

On Quantitative Narrative Analysis

Roberto Franzosi

When anthropologist Margery Wolf came across a story about a village event she had jotted down some 30 years earlier, during field work in a small village in Northern Taiwan, she discovered that her story did not quite fit the data (her notes, her recollections, her interviews). So she decided to rewrite her story. In the end, she did not have just one story; she had three stories—each with “a different perspective . . . a different style . . . different outcomes . . . yet all three involve[d] the same set of events” (Wolf, 1992, p. 7).

This chapter tells a story about stories. It looks at the same material (the concept of story itself) from three different perspectives: that of the rhetorician, the linguist, the journalist (and, ultimately, the sociologist). It spans 2,500 years of people’s tugging and pulling at stories. Contrary to Margery Wolf, these different perspectives on stories have the same outcome: Stories have a simple fundamental structure based on a sequence of 5 Ws + H (who, what, when, where, why, and how). It is this property of stories that this sociologist, knowing no rhetoric or linguistics, was lucky enough to stumble upon and use to his advantage in developing a new methodological approach to stories: Quantitative Narrative Analysis.

But enough with preambles. (They ought to be short and clear, the rhetorician tells us!) Let's get on with our story (our stories, in fact).

The Rhetorician's Story

Rudyard Kipling's (1902) "The Elephant's Child," one of the stories in his collection *Just So Stories*, ends with a brief poem, which begins as follows:

I keep six honest serving-men
 (They taught me all I knew);
 Their names are What and Why and When
 And How and Where and Who.

Kipling poetically summarized here the tradition of the so-called 5 Ws + H of narrative (who, what, when, where, why, plus how) that stretches back to antiquity, to Hermagoras of Temnos, Gorgias of Leontini, Cicero, and Quintilian, although it is probably through Aphthonius that medieval and Renaissance students became familiar with the concept.¹ Aphthonius collected a set of rhetorical exercises known as *progymnasmata* (Clark, 1952). One of the exercises Aphthonius prescribed was on narrative. He wrote the following: "The tale is concerned with six considerations: the personal agent, the thing done, at what time, in what place, in what manner, and for what cause." These, of course, would be Kipling's six honest serving-men: the Who, the What, the When, the Where, the How, and the Why. Thus, the story of the 5 Ws and H is an old story, a story better known in rhetoric under the term *circumstances*.

Both Cicero, in *De Inventione*, and Quintilian, in *Institutio oratoria*, discussed circumstances (or attributes) as mitigating or aggravating factors of a case in a trial.² They distinguished between attributes of persons and things (i.e., actions). About the attributes of "things," Quintilian wrote "for all actions, we ask why, or where, or when, or how, or with which instruments they occurred" (V, 10, 32; repeated as a summary at V, 10, 94). Yet neither Cicero nor Quintilian proposed a catching mnemonic verse, the Latin rendering of the 5 Ws and H. Apparently, it was one of Cicero's near contemporaries, the Greek rhetor Hermagoras of Temnos, active in Rome in the 2nd century BC, who came up with the verse—in fact, with the very theory of circumstance, or *peristasis* (Kennedy, 1994, p. 97). Augustine, in his *De Rhetorica*, told us as much.

What *circumstance* is . . . can be understood more easily from its partition than from its definition. There are seven parts of circumstance . . . which Hermagoras calls *moria peristaseos* (particulars of circumstance) . . . They are these: Who? (*quis*), What? (*quid*), When? (*quando*), Where? (*ubi*), Why? (*cur*), In what manner? (*quem ad modum*), By what means? (*quibus adminiculis*). (Dieter & Kurth, 1968, p. 95)

Priority aside, by Augustine's times, the 5 Ws jingle (*quis, quid, quando, ubi, cur, quem ad modum, quibus adminiculis*) had already been popular enough to be discussed starting nearly a century earlier by Victorinus, Fortunatianus, Martianus Capella, Sulpicius Victor, and Iulius Victor. Later, we would still find the jingle in Boethius (who not only adopted the division between circumstances relating to the person and to the action but the Latin jingle: *Quis, quid, cur, quomodo, ubi, quando, quibus auxiliis* [who, what, why, how, where, when, with what], see Copeland & Suiter, 2009, p. 203) and in Alcuin (see Copeland & Sluiter, 2009, p. 291; Robertson, 1946, p. 12). Indeed, by the late middle ages the distinction of circumstances for persons and things (actions) and the very jingle (*quis, quid, ubi, quibus auxiliis, cur, quomodo, quando*), albeit not always in the same order, had become commonplace, as found in Matthew of Vendôme's (Matthaeus Vindocinensis) *Ars Versificatoria* (ca. 1175/1988) (in Copeland & Sluiter, 2009, p. 565; for the original Latin, see the Munari edition, 116, 1988, p. 128), in Thierry de Chartres and John of Salisbury (Robertson, 1946, pp. 12–13). Indeed, the seven circumstances came to occupy a central role in medieval culture, used even as an exegetic code to interpret texts (McKeon, 1942, p. 7; Copeland & Sluiter, 2009, p. 192–193). And when Canon 21 of the Fourth Lateran Council of 1215 made confession with their parish priest mandatory at least once a year for all faithful ("all the faithful of both sexes, *omnis utriusque sexus*"), manual of confessional instructions for both priests and laypeople gave guidelines for thinking of sins in terms of various renderings of the jingle "*quis, quid, ubi, quibus auxiliis, cur, quomodo, quando*" (or who, what, when, where, why, and how) (Robertson, 1946, p. 7).

When English rhetorician Thomas Wilson, in his 1560 *The Arte of Rhetorique*, one of the first rhetorical texts written in the English vernacular, translated into English the Latin jingle, he coined the 5 Ws + H formula that was to stick.

Seven circumstances, which are to be considered in divers matters.

- i. Who did the deed.
- ii. What was done.

- iii. Where it was done.
- iv. What help he had to do it.
- v. Wherefore he did it.
- vi. How he did it.
- vii. At what time he did it.

The circumstances in Meter.

Who, what, and where, by what help, and by whose,

Why, how and when, do many things disclose. (Wilson, 1560/1909, p. 17)

With Wilson's 5 Ws, we leave the field of rhetoric and fast forward some 500 years when the "tools of learning," rhetoric in particular, had long since been lost (Sayers, 1948).

