," reveals that this only + can combination is one in which "only" modifies the subject, Palotás. So why not code this as "*every person other than Palotás* was not able to achieve a compromise"? The answer lies in the rationale linked to the modal usage. If the author provided a rationale for why compromise was impossible for everyone else, then this latter coding would be appropriate. However, here the rationale lies with Palotás's unique political connections (not others' lack thereof) that made his achievement possible.

Impossibility is the correct encoding of only + can usages in which "only" modifies the rationale as a whole:

> The government *can only* expect self-control from people who live on wages and salaries, when it explains why it is good for them in the long-run. (September 27, 1994, p. 3)

> The first step of the reform is going to be the revision of the state's tasks, i.e., in the long-term the state *can only* assume responsibilities that can be financed by the state at a high quality level. (December 20, 1994, p. 13)

The first quotation can be rephrased as "the government can expect self-control, only if it explains why it is good"; the latter can be rephrased as "the state can assume responsibilities, only if they can be financed at a high level." Yet these phrases do not imply that "explaining why self-control is good" is the reason why "the government can expect it" or that "financing responsibilities at a high level" is the reason why "the state can assume them." Instead, each imputes a *necessary condition* for a state activity:

- If the government does *not* explain why self-control is good, then it can*not* expect it.
- If the state does *not* finance responsibilities at a high level, then it can*not* assume them.

These latter rephrasings illustrate that "can only" conveys impossibility whenever the "only" modifies the rationale clause rather than a component of the modal clause.

Summarizing the last three paragraphs, only + can usages that convey intentionality should be encoded as follows:

- If "only" modifies "can," encode as conveying inevitability.
- If "only" modifies another part (e.g., subject, object, main verb) in the modal clause, encode as conveying possibility.
- If "only" modifies the rationale clause, encode as conveying impossibility.

As should be clear at this point, encoding during modality analysis requires considerable familiarity with grammatical structure. Fortunately, high interrater agreement is still obtainable as long as such esoteric phrases as "can only," "can just," and "can hardly" are relatively infrequent in one's sampled texts.

The only other recurrent discrepancy in coders' assignment of modal forms was with three instances of the phrase, "can hardly":

> Somebody might be an Albert Schweitzer in our small country, if he doesn't have an impact factor, quotation index, and he isn't supported by power, then he *can hardly* expect a progressive career. (September 29, 1993, p. 18)

One can appreciate the subtlety of this phrase by noting that whereas "can hardly" conveys impossibility, "can barely" conveys possibility. ("I *can* just barely fit.") Nonetheless, one coder seems not to have been aware that this phrase (in Hungarian as well as in English) conveys impossibility, not restricted possibility.

Other discrepancies seem to have been due to insufficiently careful readings of the text:

> [T]rade unions cannot make a decision even [when] they fully agree, and the MSZOSZ [a trade union] is even *less able to* do so. (January 3, 1994, p. 12)

> Even if we can say many bad things about a politician or policy, we still cannot generalize this to all politicians and policies, and we *can say even less so* that the facts exclude the existence of theoretical opportunity for an alternative policy. (April 1, 1994, p. 14)

Clearly, decision making that is "less than impossible" is itself impossible, and fact saying that (when compared to generalizing) is "less than impossible" is itself impossible.

On the other hand, coder discrepancies occasionally resulted from overly careful text readings. For example, one editorial refers to "Governing parties [that] lived in an illusion, according to which parties of the moderate right, were *not going to be able* to create an alliance" (April 8, 1997, p. 3). Although the italicized phrase here conveys impossibility, the author asserts that this impossibility was illusory. Which does one encode: the modal's explicit form or the author's repudiation of that form? Because our purpose is to encode authors' imputations of intentionality to their contemporary fellow countrymen, our encoding is of the author's conveyance of the governing parties' view that nongoverning parties' alliance creating was impossible. The author's repudiation of this is a mere descriptive (i.e., nonmodal) statement in light of subsequent historical events.

