

Sociology, narrative, and the quality versus quantity debate (Goethe versus Newton): Can computer-assisted story grammars help us understand the rise of Italian fascism (1919–1922)?

Roberto P. Franzosi

Published online: 14 October 2010
© Springer Science+Business Media B.V. 2010

Abstract At the heart of this article is a structural approach to narrative, based on the work of Propp, Labov, van Dijk, Halliday, and others. The article highlights the structural features of narrative—basically, the organization of the genre around the semantic template actor-action-actor (syntactically rendered as subject-verb-object but where, in narrative, the subject is typically a social actor and the verb a social action) and the modifiers of each element of this triplet, such as time and space of action—and shows how to implement this structure in a computer environment and how to use this methodological tool in socio-historical research (namely, the rise of Italian fascism, 1919–22). But, taking a cue from Halliday’s cover jacket of his *An Introduction to Functional Grammar*, with its representation of a color circle, the paper takes the reader on an intellectual journey from Newton to Goethe—and the quality versus quantity debate—to Goethe and Propp, to end, back home, with Simmel and Weber.

Keywords Narrative · Story grammar · Quality vs. quantity · Italian fascism

Every act of looking turns into observation, every act of observation into reflection, every act of reflection into the making of associations; thus it is evident that we theorize every time we look carefully at the world.

(von Goethe 1996, p. 90)

I, for my part, drawn in many directions as I am, cannot content myself with one way of thinking.

(von Goethe 1996, p. 46)

R. P. Franzosi (✉)

Department of Sociology/ Program in Linguistics, Emory University, 1555 Dickey Dr., Atlanta, GA 30322, USA
e-mail: rfranzo@emory.edu

The 1994 second edition of M.A.K. Halliday's *An Introduction to Functional Grammar* is characteristic for its display of a colorful frontispiece: a circle made up of two concentric rings, the outer one divided into slices of different colors (see [Appendix](#)). Words appear outside the bigger circle and inside the slices (e.g., relational, verbal, existential, being, sensing, doing).

There is nothing unusual about the use of a pictorial frontispiece as a way of visually conveying the core meaning of a book. It is a tradition that dates back to the beginning of book printing (see, Corbett and Lightbown [1979](#)). Some frontispieces, like Hobbes's powerful image of the Commonwealth monster in his 1651 *Leviathan* or Stefano della Bella's engraving of Aristotle, Ptolemy, and Copernicus in the frontispiece to Galileo's *Dialogue Concerning the Two Chief World Systems* (1632) have survived into our days, with the immediacy of their iconographic message.¹ Others, once popular, have followed the oblivion of the works to which they are attached, the meaning of their complex symbolism lost (e.g., the frontispiece of Diderot's and d'Alembert's *Encyclopédie*, 1772, with the central figure representing Truth, in a bright light—the central symbol of the enlightenment—and with its veil removed by Reason and Philosophy).²

There is also nothing unusual about Halliday's use of the "color circle," with blue, red, and yellow occupying the larger slices and violet, orange, and green occupying smaller, intermediate slices. By the twentieth century that way of representing colors in a circle had become standard (see Gage [1993](#), pp. 164–176). It was Newton who gave us, in his *Opticks* (1702), one of the first circular representations of this kind based on seven colors, followed by Claude Boutet's circle at around the same time, soon to become focused on six colors only. The division of these basic six colors into three primaries (blue, red, and yellow) and three secondaries (wholly derived from a combination of primaries; for example, green out of mixing blue and yellow) had also become standard. Writing in 1725, only a few years after the publication of the *Opticks*, J.C. Le Blon had stated: "Painting can represent all *visible* Objects with three Colours, *Yellow, Red* and *Blue*; for all other Colours can be compos'd of these *Three*, which I call *Primitive*" (cited in Gage [1993](#), p. 169). With Goethe, who left behind several works dedicated to the scientific study of colors, including his *Theory of Colours*, the color circle (or "wheel," in the Master's words) based on six colors became a standard representation (Matthaei [1971](#), pp. 41–43, 50–56). Finally, there is perhaps nothing unusual about Halliday's linking of colors to language and verbal representations. Since ancient times, the dominant four colors had been variously associated to the four elements (air, water, fire, and earth), to the four times of day, the four seasons, the four phases of the moon, the four ages of man, the four humors of the body (linked to different temperaments: optimistic, melancholic, phlegmatic, and choleric), and even to different sounds (Gage [1993](#), pp. 204, 227–246).