The Linguist's Story

It is no doubt to Vladimir Propp and the Russian formalists of the early 20th century that we would have to turn for early modern work on narrative. Barthes (1966/1975), in his influential article "An Introduction to the Structural Analysis of Narrative," acknowledges that "thinkers as early as Aristotle . . . have concerned themselves with the study of narrative forms. . . ." (p. 237), but that said, he turns to the Russian formalists to set the stage of his work ("To confine myself to the current period, the Russian formalists. . . ." Barthes, 1966/1975, p. 238). In a seminal article published in 1925, Tomashevsky (1965) thought of a story "as a journey from one situation to another" (p. 70). And those story elements that contribute to significant changes in the journey are *dynamic motifs*, while the others are *static motifs* (Tomashevsky, 1965, p. 70). Dynamic motifs are typically actions, static ones are descriptions (e.g., of the traits of characters in the story). Tomashevsky also introduced a distinction between story and plot (*fabula* versus *sjužet*)³, describing them as two different ways of telling the same story, "in chronological and causal order of events. Plot is distinct from story. Both include the same events, but in the plot the events *are arranged* and connected according to the orderly sequence in which they were presented in the work" (Tomashevsky, 1965, p. 67), rather than in chronological order. Yet it was Propp's *Morphology of the Folktale*, first published in Russian in 1928 (and published in English in 1958 and 1968), that was to become one of the most cited works on narrative. Propp's novelty was to show that the number of consequential actions ("functions" as Propp refers to them) in folktales is limited and invariant (31 functions accounting for the

plot structure of all Russian folktales, later reduced to a set of six fundamental dichotomous invariant roles—"actants"—subject-object, sender-receiver, helper-opponent by Greimas, 1971: 798–799, 805). Furthermore, the sequence of actions is invariant. (The sequence of functions is always identical. Propp, 1968, p. 22).

Not surprisingly, the formalists' quest for invariant properties of narrative would find sympathetic ears among the French structuralists. From Bremond, to Benveniste, Todorov, and Barthes, all paid tribute to the Russian formalists and to Propp in particular. Bremond (1964) opened his "*Le message narratif*," one of the early French structuralist works on narrative, with the words "The name of Propp is beginning to be no longer an unknown in France" (p. 4). He then proceeded to discuss Propp's method in his attempt to find invariant structures in all narratives and not just Russian folktales. Todorov (1966) based his distinction between *histoire* and *discours* on Tomashevsky's story versus plot (p. 126–127). But while focusing on the invariant *macro*-structures that characterize narrative (in particular, the different sequential order of story and plot), Todorov (1996b) added an innovation of his own by focusing on the *micro*-structures of narrative—on the narrative clause. "The minimal schema of the plot can be shown naturally by a clause" (p. 74). The narrative clause is characterized by "two entities . . . a) the agents [that] correspond to proper nouns . . . serve[ing] as subject or object of the clause . . ." ("The appearance of an object depends upon the transitive or intransitive character of the verb," he would write in his *Grammaire du Décaméron*; 1969a, p. 28) and b) the predicate, which is always a verb." (Todorov, 1969b, p. 74) But these verbs, in narrative, "have a semantic characteristic in common: they denote an action" (Todorov, 1969b, p. 74). Todorov (1969a) would refer to this elementary Subject-Verb-[Object] structure as a "narrative grammar" (p. 27), the rhetorician's story of the 5 Ws having been reduced to 2: Who and What.

While the French structuralists were busy working on the invariant properties of narratives on the heels of the Russian formalists, on the other side of the Atlantic, a young sociolinguist, William Labov, was about to make his mark. In time, his work would become *the* most cited work on narrative. Contrary to the formalists and the structuralists, Labov searched for the "fundamental structures" of narrative not on folktales or literary fiction ("the products of expert storytellers") but "in oral versions of personal experiences," in "the actual narratives of . . . unsophisticated speakers" (Labov & Waletzky, 1967, p. 12). Contrary to the formalists and the structuralists, Labov was also primarily "concerned with the smallest unit of linguistic expression which defines the functions of narrative—primarily the clause," (Labov & Waletzky, 1967, p. 13) rather than with larger discursive units. And that "narrative clause . . . is one of the simplest grammatical patterns"

(Labov, 1972, p. 375). Basically, it is a structure based on “a series of eight elements . . . the first of these eight is the sentence adverbial [so that, then], the second the subject noun phrase [namely, in narrative, an actor], the third through eighth the verb phrase [for mode and instruments, location and time with the verb as preterit verb, i.e. an action in the past tense]” (Labov, 1972, p. 375–376). The narrative clause then, for Labov as well, is nothing but the 5 Ws + H.

The linguist’s story, then, meets the rhetorician’s story around the 5 Ws + H.

The Journalist’s Story

“So what’s the story?’ The night city editor asks. . . .” In a much-cited upbeat account of his experience as a journalist for *The New York Times*, Robert Darnton (1975) left no doubt about what writing news is all about (p. 178). The very title of his article, “Writing News and Telling Stories,” betrayed what Darnton had in mind. By the early 1900s, every news-writing handbook, a genre quickly becoming ever more popular, would tell its readers the same story: that news writing is nothing but story writing. Not that you need to attend journalism schools, where the professors responsible for the news-writing manuals work, to learn that. To quote Darnton (1975), “Although some reporters may learn to write in journalism schools . . . most of them (including many journalism-school graduates) pick up news writing in the course of an apprenticeship. . . . The copy boy internalizes the norms of the craft” (p. 186) by doing. And that involves learning how to tell a story. “Almost every article develops around a core conception of what constitutes ‘the story’” (Darnton, 1975, p. 180). These conceptions, of course, are not necessarily conscious. “We were blind to the archaic element in journalism. But our very conception of ‘news’ resulted from ancient ways of telling ‘stories.’” (Darnton, 1975, p. 191) Perhaps just as unconsciously, Darnton used the word *article* only a handful of times, but tens of times used the word *story*. Not surprisingly, in *The New York Times Manual of Style and Usage* from 1999, under story we read the following: “story. The preferred word for a newspaper or magazine report is ARTICLE(S)” (Siegal & Connolly, 1999, p. 320).