Note that this modal usage would be excluded from our data set, if it were to refer to Hungarians' intentions during the distant past (i.e., if the intentions were not of *contemporary* Hungarians). At issue is "how distant" the past must be for an intention to be no longer considered contemporary. Here, the reference is to the intentions of members of the socialist and liberal–democrat coalition at the beginning of a still-ongoing political campaign. If our analysis were restricted to intentionality during this specific campaign, such a modality reference could not be considered contemporary. Yet, given our lengthier 7-year time frame, it seems reasonable to retain it as a datum for analysis.

Nonetheless, this case should not lead one to routinely ignore authors' references to illusion. Consider the following illustration from among the twenty-seven discrepancies in assigning modal forms:

> [*C*]*an* we expect with such low salaries that the best teaching and psychological talents are going to choose caretaker careers? Obviously this is only a vain illusion. (July 12, 1991, p. 8)

Here, by applying the term "illusion" to the expectation referred to in the rhetorical question, the author makes clear the meaning he wishes to

convey—namely, that we cannot (i.e., it is impossible for us to) hold such an expectation.

In sum, we can report exceptionally high interrater agreement in coders' assignments of modal forms to identical instances of modal auxiliary verbs. However, improvement in this agreement is possible if "can only" instances are encoded in accordance with the clause or sentence-part modified by "only" and if "can hardly" instances are consistently encoded as conveying impossibility. Beyond this, minor additional improvements require coders' remaining sensitive to the overall objective of encoding "imputed intentionality" as well as to other nuances in authors' language use.

CLASSIFYING RATIONALES

As already mentioned, every modal usage is (at least in principle) justifiable in terms of some rationale. That is, after asserting that some action or situation is possible, impossible, inevitable, or contingent, authors may always be asked to explain this possibility, impossibility, inevitability, or contingency. Why was it possible for Palotás to achieve the compromise? Why is it impossible for us to expect talented teachers to choose caretaker careers? Newspaper editorials afford authors opportunities to spell out precisely what their reasons might be for such imputations of intentionality, for example, in Palotás's ability to achieve political compromise or in our inability to expect others' choices. Yet such opportunities may not be taken either because authors refrain from making their rationales explicit or because they assume a diagnostic or descriptive stance toward their modal statements. (For example, an author might simply presume expert knowledge of all political actors and simply pronounce Palotás to have been the only politician "with an innate capacity" to achieve compromise.) In these cases, encoding would be inappropriate, as discussed at length in this article's second section.

Generally speaking, the number of rationale categories should be kept to a minimum. This not only improves interrater agreement but also reduces the number of cells in one's contingency tables at the time statistical analyses are performed. The sparser one's table, the less powerful hypothesis tests will be. Our initial set of categories for classifying rationale types consisted of "political, economic, cultural, and security related." To this, "welfare related" was later added to differentiate a fifth set of rationale instances poorly captured by these four. Coders were provided the following guidance when encoding rationales:

TABLE 2
Two Coders' Assignments of Rationale Instances to Five Categories

		Coder 2						
		Politics	*Economy*	*Culture*	*Security*	*Welfare*	*Not Coded*	*Total*
Coder 1	Politics	141	5	11	0	0	8	165
	Economy	0	120	4	0	3	3	130
	Culture	10	1	80	0	11	3	105
	Security	5	3	0	36	0	1	45
	Welfare	1	1	0	0	24	2	28
	Not coded	5	4	3	0	0	0	12
	Total	162	134	98	36	38	17	485

- *Political* rationales account for the possibility, impossibility, inevitability, or contingency of actions or situations as resulting from *activities by politicians and political bodies toward strengthening or maintaining their power*.
- *Economic* rationales explain them as occurring because of *aspects of the market* (e.g., efficiency or profitability) *and segments of the economy* (e.g., agriculture or industry).
- C*ultural rationales* account for them as resulting from a *nation's heritage, language, morality, and so on.*
- *Security-related* rationales explain them in terms of *safety, order, or the military regarding protection of the nation's citizenry.*
- *Welfare-related* rationales justify them in terms of *national well-being regarding health, education, unemployment benefits, elderly care, and so on.*

Table 2 provides a summary of two coders' agreements and disagreements in assigning these rationale classifications to the 485 rationales coded by at least one of them.