But certainly new is Halliday's association of various colors to different linguistic and verbal representations. Halliday writes ([1994](#), p. 106):

Language enables human beings to build a mental picture of reality, to make sense of what goes on around them and inside them. Here again the clause

¹ For a discussion of the frontispiece of the *Leviathan*, see Corbett and Lightbown ([1979](#), pp. 219–230).

² The iconographic representation of some frontispieces became, with time, so complex, as to require an accompanying explanatory text (Corbett and Lightbown [1979](#), p. 46–47).

plays a central role, because it embodies a general principle for modelling experience—namely, the principle that experience is made up of PROCESS. Our most powerful impression of experience is that it consists of ‘goings-on’—happening, doing, sensing, meaning, and being and becoming. All these goings-on are sorted out in the grammar of the clause. Thus as well as being a mode of action, of giving and demanding goods-&-services and information, the clause is also a mode of reflection, of imposing order on the endless variation and flow of events. The grammatical system by which this is achieved is TRANSITIVITY. The transitivity system construes the world of experience into a manageable set of PROCESS TYPES.

For Halliday, human beings experience their inner and outer world and represent these experiences linguistically as processes in the clause, with three primary types of processes—doing (or material processes, further divided into happening, creating/ changing, and doing (to)/acting), sensing (or mental, further divided into seeing, feeling, and thinking), and being (or relational, further divided into symbolizing, having identity, and having attribute), corresponding to the three primary colors of red, blue, and yellow—and three more types of secondary processes (verbal, existential, and behavioural, located at the boundaries between different combinations of primary processes “sharing some features of each, and thus acquiring a character of their own”—verbal between mental and relational, existential between relational and material, and behavioural between material and mental). Halliday continues:

There is no priority of one kind of process over another. But they are ordered; and what is important is that, in our concrete visual metaphor, they form a circle and not a line. (More accurately still, they could be shown to form a sphere ...).³

For Halliday, then, the frontispiece is not just a colorful cover for his book, with no or little relationship to its content; on the contrary, the frontispiece represents metaphorically his scientific understanding of language. “The grammar construes experience like a colour chart, with red, blue and yellow as primary colours and purple, green and orange along the borders ...” (Halliday 1994, p. 107).

Processes and text genres

Doing, Being, and Sensing. Those, then, are the three fundamental categories of processes for Halliday. The question is: Are different types of text genres characterized by different combinations of the fundamental processes (or of their subcategories)? Or, to put it differently, do certain types of texts preferentially exhibit certain types of processes? Halliday does not entertain that question. But Chatman writes: “[Narrative] statements are of two kinds—process and stasis—according to whether someone did something or something happened; or whether something simply existed in the story. Process statements are in the mode of DO or HAPPEN ... Stasis statements are in the

³ Linear had indeed been the representations of colors until the late seventeenth century, in a tradition dating back to Aristotle. But spherical representations were not absent, particularly in modern times (e.g., the 1611 color sphere of the Swedish mathematician Sigfrid Forsius or Runge’s later color sphere of 1810 (see Gage 1993, pp. 166–167, 194).

mode of IS.” Narrative, in other words, is characterized by verbs of doing rather than being. Toolan (1988, p. 157) pushes that distinction further:

what is said ... will not be the core of a story; that, rather, what is done ... will be. The “what is done” then becomes (or may become) the core narrative text of clauses—actions—while the “what is said” becomes evaluative commentary on those actions.