By the early 1900s, every news-writing handbook, written by the professors, insisted on the 5 Ws as the basic rule of good journalistic style. Writing in 1903, Shuman put it this way:

Every newspaper report should answer the questions, ‘What? Who? Where? When? Why?’ and should do it in the first paragraph as nearly as possible. This is the first and greatest commandment in the matter of journalistic style, and the penalty for breaking it is the waste-basket and swift oblivion. (p. 60)

A few years later, in his *The Writing of News: A Handbook*, Ross (1911) opened a section titled “What the Lead Should Contain” with the following familiar questions: “Who? What? When? Where? Why? It is a standard rule that the news lead should answer these questions about the story” (p. 59). Hyde (1912) put it no differently: “There are certain questions which arise in the reader’s mind. . . . The questions usually take the form of *when? where? what? who? how? why?*” (p. 38). For Bleyer (1913):

The beginning, or “lead,” of the story . . . should tell the reader the nature of the event, the persons or things concerned, as well as the time, the place, the cause, and the result. These essential points are given in answer to the questions: What? Who? When? Where? Why? How? (p. 66).

Even school children were being taught the formula in those days. As newspaper writing was starting to spread in high school curricula, Flint (1917) provided teachers with a ready-made outline of a possible course. For week seven, we find the following annotation:

Newspaper Writing. The news story. The five W’s: who, what, when, where, why. How their relative importance varies. Conditions that give pre-eminence to one element. . . . *Assignments*. Selection of examples of news stories illustrating the different ways of handling the five W’s. (Flint, 1917, p. 47)

Still, in 1941, editor Morton Sontheimer, in his autobiography, cited the *Scripps-Howard Style Book*, saying it was “about the best I have ever read, as far as in-house style books” (p. 48) and putting it this way: “A news story should tell: *Who, What, When, Where . . . and Why*” (p. 85).

That is perhaps why it is surprising to read, only a year later in 1942, Mott’s categorical conclusion that “The old-fashioned lead of the five W’s and the H, crystallized largely by Pulitzer’s ‘new journalism’ and sanctified by the schools, is widely giving way to the much more supple and interesting feature lead, even on straight news stories” (p. 65). The 5 Ws and H may be dead, but don’t rush to the funeral. The formula is still being repeated like a mantra in every news-writing manual, particularly in relation to the “lead” (e.g., Bliss & Patterson, 1978, p. 130; Hough, 1988, p. 58; Keeble, 1998, p. 106; Porter & Ferris, 1988, p. 94; Rich, 2010, p. 37).

Perhaps, old habits die hard. Perhaps more to the point, for as long as news writing is story writing, you will find there Kipling’s (1902) “six honest serving-men”: the 5 Ws + H. Not surprisingly, scholars studying news found that news does, indeed, tell stories.⁴ In Bell’s (1991) words, “Journalists do not write articles. They write stories. . . . Journalists are professional storytellers of our age” (p. 146). These scholars also found that news stories have

deep semantic macrostructures based on a hierarchy of events.⁵ They found that news stories have surface syntactic structures based on the 5 Ws, particularly in the lead.⁶

John King's Story

Of John King we know very little. The little we know comes from an unsigned newspaper story, like Darnton would have it, that appeared in *The Eastman Times-Journal* on July 22nd, 1909.⁷ A Negro. A farmer. He gave his name as John King before being lynched on Tuesday, July 20th, 1909, near Cadwell, Georgia. Like many like him, he came "to his death at the hands of unknown parties," as the coroner would report after his investigation. From the records of the 1900 Census, we know that John King was originally from North Carolina, where he had been born in August 1863. He could read and lived on a rented farm. He was married to Polly, also from North Carolina, and 14 years his younger. She could read and write. Together, they had eight children, only four surviving. All the children (two boys, John, 17, and Sweetie, 13, and two girls, Edith, 15, and Rosa, 10) were still living at home. None of them had had any schooling. Since John was born in Mississippi, the family must have arrived in Georgia from North Carolina through Mississippi; an unfortunate move as it turned out. Here is the story of John King's unfortunate final hours.

07/22/1909 *The Eastman Times-Journal*

NEGRO LYNCHED NEAR CADWELL TUESDAY AFTERNOON

John King Strung up to a Tree and Riddled With Bullets at the Hands of Unknown Parties

John King, a Negro said to be from Laurens county, was lynched near Little Ocmulgee church, about five miles from Eastman, Tuesday afternoon by an infuriated posse of citizens, who took the Negro from the officers, tied him to a tree and riddled his body with bullets.

The offence for which King met his death was insulting conduct toward Misses Estelle, Nora and Ruth Nicholson, highly respected daughters of Mr. Sam Nicholson of Eastman, on the public highway Tuesday

morning while these young ladies were en route to the home of Mr. Frank Holland who resides in the country.

When the young ladies were about two and one-half miles from town, they met King in the road and passed him, but noticed that...the Negro turned around and began to follow them. King...overtook the horse, grabbed him by the bridle and told the young ladies to get out; that the horse belonged to him, and that he was going to have it. The young ladies remained in the buggy and by their screams...the Negro became frightened and ran....

Deputy Sheriff George Steele, Messrs. Ben Taylor, Joe and John Peacock, Sam Rogers and a number of other citizens were soon on the Negro's trail, and as the chase progressed many others joined in until there were probably 150 or 200 men in pursuit of the criminal.

King was captured by Messrs. Geo. Steele, Ben Taylor, Sam Rogers, John Lord and Joe and John Peacock near Cadwell....The Negro acknowledged that he was the man who had disturbed the young ladies....He said his name was John King and that he had been working on a farm near Cadwell. When near Little Ocmulgee church the parties having King in charge were surrounded by a large posse of infuriated citizens, who took King from the buggy, carried him into the woods and riddled his body with bullets....

Coroner C.A. Armstrong held an inquest over the body Tuesday night, finding that he came to his death at the hands of unknown parties.

We find in this story the familiar 5 Ws—the *who* (the negro John King, the infuriated posse of citizens, the officers, Mr. Sam Nicholson, the “highly respected” Misses Estelle, Nora, and Ruth Nicholson), the *what* (lynch, string up, riddle with bullets), the *when* (07/22/1909, . . . Tuesday afternoon), the *where* (near Cadwell, . . . near little Ocmulgee church, about five miles from Eastman), the *why* (he claimed it was his horse, insulting conduct), and even the *how* (with bullets).