Considering only the 456 rationales (and modal forms) encoded by both coders, the kappa statistic for this table equals .84 ($z = 479.80$). Although the table's most frequently interchanged categories are those of "politics" and "culture," on closer examination, a more fundamental explanation for discrepancies in rationale-assignment emerges. The table's most recurrent source of discrepancy is the erroneous assignment of a modal form's topic to its rationale. For example, after criticizing contemporary attempts to apply the un-Hungarian label to political dissent, an editorialist writes that a citizen's opinion may

> be right, wrong, prejudiced, mistaken, undecided, or anything, but it cannot evaluate the nationality of the person at all. The government could have a Hungarian opinion, but a citizen *cannot* have a Hungarian point of view. A citizen is more and less than that at the same time. That is the difference

> between an opinion and the chanting choir of citizens in a military order. (August 5, 1993, p. 3)

In contrast, another writer argues that

> the current government should pay at least as much attention as its predecessor did, to the opinion and needs of the 3 million Hungarians who live abroad . . . otherwise the government *cannot* clear itself from the accusation that it signed an agreement "about them, but without them," and such an accusation would be a very sensible issue, given that the leading party of the coalition is the successor of [the pre-1989 communist] party, that committed hardly reparable sins against Hungarian minorities between 1956 and 1988–89. (August 17, 1994, p. 11)

Note how the first quotation indicates that a single Hungarian (*cultural*) viewpoint is impossible for *political* reasons (namely, because citizens do not "chant" their political leaders' opinions). The second, contrasting quotation indicates that a *political* act (a party's clearing itself of having been insensitive to Hungarian expatriates) is impossible for *cultural* reasons (namely, given the party's historical ties to a regime that was undeniably insensitive in this way). At least fifteen discrepancies in rationale assignment appear because of a coder's nonrecognition that the subject matter of a modal usage is different from that of its rationale.

Beyond this, most discrepancies seem to have resulted from differences in how coders conceptualized the five rationale categories. For example, coder 2 assigned the politics category to rationales involving agreements among politicians from different countries, whereas coder 1 generally assigned security to such rationales. Moreover, coder 1 tended to assign culture to all rationales involving human suffering, whereas coder 2 generally assigned welfare in these cases. So despite its being acceptably high, we conclude that coders' interrater agreement could be improved with more rigorous definition of our rationale categories. Accordingly, we provide the following revised guidance for coders' classifications of rationales:

- *Cultural rationales* are ones *grounded in a country's domestic past* (i.e., those historical events or traditions that "make us who we are"). They may also involve writers' *references to a commonly accepted morality or to their own judgments* (presumably, ones like those of their readers and most other citizens). The following are three illustrations from our data of cultural rationales that reference the past:
  - My unexaggerated speech about majority participation in the 1956 uprising is possible, because at the time "people were smiling, happy, with faith in the birth of a free Hungary, [and] were rejoicing with each other."

- Many people's creation of personal autonomy is impossible, because "a significant part of the society got stuck between the traditional and a civil lifestyle."
- People's acting without consequences was possible, because of "the era . . . [that] is [now] over."

Now consider three illustrations of self-referential cultural rationales:

- My belief that a politician's impartiality as a publisher is possible, because this is someone "very sympathetic to me."
- The prohibiting of the people's faith and efforts is impossible, because of the "spiritual abilities [with which] [o]ur country has never been poorly endowed."
- Sufficiently strong emphasis of my understanding is impossible, because "I know what it is: not to know what is right, when I act appropriately as a human being."