Narrative texts are characterized not only by processes of doing rather than being, but also by processes of doing rather than sensing (or of acting rather than saying, to use subcategories of doing and sensing). Indeed, this “doing versus saying” distinction, implicitly or explicitly, is at the core of linguistic theories of narrative structures. “In general, narrative theorists rather tend to analyse the course of action to which they limit their story” (Bal 1977, p. 89). “There is no structural analysis of narrative that does not borrow from an explicit or implicit phenomenology of ‘doing something.’” (Ricoeur 1984, p. 56; see also Genette 1980, pp. 164, 169). Not all sentences of a narrative are about action and doing. “Non-story elements may be found in a narrative text just as story elements may be found in a nonnarrative text.” (Rimmon-Kenan 1983, p. 15). Descriptive, expository, and evaluative clauses are typically part of an even minimal narrative. There is never either a pure narrative or nonnarrative text. Narrative texts are those where “the distinctive traits of the narrative genre ... are quantitatively predominant” (Bal 1977, p. 13; also pp. 89–90).

Narrative and sequences (and ways of seeing)

Narrative texts thus seem to be characterized by processes of doing rather than being or sensing. Yet, is an emphasis on doing, on van Dijk’s discourse of action (van Dijk 1980, p. 13) or Genette’s narrative of events (Genette 1980, pp. 164, 169), enough to produce a narrative text?

For Labov (1972 pp. 360–361, emphasis added), “a minimal narrative ... [is] a *sequence* of two clauses which are temporally ordered.... The skeleton of a narrative ... consists of a series of temporally ordered clauses.” “Narrative [is] one method of recapitulating past experience by matching a verbal *sequence* of clauses to the *sequence* of events which (it is inferred) actually occurred.” (Labov 1972, p. 1, added; see also Labov and Waletzky 1967) Indeed, time and sequence (or the chrono-logical ordering of events, i.e., of clauses characterized by processes of action) are fundamental elements in this definition of narrative (and, implicitly, of narrative clauses as characterized by processes of doing).

This focus on time and sequence goes back to the Russian formalists of the beginning of the twentieth century (e.g., Propp and Tomashevski) with their distinction between narrative story and plot (*fabula* versus *sjužet*). “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.” (Tomashevski 1965, p. 67) “The sequence of events has its own laws—Propp wrote (1958, p. 22; emphasis in original)—Theft cannot take place before the door is forced. Insofar as the tale is concerned it has its own entirely particular and specific laws. The sequence of elements ... is strictly *uniform*.”

The French structuralists adopted the Russian formalists' basic distinction of *fabula* versus *sjužet* as *histoire* versus *discours*, later rendered in English as story versus discourse (see Benveniste 1971, pp. 206–208; Barthes 1977; Chatman 1978, p. 19; Genette 1980, pp. 25–27, Toolan 1988, pp. 11–12). A story, then refers to a skeletal description of the fundamental events in their natural logical and chronological order or sequence (perhaps, with an equally skeletal listing of the roles of the characters in the story) (Prince 1973, p. 23; Bal 1977, pp. 4, 7; Todorov 1977, p. 111; Rimmon-Kenan 1983, p. 19; Cohan and Shires 1988, pp. 53–54; Toolan 1988, p. 9).

Chronological sequence is a crucial ingredient of any definition of story. And without story there is no narrative (Rimmon-Kenan 1983, p. 15). Yet, temporal ordering is a necessary but not sufficient condition for a story (Rimmon-Kenan 1983, p. 19). The events in the sequence must also: 1. disrupt an initial state of equilibrium setting in motion an inversion of situation, a change of fortunes; and 2. be coherently connected. Not all events of a sequence equally contribute to a change in situation (Tomashevski 1965, p. 70; Prince 1973, p. 28; Todorov 1990, p. 30). Some events are more consequential than others—referred to as dynamic motifs, rather than static ones, by Tomashevski (1965[1925], p. 70), cardinal functions (or nuclei) and catalysers by Barthes (1977[1966], pp. 93–94), kernel and satellite events by Chatman (1978, pp. 32, 53–56) and by Rimmon-Kenan (1983, p. 16). In simple narratives, with minimum plot development, the sequential connection of basic narrative units to one another occurs mainly through temporal order (minimum plot, maximum story). In more complex narratives (maximum plot), those same basic narrative units are connected in a variety of ways through enchainment, embedding, and joining, rather than simply through temporal ordering (Bremond 1964, 1966; Todorov 1981, pp. 52–53; Rimmon-Kenan 1983, p. 23).