Quantitative Narrative Analysis (QNA)

In the early 1980s, in search of the actor in the study of social protest and violence, I started working on a new methodological approach focused on social actors rather than variables.⁸ Key to this approach was the idea of taking stories of conflict, coming mostly from newspaper articles⁹, and

turning their words into numbers. Indeed, for a time, several of my publications bore the title “From Words to Numbers” for the technique (Franzosi, 1989, 1994, 2004). Then, I settled on Quantitative Narrative Analysis for its easy acronym, QNA (Franzosi, 2010). QNA involves four steps:

1. Identify and assemble a set of narrative documents for analysis (e.g., newspaper articles of conflict and violence; but could also be police reports, personal narratives as told in in-depth interviews, web narratives, or fictional narratives).
2. Design a “story grammar” (i.e., a basic linguistic structure that would capture in a set of categories the information contained in narrative documents—namely, information on actors, their actions, and the attributes of actors and actions).
3. Design a computer program to store both the grammar and the coded information so as to make possible large-scale socio-historical projects (PC-ACE, Program for Computer-Assisted Coding of Events);
4. Analyze the data with statistical tools most appropriate for the narrative nature of the data (i.e., actors acting in time and space, rather than variables; for example, network models or GIS [Geographic Information System] tools, instead of regression models).

Step 1: Narratives

Newspapers are full of stories: thick descriptions of actors and actions, actors involved in crime, in politics, in social protest, in economic and financial transactions—because news writing is story writing, as Darnton told us. In my research, I have relied on newspapers as data sources for three different projects dealing with different substantive problems: industrial versus service sector strikes in Italy (1986–1987) (Franzosi, 1997); the rise of Italian fascism (1919–1922) (e.g., Franzosi, 1999, forthcoming); and lynching in Georgia (1875–1930). Let me focus here on my project on lynching and its gruesome stories.

Step 2: Story Grammar

Perhaps not surprisingly, given the nature of the documents I was working with (newspaper articles) and journalism’s 5 Ws + H, in my attempt to measure the role of social actors in historical processes, I stumbled upon a basic subject-verb-object (SVO) structure plus minimal modifiers (e.g., number of actors, time and space of the verb/action, reason). Knowing neither rhetoric nor linguistics, and having never heard of the 5 Ws of American journalism, I stumbled upon this structure by sheer brute force (or at least by imitation,

watching what others around me, Charles Tilly in particular, were doing with newspapers; indeed, “the new never strays too far from the old”) and luck (to be working with the only text genre with invariant properties: narrative, as captured by newspaper articles on conflict and violence; Franzosi, 2004, pp. 109–113). I would later discover that others who, like me, knew no rhetoric and no linguistics, had labeled this structure *story grammar*.¹⁰

But what is a story grammar? “Basically, a story grammar,” I write in *From Words to Numbers* (Franzosi, 2004), “is nothing but the simple linguistic structure subject-action-object or actor-action-actor with their respective modifiers (for example, number of actors involved, type of actor, time and space of action, reason, outcome)” (p. 5). Contrary to content analysis, the technique traditionally used to analyze texts quantitatively in the social sciences, the categories of a story grammar “are formally and explicitly related to one another throughout the coding scheme via a set of rewrite rules (e.g., subjects are linked to actions, actions to objects, and subjects, actions, and objects are linked to their modifiers).”¹¹ A rewrite rule, symbolized by a right-pointing arrow (→), takes any object of the grammar and rewrites it in terms of other, constituting objects (Franzosi, 2010, p. 23–24). Thus, if we call the *semantic triplet* the basic SVO structure, the following rewrite rule shows how this triplet can be rewritten:¹²

<semantic triplet> → {<subject>} {<verb>} [{<object>}]

In turn, the elements of the semantic triplet can be rewritten according to their attributes, down to their *terminal* symbols (those found in the language itself). For instance, <subject> could be rewritten as follows:

<subject> → <actor> [{ <attributes> }]
 <actor> → mob | Negro | sheriff | ...
 <attributes> → [{ <gender> }] [{ <race> }] [{ <organization> }]...
 <gender> → male | female |
 <race> → black | white | ...
 <organization> → police | federal authorities | ...

A verb and its attributes (“circumstances”)¹³ would instead look like this:

<verb> → <verbal phrase> [{<circumstances>}]
 <verbal phrase> → bring | burn | shoot | kill | hang | ...
 <circumstances> → [{<time>}] [{<space>}] [{<reason>}] [{<instrument>}]
 [{<outcome>}]

A story grammar, then, is nothing but the rhetoricians' 5 Ws + H. Even the names of the objects of the grammar betray these ancient roots. After all, Cicero, in his youthful work *De Inventione*, listed the attributes (*attributa*) of persons as name, nature, manner of life, fortune, habit, feeling, interests, purposes, achievements, accidents, and speeches made¹⁴ and the attributes of actions as place, time, occasion, manner, and facilities.¹⁵ Quintilian listed the attributes of persons as family, ethnicity/race, country, sex, age, education, physical aspect, fortune, marital status, character, occupation, accidents, speeches made, feelings/emotions, and name; and for the attributes of action, he listed reason, time, place, opportunity, instrument, and how.¹⁶ Sulpicius Victor later similarly talked about circumstances (*circumstantia*) of people as sex, nature, age, discipline, position [*fortuna*], education, and name (line 25, in Halm, 1863, p. 326). And so does Alcuin for whom the "attributes of persons are: name, nature, way of life, position, habitus, emotion, interests, plans, deeds, what has happened to someone [*casus*], words" (*De Inventione* I-24-34, Boethius 4, p. 203 in Copeland & Sluiter, 2009).¹⁷

Step 3: PC-ACE (Program for Computer-Assisted Coding of Events)

The relational nature of a story grammar makes it possible to implement a story grammar in a relational database management system (RDBMS). To this purpose, I developed a specialized software—PC-ACE (Program for Computer-Assisted Coding of Events)—that utilizes computer-assisted story grammars as the main tool to collect, organize, and store large bodies of narrative data.¹⁸ In fact, without computer software, there can be no substantive applications of a story grammar approach to narrative on a large scale (Franzosi, 2010, pp. 59–60). The sheer complexity and sophistication of coding schemes such as story grammars would limit their use to trivial, illustrative examples.¹⁹

Step 4: Data Analysis

The story of John King is one of 376 stories of lynching that occurred in the state of Georgia between 1875 and 1930, as reported in 1,332 articles from 212 different newspapers. Using a complex story grammar, implemented in PC-ACE, I coded these stories. "I" is a convenient euphemism: Silvia Girardi and Stefania Vicari did the data coding and verification. The result, based on 80% of the newspaper articles coded, was over 6,500 semantic triplets.