The following are three cultural rationales that refer more directly to morality:

- Someone's understanding of capitalism is impossible, because "his questions are moral questions."
- Distinguishing between legality and morality regarding the death penalty is impossible, because "[t]his is, after all, a question of conscience."
- Acting for the country in ways worthy of obscenely high compensation is impossible, because such compensation would not be "acceptable in the eye of Hungarian citizens."

Thus, in general, cultural rationales reference nations' uncontestable historical and ethical foundations—the source of their social order.

As depicted in Table 3, the other four rationale categories reference citizens' "application" or "manipulation" of their social order either "by using its internal workings" or "in response to external threats." Following Swidler's (1986) distinction (based on Peter Stromberg's work) between ideology and common sense, our conceptualization of political, economic, security-, and welfare-related rationales is as involving ideological contestation whereas that of cultural rationales is as proffering common sense (see Geertz 1975). The following are our revised guidelines for encoding noncultural rationales for modal usages within a nation's discursive depictions of intentionality among its citizens:

- *Political* rationales account for the possibility, impossibility, inevitability, or contingency of actions or situations as resulting from *manipulations of the national order's internal workings by those of its citizens empowered to do so in ways consistent or inconsistent with the will of the nation's electorate*. Thus, according to this definition, interactions between politicians of different

TABLE 3
Guidelines for Assigning Rationale Instances to the Categories of Politics, Economics, Security, and Welfare

	Citizens'	
their social order	*application of*	*manipulation of*
by using its internal workings	**Economics** • Regarding *global* markets • Producers vs. consumers • Competition • Budget constraints • Supply vs. demand • Technological developments	**Politics** • Regarding *national* accountability • Leaders as public servants • Leaders: political, corporate, special interest, union, lobby • Ineptitude vs. corruption • Vested interests vs. public trust • Reappointment and reelection
in response to external threats	**Welfare** • Threat *management* • Recipients vs. providers of services • Recipients' abuse vs. restraint • Providers' neglect vs. responsibility • Services: subsistence, health, education, employment, environmental conservation, elderly care, etc.	**Security** • Threat *response* • Weak citizens vs. strong military • Threat containment vs prevention • Domestic vs. foreign violence • Affinity vs. animosity re noncitizens • Citizens' (expatriates') safety abroad • military viability (expenditures)

countries are not inherently political. Moreover, not only might political rationales involve politicians and political parties, they may also involve other social servants such as unions, lobbyists, and corporate leaders. Unlike cultural rationales, the issue here is not with the morality of these leaders' activities but with the extent to which their activities conform to public sentiments. The extent of this conformity may be depicted as arising because of the leaders' ineptness, their corrupt self-interest, their desire to maintain power (e.g., via reelection or ongoing appointment), or their sincere desire to maintain the public trust or conform to the will of the electorate. Illustrations of political rationales from our data are as follows:

- Signing a petition is possible, because "today is not too late to stop . . . government-initiated monopolization of . . . the media."
- Wiretappers' unobstructed hearing is possible, because "it is . . . in the interest of [government-sponsored] eavesdroppers to have a good [phone] connection."
- Nonpunishment of corrupt politicians is possible, because "this system must be changed . . . [so that it] serves the electorate."