Whether sequential or consequential, the events must be narrated coherently; the narrative units must be coherently stitched together (on semantic coherence, see Agar and Hobbs 1982; van Dijk 1983, p. 25). And that coherence occurs at different levels: local (when each sentence in a text is meaningfully well formed and meaningfully related to its adjacent sentences) and global (when the sentences in a text provide a coherent whole). Coherence can also be thematic (when each sentence or sequence of sentences in a text is meaningfully related to one or more dominant theme or topic appearing in the text).

And simplex or complex, a narrative is never innocent. White (1987, p. 7) has shown how even the simplest of medieval historical narratives, the chronicle, with its list of events in chronological order, betrays the medieval narrator's point of view. The chronicler's seemingly bizarre selection of events of different importance and duration reflects the medieval "religious *weltanschauung* where only the "other world" counts and God's intentions for this world are inscrutable" (Franzosi 2004a, p. 152). A basic, fundamental principle of silence and emphasis operates in the production of any text (including our own scientific texts) (Franzosi 2004a,b, pp. 22–23, 147–154, 175–177, 219–222). Certain events are selected in and others are selected out in story-telling (and, more generally, in the production of texts). There is a difference in meaning (not to mention for the patient, as semantic role in the sentence) if we say "socialists attacked fascists" or "fascists attacked socialists." There is a difference in meaning if we alter (even slightly) a sequence of events: "socialists hold a rally" "socialists and fascists clash" versus "socialists and fascists clash" "socialists

hold a rally.” There is a difference in meaning if we use different words in a sentence: “fascists attacked a socialist” versus “thugs attacked an innocent socialist.” There is a difference in meaning if we indicate a character trait of an actor with the use of syntactic adjectives: “fascists attacked a socialist” versus “fascists attacked an innocent socialist.” There is a difference in meaning if we use active versus passive syntactic forms: “fascists attacked a socialist” versus “a socialist was attacked by fascists.” While in the active form, syntactic subject and semantic agent and syntactic object and semantic patient coincide, in the passive form the syntactic subject is now the patient and the syntactic object is the agent. And while there cannot be a well-formed active sentence without syntactic subject, the object is syntactically optional in a passive form. It can be omitted. “A socialist was attacked.” But in so doing, the agency of the action is denied (on these issues, see Trew 1979, and Franzosi 2004a,b, pp. 122, 176) Thus, even subtle variations in sentence construction may lead to different meanings, as German mathematician and analytic philosopher Gottlob Frege concluded in a seminal article on “Sense and Reference”: “a subordinate clause may not always be replaced by another of equal truth value without harm to the truth of the whole sentence structure.” (Frege 1948, p. 230)

The combination of syntactic principles and semantic principles of silence and emphasis and of different ways of telling the same story through different plots ends up producing very different meanings. Which meaning we choose to convey through the conscious or even unconscious use of these linguistic principles (which reality we choose to construct) ultimately depends upon our point of view, our different ways of seeing. Any attempt to build rigor in the process of measurement of meaning that is not fully aware of these principles dangerously risks to become blind to its own rhetoric (on these general issues, see Franzosi 2004a, *passim*).

Inside Halliday’s color circle: a linear representation of processes

The picture in the frontispiece of Halliday’s 1994 edition of his *Functional Grammar* is reproduced, without colors, inside the book (p. 108), where a whole chapter (chapter 5: “Clause as representation”) is dedicated to an explanation of its meaning. There, Halliday zooms in the various aspects of his color circle. There, Halliday (1994, p. 108) provides a definition for processes:

A process consists, in principle, of three components:

- (i) the process itself;
- (ii) participants in the process;
- (iii) circumstances associated with the process.