Table 4.1 shows a list of the most frequent actors appearing in the database.^{21 22} African American males are protagonists of the stories told by the

newspaper articles, together with (black and white) females and mobs. If we zoom in on the victims of lynching mobs, we find out that they were black young men, on average 26 years old (the youngest being 10 and the oldest 74), who typically worked as farm hands/cotton pickers or worked small farms of their own (rented or owned).

The results shown in Table 4.2 of the most frequent actions, perhaps not surprisingly, show that lynching events involve a great deal of violence (primarily against people) and coercion, of movement (going, searching, coming), and control. Acts of communication and facilitation/help are also frequent.

Table 4.1 Most frequent aggregated actors in the Lynching Database

Aggregated actor	Frequency ²⁰
Negro	3,550
Police	1,203
Mob	1,120
White women	726
White men	609

Table 4.2 Most frequent aggregated actions in the Lynching Database

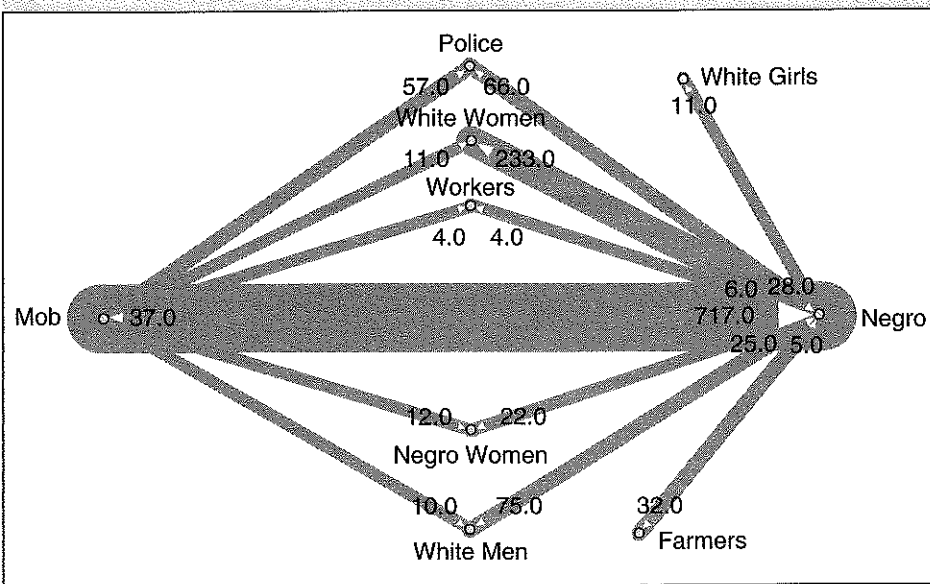
Aggregated Action	Frequency
Violence against people	1,248
Going	988
Force/coercion	620
Communication	369
Search	329
Control	286
Doing	271
Law	254
Sexual violence	137

(Continued)

Table 4.2 (Continued)

Aggregated Action	Frequency
Request	126
Assembling	116
Coming	99
Violence against things	99
Facilitation/help	95

Figure 4.1 Network graph of lynching violence (Georgia, 1875–1930)



We can intersect actor and actions to find out who did what (and to whom). Thus, which crimes did lynch victims commit? In 20% of the cases, they had committed (or allegedly committed) sexual assaults (“outraging” women—white women in 95% of the cases and typically very young, with the average age being 16, the youngest being 6 years old, and the oldest 80); some 40% had committed non-sexual, violent assaults; and in 13% of the cases, they had committed property crimes (e.g., arson, robbery). We can

apply network graphs to map visually the relations between social actors around particular spheres of action (e.g., violence, communication). In network graphs, the thickness of the line is roughly proportional to the number of actions of a given type (*relations* in network jargon, e.g., relations of violence) between any two social actors (*nodes*), and the arrows measure the direction of the relation (e.g., who is violent to whom). The graph of Figure 4.1 depicts one such graph centered on the sphere of action of violence. The graph makes clear that Negroes are at the receiving end of much violence, particularly by the mob (717 actions of violence by the mob against Negroes), but the Negroes themselves are agents of violence, in particular against women (white women and girls, with 233 and 11 actions respectively, but also black women in 22 cases). The graph also shows that the police (conveniently aggregating here such figures as the sheriff, deputy-sheriff, and other law-enforcing agents) are often victims of violence by both the mob and African Americans.

The results presented here, however exploratory and illustrative, do show that QNA, as an approach to text involving computer-assisted story grammars and network models (and other tools of analysis based on the narrative properties of text, such as sequence analysis or GIS tools) provides a way to measure social actors and their actions for socio-historical research.

Epilogue

“The conclusions, called epilogues by the Greeks, have three parts: summary, amplification, and compassion.” Thus we read in the *Rhetorica ad Herennium* (II, 47). I will spare you compassion—if not for the soul of John King, an African American lynched on July 22, 1909, in Georgia, guilty of having frightened three highly respectable white young ladies on a public road (after all, as the anonymous writer of the *Rhetorica* reminds us, “nothing dries faster than tears” *Ibid.* II, 50)—and the commonplaces (*loci communes*) of amplification used to stir feelings. I will focus here on the summary.

At the heart of this chapter is a methodological innovation that I have developed in the analysis of narrative texts: Quantitative Narrative Analysis (QNA). QNA is based on the use of a story grammar in a computer-assisted environment (PC-ACE, Program for Computer-Assisted Coding of Events). I show how computer-assisted story grammars are not just methodological toys with desirable properties but produce innovative substantive results (sociologically and historically) when applied to hundreds or even thousands of newspaper stories (e.g., in my projects on lynching in the Jim Crow South, 1875–1930, or on the rise of Italian fascism, 1919–1922). Story grammars

are nothing but the 5 Ws + H of journalism (who, what, when, where, why, and how). I trace this construct back to classical rhetoric, to Hermagoras, Cicero, Quintilian, and the theory of *stasis* and circumstances, to the medieval and Renaissance rhetoricians who gave it the Latin jingle “*quis, quid, cur, quomodo, ubi, quando, quibus auxiliis*” (in my favorite order) and, later, to the Renaissance rhetoricians who translated the Latin jingle in its vernacular form of the 5 Ws and H (who, what, when, where, why, and how). When 20th century linguists and literary critics started working on narrative, rhetoric had long since ceased to be part of school curricula. And to some extent, they re-invented the wheel. But they also brought in a trickle of innovation: the role of negatives and modality in slanting a narrative (not to mention passivization and nominalization), the relationship between narrator and narratee, point of view, and sequences (story and plot).