  - Punishment of lying news reporters is impossible, because "the rude and inhuman laws of the system distort all the participants."
  - Ignoring potential corruption by appointed task groups is impossible, because "an ad hoc task group . . . can hide opportunities for abuse."
  - Curriculum change is inevitable, because "current [education policy] . . . is an example of wrong [political] compromises."
- *Economic* rationales account for the possibility, impossibility, inevitability, or contingency of actions or situations as resulting from *participation in (i.e., using the internal workings of) the global economic order*. Whereas the ultimate legitimacy of democratically elected politicians' activities lies with their nation's electorate, the viability of economic activities is played out in markets that often span the borders of many nations. Economic rationales incorporate references to market processes such as technological developments, budget constraints, supply versus demand, domestic or international competition, and so on. Some illustrations of economic rationales from our data are as follows:
  - Our [sarcastic] gratitude for Western acceptance of our labor is possible, because there was "selling [of] Hungarian firms to Western capitalists for nothing."
  - Our experiences of sickening displays of wealth are possible, because "it is drawn from [black market] resources drawn away from the economy."
  - Our choice between the collapse either of the budget or of public welfare is inevitable, because there is an imminent "shadow of inflation."
  - Excess public-sector employment is impossible, "because it would burden the city with many million forints of additional expenditures."
  - The government's watering down of the Bokros-package [of strict social and economic measures] is impossible, because "Hungary . . . should position itself as a strong discussion partner with its [economic] performance."
- *Security-related* rationales account for the possibility, impossibility, inevitability, or contingency of actions or situations as resulting from *manipulation (or mobilization) of the national social order in response to impending or manifest threats to this order*. They refer to the protection of law-abiding citizens from others' violence, with the ultimate counterthreat of police or military force. Protection may enlist noncitizens with whom one has developed affinities, in alliances against others toward whom one has developed animosity. Defenseless citizens may be expatriated or locally resident. Although threats may be hypothetical or ongoing (or range in severity from petty crimes to genocide), they are always depicted as originating with humans, namely domestic criminals or hostile foreigners. The following are illustrations of security-related rationales from our data:
  - In its diplomacy, the government's going below a reasonable minimum is impossible, because this would constitute "abdicating responsibility for the interests of Hungarian minorities [in Romania and Slovakia]."
  - Our negotiations with Slovakia regarding a water-power plant have been impossible, because of "emotional, historical, and national scars of the past."
  - Our accurate guessing of imminent political impasses with bordering nations is possible, because "[c]entury-long historical debates will not be closed."

- Our hope that European institutions will force Romanian acquiescence is possible, because "[p]eople here always trust more in [the security potential from] European civilization than in Europe itself."
- Our worrying about the fate of our mariners is inevitable, because "Hungarian boats on the Danube . . . are [being] pirated by Serbs."

Note how the middle three of these illustrations refer to the historical origins of negative or positive sentiments from citizens of one country toward citizens of another country. In contrast to cultural rationales that may reference the *domestic* historical origins of a nation's character, security-related rationales often reference the *international* historical origins of animosities with or affinities to other nations' citizens.

- *Welfare-related* rationales account for the possibility, impossibility, inevitability, or contingency of actions or situations as resulting from *applying the nation's social order (i.e., providing its services) for the management of known threats*. Citizens may either be "service providers" or "service recipients." Recipients may be depicted as abusing the system or restrained in their use of it. Providers may be depicted as negligent or responsible. Known threats for which services may be provided include (in part) subsistence, health, education, employment, environmental conservation, and elderly care. Some illustrations follow:
  - Our tolerance of political ineptitude is impossible, because "the situation of our health care . . . threaten[s] us with further deterioration."
  - [Unsuccessful] attempts to save patients are inevitable, because physicians "are not like God. . . . Even the smallest surgical procedures have risk factors."
  - After having been publicly abused in the media, a citizen's regaining of lost prestige or enthusiasm is impossible, because the media "do not inform the readers [of the citizen's] professional and ethical exoneration from the charges."
  - Accounting for some Hungarians' unauthorized access to government subsidies is inevitable, because "only those people should receive some kind of government [subsidies], who really need it."
  - Hungarian society's acceptance of "the new capitalists" is possible, because "[a]ccording to Imre Nagy, the capitalist is going to become a responsible citizen."

Like political rationales, welfare-related rationales reference a nation's *domestic* affairs. At stake with the former are the motivations of leaders when deciding which services the nation is to provide. Yet with welfare-related rationales, the quality (not the existence) of services is at stake. Thus, the rationale is welfare related, not political, if political acts are criticized for their impact on social services. The converse holds if they are criticized for not heeding public opinion.

CONCLUSION

This article has "opened our shop doors" to researchers interested in applying modality analysis within studies of rhetorical references to intentionality—especially to those involved in historical-comparative analyses of national discourse. By critically reviewing discrepancies in applications of our coding scheme, we hope to have both improved our methodology and conveyed it more clearly for others' use. For instance, when identifying modal usages, we have provided guidelines for improved discernment of imputed intentionality (as distinct from descriptions or diagnostic statements) and of citizens referenced as (possibly anthropomorphized) semantic subjects. When identifying rationales, we recommend seeking them within rhetorical but not serious questions. And when classifying modal usages, we provide guidelines for encoding phrases in which the modal auxiliary verb "can," is associated with adverbs such as "only," "just," "hardly," and "barely." But most importantly, our recommendations on rationale classification afford much greater depth and clarity to our definition of "national discourse" (i.e., the subject matter of the editorials we study). Let us be specific.

Analyzing national discourse not only requires obtaining a sample of representative texts or transcripts, it also requires developing a nominal definition for this type of discourse. Our definition emerges from the domain of rationales that we associate with a nation's discourse. Implicit in this domain (as delineated in Table 3) is a specific definition of nation.

Based on our rationale scheme, nation can be understood as a collection of citizens both (1) with an elected leadership that makes policies regarding services provided to these citizens and (2) within an environment that allows economic participation or that threatens the survival of the nation itself (i.e., of its leadership election, service provision, and/or economic participation). Citizens are thus people with electoral power, economic access, and rights to services. Leadership positions are contingent on citizens' satisfaction with their leaders' decisions. Services are contingent on providers' responsibility in providing them and on citizens' restraint in using them. Whereas anyone may participate in the world economy or may threaten a nation's social order, only citizens may vote or receive their nation's services.

In confining our research to this definition, note that we have thereby limited our analyses to discourse within democracies. Keeping in mind Stalin's proviso, "It does not matter who votes; what matters is who counts the votes," this definition allows us to speak of a nation's "degree" of democracy (i.e., the extent to which its politicians respond to citizens' votes/preferences). Likewise based on this definition, one may refer to the degrees to which nations ensure that citizens are provided services. Further

variations are opened for consideration among nations' cultural origins and their economic and security-related challenges. In short, emerging from our exercise in clarifying rationale categories is a concept of nation that has broad applicability to the countries of today's world.

Yet repeated appeals to political rationales could just as well be signs of a vibrant democracy or of one in jeopardy. Welfare- and security-related rationales could likewise be used to convey hope or despair, as could cultural or economic rationales. So we return to the kernel of modality analysis and to the modal usages to which these rationale types are discursively linked. One might argue, for example, that a nation's economic discourse is unproblematic to the extent that economic rationales account for citizens' possibilities. Or its welfare discourse may be generally acceptable to the extent that welfare-related rationales account for citizens' inevitabilities (say, by referring to what they must do as responsible providers and users of services).

Of course, modality analyses might be applied to populations of discourse other than national ones. Research on rhetorics of intentionality within corporate discourse or doctor–patient conversations, for example, would involve the development of corresponding rationale schemas that capture the discursive domain at hand. At issue, then, is how employees and employers, patients and doctors, and so on, appeal to these rationales in persuading and being persuaded of things possible, impossible, inevitable, or contingent for each other. Yet, at this point, we approach the limits of this article's methodological scope and close with our hopes of having provided sufficient insight into the variety of substantive applications that modality analysis affords.

REFERENCES

Berelson, B. 1952. *Content analysis in communication research*. New York: Free Press.

Geertz, C. 1975. Common sense as a cultural system. *Antioch Review* 33 (1): 5–26.

Holsti, O. R. 1969. *Content analysis for the social sciences and humanities*. London: Addison Wesley.