And in this story, where is *my* innovation? Not in the use of a story grammar (the 5 Ws + H structure) known to rhetoric, from classical to medieval and renaissance times. I was lucky to be using newspaper articles as sources of socio-historical data and that these texts basically tell stories (real stories). Not surprisingly, I found there the 5 Ws + H; after all, these types of texts are based upon this structure. My contribution (the novelty in my story) lies elsewhere:

1. In taking this simple 5 Ws + H structure (a story grammar) and finding an appropriate computer data model for this structure (namely, relational database management systems [RDBMS]).
2. In developing a computer program (PC-ACE [Program for Computer-Assisted Coding of Events]) that would allow me to apply this methodological technique in large-scale socio-historical problems.
3. In using PC-ACE to store information on the 5 Ws + H from thousands of newspaper stories on lynching events in Georgia (1875–1930) (and on the rise of Italian fascism, 1919–1922).
4. In applying statistical techniques of data analysis isomorphous to the narrative nature of the data (i.e., actors and actions, sequences, time and place; e.g., network models that graphically map the relationships among social actors, GIS, tools for the spatio-temporal mapping of social actions).

Yet not even here can I really claim innovation. Neither one of those two elements of my story (computer data models and statistical models) are *mine*. They were there for me to use to my advantage. Which brings my story to a sobering end: the limits to innovation (at least *my* innovation). All I did was put *together* for *new* purposes things already *separately* out there (that would be using metaphor, the rhetoricians would tell us). If the point

of storytelling is, ultimately, that of persuasion (and in this, sentiments and emotions play a part),²³ I do hope to have persuaded you (perhaps even through the emotional appeal of a false-modest characterization of the hero of the story) of the following:

1. QNA represents a powerful approach for the analysis of narrative texts.
2. The main ingredients of QNA (the 5 Ws + H structure and their sequential organization) have been around for centuries.
3. The new never strays too far from the old.

And that is the end of this story.

Acknowledgments

The project on lynching in Georgia (1875–1930) was made possible by funding from the University Research Committee at Emory University. Woody Beck kindly provided the cross-references between each lynching event and newspaper articles. Justin Clayton did most of the work to locate the articles and convert them into pdf files. Silvia Girardi and Stefania Vicari entered and verified the data in PC-ACE, and Stefania Vicari and Gianluca De Fazio did some of the analyses. Stewart Tolnay provided the census information on John King from his University of Washington Lynching-Census Record Linkage Project. James Holstein and Jaber Gubrium offered very helpful comments on a first draft of this chapter.

References

- Barthes, R. (1966/1975). An introduction to the structural analysis of narrative [Special issue]. *New Literary History*, 6(2), 237–272.
- Bell, A. (1991). *The language of news media*. Oxford: Blackwell.
- Bell, A. (1998). The discourse structure of news stories. In A. Bell & P. Garrett (Eds.), *Approaches to media discourse* (pp. 64–104). Oxford: Blackwell.
- Bleyer, W. G. (1913). *Newspaper writing and editing*. Cambridge, MA: The Riverside Press.
- Bliss, E., Jr., & Patterson, J. M. (1978). *Writing news for broadcast* (2nd ed.). New York: Columbia University Press.
- Bremond, C. (1964). Le message narratif. *Communications*, 4, 4–32.
- Chatman, S. (1978). *Story and discourse. Narrative structure in fiction and film*. Ithaca, NY: Cornell University Press.