Horn, L. R. 1989. *A natural history of negation*. Chicago: University of Chicago Press.

Hruschka, D. J., D. Schwartz, D. Cobb St. John, E. Picone-Decaro, R. A. Jenkins, and J. W. Carey. 2004. Reliability in coding open-ended data: Lessons learned from HIV behavioral research. *Field Methods* 16 (16): 307–31.

Krippendorff, K. 2004. *Content analysis: An introduction to its methodology*. Thousand Oaks, CA: Sage.

Lombard, M., J. Snyder-Duch, and C. C. Bracken. 2002. Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human Communication Research* 28 (4): 587–604.

Markoff, J., G. Shapiro, and S. Weitman. 1974. Toward the integration of content analysis and general methodology. In *Sociological methodology, 1975*, ed. D. R. Heise, 1–58. San Francisco: Jossey-Bass.

Popping, R. 1984. AGREE, A package for computing nominal scale agreement. *Computational Statistics and Data Analysis* 2 (2): 182–85.

———. 2000. *Computer-assisted text analysis*. London: Sage.

Roberts, C. W. 1989. Other than counting words: A linguistic approach to content analysis. *Social Forces* 68 (1): 147–77.

———. 1997. Semantic text analysis: On the structure of linguistic ambiguity in ordinary discourse. In *Text analysis for the social sciences: Methods for drawing statistical inferences from texts and transcripts*, ed. C. W. Roberts, 55–77. Mahwah, NJ: Lawrence Erlbaum.

———. 2008. *"The" fifth modality: On languages that shape our motivations and cultures*. Leiden, Netherlands: Brill.

Roberts, C. W., R. Popping, and Y. Pan. Forthcoming. Modalities of democratic transformation: Forms of public discourse within Hungary's largest newspaper, 1990–1997. *International Sociology*.

Roberts, C. W., C. Zuell, J. Landmann, and Y. Wang. 2008. Modality analysis: A semantic grammar for imputations of intentionality in texts. *Quality and Quantity*. (Prepublished September 12, 2008; DOI: 10.1007/s11135–008–9194–7.)

Scott, W. A. 1955. Reliability of content analysis: The case of nominal scale coding. *Public Opinion Quarterly* 19 (3): 321–25.

Shapiro, G., and J. Markoff. 1998. *Revolutionary demands: A content analysis of the cahiers de doléances of 1789*. Stanford, CA: Stanford University Press.

Swidler, A. 1986. Culture in action: Symbols and strategies. *American Sociological Review* 51 (2): 273–86.

Van der Auwera, J. 1996. Modality: The three-layered scalar square. *Journal of Semantics* 13 (3): 181–95.

ROEL POPPING is in the Department of Sociology at the University of Groningen, Netherlands. His research is on historical shifts in public opinion, values, and scientific knowledge, primarily within the context of post-1989 Central and Eastern Europe. His book, Computer-assisted Text Analysis, *was published by Sage in 2000. He has articles forthcoming in* Quality & Quantity *("Some Views on Agreement to Be Used in Content Analysis Studies") and* International Sociology *("Modalities of Democratic Transformation: Forms of Public Discourse within Hungary's Largest Newspaper, 1990–1997").*

CARL W. ROBERTS is an associate professor with a joint appointment in the Sociology and Statistics Departments at Iowa State University. His current research on theory ("The" Fifth Modality: On Languages That Shape Our Motivations and Cultures, Brill, 2008) and intercultural comparisons ("Modalities of Democratic Transformation: Forms of Public Discourse within Hungary's Largest Newspaper," International Sociology, *forthcoming) 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 *and* Sociological Methodology as *well as in his edited collection,* Text Analysis for the Social Sciences *(Lawrence Erlbaum, 1997). His empirical work investigates cultural variations in social discourse based on sampled texts ranging from medieval documents to editorials in contemporary foreign and domestic newspapers.*