- Clark, D. L. (1952). The rise and fall of Progymnasmata in sixteenth and seventeenth century grammar schools. *Communication Monographs*, 19(4), 259–263.
- Conboy, M. (2007). *The language of the news*. London: Routledge.
- Copeland, R., & Sluiter, I. (Eds.). (2009). *Medieval grammar and rhetoric. Language arts and literary theory, AD 300–1475*. Oxford: Oxford University Press.
- Cotter, C. (2010). *News talk. Investigating the language of journalism*. Cambridge: Cambridge University Press.
- Darnton, R. (1975). Writing news and telling stories. *Daedalus*, 104(2), 175–194.
- Dijk, T. V. (1983). Discourse analysis: Its development and application to the structure of news. *Journal of Communication*, 33, 20–43.
- Dijk, T. V. (1986). News schemata. In C. Cooper & S. Greenbaum (Eds.), *Studying writing: Linguistic approaches* (pp. 155–186). Beverly Hills: Sage.
- Dieter, O. A. L., & Kurth, W. C. (1968). The *de rhetorica* of Aurelius Augustine. *Speech monographs* (now *Communication monographs*), 35(1), 90–108.
- Flint, N. L. (1917). *Newspaper writing in high schools: Containing an outline for the use of teachers*. Lawrence: University of Kansas Press.
- Franzosi, R. (1987). The press as a source of socio-historical data. *Historical Methods*, 20, 5–16.
- Franzosi, R. (1989). From words to numbers: A generalized and linguistics-based coding procedure for collecting event-data from newspapers. In C. Clogg (Ed.), *Sociological methodology* (Vol. 19, pp. 263–298). Oxford: Basil Blackwell.
- Franzosi, R. (1994). From words to numbers: A set theory framework for the collection, organization and analysis of narrative data. In P. Marsden (Ed.), *Sociological methodology* (Vol. 24, pp. 105–136). Oxford: Basil Blackwell.
- Franzosi, R. (1996). A sociologist meets history: Critical reflections upon practice. *Journal of Historical Sociology*, 9(3), 354–392.
- Franzosi, R. (1997). Labor unrest in the Italian service sector: An application of semantic grammars. In C. W. Roberts (Ed.), *Text analysis for the social sciences: Methods for drawing statistical inferences from texts and transcripts*. Hillsdale, NJ: Lawrence Erlbaum.
- Franzosi, R. (1998). Narrative analysis—Why (and how) sociologists should be interested in narrative. In J. Hagan (Ed.), *The annual review of sociology* (pp. 517–554). Palo Alto, CA: Annual Reviews.
- Franzosi, R. (1999). The return of the actor. Networks of interactions among social actors during periods of high mobilization (Italy, 1919–22). *Mobilization*, 4(2), 131–149.
- Franzosi, R. (2004). *From words to numbers. Narrative, data, and social science*. Cambridge: Cambridge University Press.
- Franzosi, R. (2008). *Content analysis* (Benchmarks in *Social Research Methods* series). 4 vols. Thousand Oaks, CA: Sage.
- Franzosi, R. (2010). *Quantitative narrative analysis*. Los Angeles: Sage.
- Greimas, A. (1971). Narrative grammar: Units and levels. *Modern Language Notes*, 86, 7793–7806.
- Halliday, M. A. K. (1985/1994). *An introduction to functional grammar*. London: Arnold.
- Halm, K. (Ed.). (1863). *Rhetores Latini minores. Ex codicibus maximam partem primum adhibitis emendabat Carolus Halm*. Leipzig: Teubner.
- Hough, G. A. (1988). *News writing* (4th ed.). Boston: Houghton Mifflin.
- Hyde, G. M. (1912). *Newspaper reporting and correspondence: A manual for reporters, correspondents, and students of newspaper writing*. New York: D. Appleton.
- Keeble, R. (1998). *The newspapers handbook* (2nd ed.). London: Routledge.
- Kennedy, G. (1994). *A new history of classical rhetoric*. Princeton, NJ: Princeton University Press.
- Kipling, R. (1902). The elephant's child. In R. Kipling, *Just so stories for little children*. London: Macmillan.
- Labov, W. (1972). *Language in the inner city*. Philadelphia: University of Pennsylvania Press.
- Labov, W., & Waletzky, J. (1967). Narrative analysis: Oral versions of personal experience. In J. Helm (Ed.), *Essays on the verbal and visual arts* (pp. 12–44). Seattle: University of Washington Press. Reprinted in *Journal of Narrative and Life History*, 1997, 7(1), 1–38.
- Manoff, R. K. (1987). Writing the news (by telling the “story”). In R. K. Manoff & M. Schudson (Eds.), *Reading the news* (pp. 197–229). New York: Pantheon.
- Matthew of Vendôme's (Matthaeus Vindocinensis). (ca. 1175/1988). *Opera. Vol. III. Ars Versificatoria* (F. Munari, Ed.). Rome: Edizioni di storia e letteratura.
- McKeon, R. (1942). Rhetoric in the Middle Ages. *Speculum. A Journal of Medieval Studies*, 17(1), 1–32.
- Mott, F. L. (1942). Trends in newspaper content. *Annals of the American Academy of Political and Social Science*, 219, 60–65.
- Porter, B., & Ferris, T. (1988). *The practice of journalism. A guide to reporting and writing the news*. Englewood Cliffs, NJ: Prentice Hall.
- Prince, G. (1973). A grammar of stories: An introduction. *De proprietatibus litterarum; Series minor* (Vol. 13). Paris: Mouton.
- Propp, V. (1968). *Morphology of the folktale* (2nd ed.). Austin: University of Texas Press.
- Reah, D. (2002). *The language of newspapers* (2nd ed.). London: Routledge.
- Rich, C. (2010). *Writing and reporting news: A coaching method* (6th ed.). Boston: Wadsworth.
- Robertson, D. W., Jr. (1946). A note on the classical origin of “circumstances” in the medieval confessional. *Studies in Philology*, 43(1), 6–14.
- Ross, C. G. (1911). *The writing of news: A handbook with chapters on newspaper correspondence and copy reading*. New York: Henry Holt.
- Sayers, D. (1948). The lost tools of learning. *The Hibbert Journal. A Quarterly Review of Religion, Theology, and Philosophy*, XLVI(1), 1–13.

- Shuman, E. L. (1903). *Practical journalism. A complete manual of the best newspaper methods*. New York: D. Appleton and Company.
- Siegal, A. M., & Connolly, W. G. (1999). *The New York Times manual of style and usage*. New York: Three Rivers Press.
- Sonthheimer, M. (1941). *Newspaperman: A book about the business*. New York: Whittlesey House.
- Steiner, P. (1984). *Russian formalism: A metapoetics*. Ithaca, NY: Cornell University Press.
- Todorov, T. (1966). Les catégories du récit littéraire. *Communications*, 8, 125–151.
- Todorov, T. (1969a). *Grammaire du décameron*. The Hague: Mouton.
- Todorov, T. (1969b). Structural analysis of narrative. *NOVEL: A forum on fiction*, 3(1), 70–76.
- Tomashevsky, B. (1965). Thematics. In L. T. Lemon & M. J. Reis (Eds.), *Russian formalist criticism* (pp. 61–95). *Four essays translated and with an introduction by Lee T. Lemon and Marion J. Reis*. Lincoln: University of Nebraska Press.
- Vollgraff, W. (1948). Observations sur le sixième discours d'Antiphon. *Mnemosyne*, 1(4), 257–270.
- Wilson, T. (1560/1909). *The art of rhetoric* (C. H. Mair, Ed.). Oxford: Clarendon Press.
- Wolf, M. (1992). *A thrice-told tale: Feminism, postmodernism, and ethnographic responsibility*. Stanford: Stanford University Press.

Notes

1. For a history of the concept, see Robertson (1946) and Vollgraff (1948, pp. 265–267).
2. Robertson (1946, pp. 9–10). Discussion of circumstances is also found in the *Rhetorica ad Herennium*.
3. On this terminology, and on the role of Tomashevsky and Šklovskij, the self-declared “founder of the Russian school of formal method,” see Steiner (1984, pp. 44–52).
4. Bell (1991, pp. 146–174); Cotter (2010, pp. 135–170); Manoff (1987).
5. See Bell (1991, p. 167; 1998, p. 73); van Dijk (1983, 1986).
6. See Bell (1991, pp. 175, 182, 190–202; 1998, pp. 69–88); Conboy (2007, p. 51); Cotter (2010, pp. 161, 165); Manoff (1987); Reah (2002, pp. 26–28).
7. We also have two short, identical articles published on July 21, 1909, by *The Atlanta Constitution* and, the day after, on July 22, 1909, by the *Montgomery Monitor*. *The Atlanta Constitution* gives the name of the victim as John Green and the place of lynching as Eastman in Dodge County, 13 miles away from Cadwell, in Laurens County. All other information coincides. However, there is no census record for a John Green, while there is one, with a close match, to John King.
8. For that story, I invite the curious reader to read my “A sociologist meets history” (Franzosi, 1996) and the Acknowledgments to my *From Words to Numbers* (Franzosi, 2004).

9. There is a long tradition of using newspaper articles as sources of socio-historical data, particularly in the study of social protest and conflict, revolutions, and social movements (see, for all, Franzosi, 1987, 2004, pp. 167–173).

10. It was mostly psychologists who referred to the 5 Ws + H structure as *story grammar*. Todorov (1968) referred to it as “grammar of stories” and so did Prince (1973) in his work on stories. For the story of the concept of story grammar, see Franzosi (2004, pp. 41–55).

11. Franzosi (2010, p. 35). On content analysis, see Franzosi (2008); for a comparison of content analysis and quantitative narrative analysis, see Franzosi (2004, pp. 59–61, 91–92; 2010, pp. 34–36).

12. The angular brackets <> denote elements that can be further rewritten; while *terminal elements* (i.e., the words or linguistic expressions found in the text) have no <>. Curly brackets {} denote elements that can occur more than one time; while square brackets [] denote optional elements. Thus, in the clause *victim screams* there is only one participant (the agent), while the clause *mob kills Negro* has two participants (the agent, mob, and the recipient or patient, Negro). As a result, the grammar requires only the first participant; the second is optional.

13. Halliday (1985/1994), in his *An Introduction to Functional Grammar*, refers to the circumstances of a process as “the location of an event in time or space, its manner, or its cause” (p. 150). Circumstances refer to such notions as “when, where, how and why” something happens (p. 150). The term *circumstances*, then, for Halliday, has a more restricted meaning than for the rhetorician for whom circumstances included the who and the what.

14. *Nomen, naturam, victum, fortunam, habitum, affectionem, studia, consilia, facta, casus, rationes* (I, xxiv, 34).

15. *Locus, tempus, modus, occasio, facultas* (I, xxvi, 38).

16. *Causa, tempus, locus, occasio, instrumentum, modus* and *ubi, quando, quo modo, per quae facta* (Quintilian, *Institutio oratoria*, V, 10, 24–30 and V, 10, 32).

17. Modern narrative theorists have argued that the personal qualities of a character, or *traits*, are typically, although not exclusively, expressed grammatically as adjectives (Chatman, 1978, p. 125; Todorov, 1969a, p. 31).

18. The latest release of PC-ACE is available in the public domain for free download at www.pc-ace.com.

19. For a review of the desirable properties of computer-assisted story grammars see, for all, Franzosi (2010, pp. 33–41).

20. All frequencies are measured as number of triplets in which the objects (e.g., actors, action) appear.

21. For the purpose of data analysis, both actors and actions have been aggregated into larger categories (on these issues, see Franzosi, 2004, pp. 293; Franzosi, 2010, pp. 103–104). Thus, the aggregated action “violence against people” includes such verbal phrases as *kill, wound, hang, rape, riddle with bullets, torture*, and so on.

22. We use PC-ACE Query Manager (a GUI tool—Graphical User Interface—to help the user) to extract information from a relational database in general ways that

convert words into frequency distributions (i.e., numbers). These numbers can be analyzed using standard multivariate statistical techniques (namely, factor analysis, Franzosi, 2004, 113–115; regression models, Franzosi, 2004, pp. 115–117; or logistic models, Franzosi, 1994). However, the nature of the data collected via a computerized story grammar (PC-ACE) lends itself to a variety of novel approaches to the analysis of qualitative data: network analysis, sequence analysis, and Geographic Information Systems. All these approaches take advantage of the underlying narrative structure of the data. In particular, there is a homologous relationship between story grammars (with subjects/actors related to objects/actors via a given action; “mob burns negro”) and network models (Franzosi, 1998, 1999, 2004). Network models graphically map the relationships among actors via directed graphs. It is an ideal tool to map the network of social actors involved in lynching and their reciprocal roles.

23. On persuasion and the role of sentiments and emotions, see Quintilian (*Institutio oratoria*, IV, 2, 21, IV, 2, 31 and IV, 2, 111).

PART II

Analyzing Storytelling

Varieties *of* Narrative Analysis

James A. Holstein

Marquette University

Jaber F. Gubrium

University of Missouri

Editors



Los Angeles | London | New Delhi
Singapore | Washington DC



Los Angeles | London | New Delhi
Singapore | Washington DC

FOR INFORMATION:

SAGE Publications, Inc.
2455 Teller Road
Thousand Oaks, California 91320
E-mail: order@sagepub.com

SAGE Publications Ltd.
1 Oliver's Yard
55 City Road
London EC1Y 1SP
United Kingdom

SAGE Publications India Pvt. Ltd.
B 1/1 Mohan Cooperative Industrial Area
Mathura Road, New Delhi 110 044
India

SAGE Publications Asia-Pacific Pte. Ltd.
33 Pekin Street #02-01
Far East Square
Singapore 048763

Acquisitions Editor: Michele Sordi
Associate Editor: Lauren Habib
Editorial Assistant: Kalle Koscielak
Production Editor: Eric Garner
Copy Editor: Jenifer Dill
Typesetter: C&M Digital (P) Ltd.
Proofreader: Laura Webb
Indexer: Judy Hunt
Cover Designer: Candice Harman
Marketing Manager: Helen Salmon
Permissions Editor: Adele Hutchinson

Copyright © 2012 by SAGE Publications, Inc.

All rights reserved. No part of this book may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without permission in writing from the publisher.

Printed in the United States of America

Library of Congress Cataloging-in-Publication Data

Varieties of narrative analysis / James A. Holstein, Jaber F. Gubrium, editors.

p. cm.
Includes bibliographical references and index.

ISBN 978-1-4129-8755-4 (cloth : alk. paper)

1. Narrative analysis (Research method)
I. Holstein, James A. II. Gubrium, Jaber F.

H61.295.V37 2012
001.4'33—dc22 2011011900

Contents

Acknowledgments	xi
About the Editors	xiii
Introduction: Establishing a Balance	1
Reemergent Interest	2
Adjusting the Balance	3
Theoretical Sensibilities and Analytic Strategies	5
References	10
PART I: Analyzing Stories	
1. Exploring Psychological Themes Through Life-Narrative Accounts	15
<i>Dan P. McAdams</i>	
Narrative in the Context of Discovery	17
Narrative in the Context of Justification	20
Using Narratives to Test Extant Theories	25
Conclusion	29
References	30
2. Practicing Dialogical Narrative Analysis	33
<i>Arthur W. Frank</i>	
Five Commitments	34
Doing Dialogical Narrative Analysis	37
Animating Interest	37
Varieties of Fieldwork	38
Collecting Stories	40
Deciding What Is a Story	41
Selecting Stories for Analysis: Practicing <i>Phronesis</i>	43
Opening Up Analysis	44
Building a Typology	46