In this interpretation of what is going on, there is doing, a doer, and a location where the doing takes place. This tripartite interpretation of processes is what lies behind the grammatical distinction of word classes into verbs, nouns, and the rest. ...

The process proper, in other words, is “typically realized by” a verbal group, a participant by a nominal group, and a circumstance by an adverbial group or prepositional phrase (Halliday 1994, p. 109). Circumstances refer to the “circumstances associated with or attendant on the process,” namely “the location of an event in time and space, its manner, or its cause,” to the when, where, how, and

why.⁴ Participants typically map grammatically onto Subjects and Complements. A circumstantial element is often itself a process, “a process that has become parasitic on another process,” in the form of subordinate clause (Halliday 1994, p. 151). There, Halliday turns the circular representation of processes of the frontispiece into a linear representation of individual processes.

The lion	chased	the tourist	Lazily	through the bush
Participant	Process	Participant	Circumstance	Circumstance
Nominal group	Verbal group	Nominal group	Adverbial group	Prepositional phrase

There, he links syntax and semantics via a theory of semantic roles, where participants are assigned different roles depending upon the different types of processes.⁵ Thus, participants become actor and goal in material processes, behavior in behavioural processes, sensor and phenomenon in mental processes, sayer and target in verbal processes, carrier and attribute, identified and identifier, token and value in relational processes, and existent in existential processes (see Halliday’s summary table of p. 143). Thus, in the material clause “the lion chased the tourist,” the two participants “lion” and “tourist” can be better specified as “actor” and “goal;” while in the clause of mental process type “I believe you,” the participants “I” and “you” are respectively sensor and phenomenon.

Different types of processes result in different types of linear templates, or, better, the types of participants and circumstances vary with the type of process. Thus, in processes of doing, the main participants are Actors and Goals (but these are more likely to be sensor and phenomenon in processes of sensing). At the linguistic level, processes, for Halliday, provide the basic template of the clause, assigning not only different types of participants, but also different numbers of participants. The verb “kill” calls for two participants: agent (the killer) and goal (the killed). But the verb “strike” has only one participant: the agent. To use a biological analogy, verbs contain the basic genetic information of the entire sentence, as captured by transitive and intransitive structures. In Chafe’s words: “The verb determines what the rest of the sentence will be like” (Chafe 1970, p. 97; on these issues, see Franzosi 2004a,b: 252–255).

Social relations

The privileged position of the verb over noun in language translates in a privileged position of action over actors and agents in narrated reality (although, of course, there cannot be action without actors). In turn, action implies interaction, relationships among social actors. “To act is always to act ‘with’ others” (Ricoeur 1984, p. 54). And that interaction constitutes the “*specificum sociologicum*,” in von Wiese’s words (Franzosi 2004a,b, pp. 260–262). A sociology of social relations brings us back to the work of Georg Simmel and his heirs (Franzosi 2004a, pp. 255–266). As

⁴ Halliday (1994, p. 150). On circumstances and their historical development in rhetoric, see Franzosi (2004a,b, pp. 124, 217, 255, 382 n. 67).

⁵ On semantic roles, see Franzosi (2004a,b, pp. 123–124, 253, 361 n. 44, 48, 49); see, also Chafe (1970) and Palmer (1994).

French sociologist Alain Touraine (1984) was to put it: “[T]he object of sociology is to explain the behavior of actors by the social relations in which they are placed. ... It is the relation, not the actor, that we must study.” Our understanding of action/interaction, as Ricoeur points out, fundamentally requires an understanding of a whole “conceptual network ... about ‘what,’ ‘why,’ ‘who,’ ‘how,’ ‘with whom,’ or ‘against whom’ in regard to any action.” (Ricoeur 1984, p. 55) Burke similarly wrote of “the five key terms of dramatism,” of the “pentad of key terms”: Act, Scene, Agent, Agency, Purpose, “what was done (act), when or where it was done (scene), who did it (agent), how he did it (agency), and why (purpose).” (Burke 1969, p. xv) But Ricoeur’s conceptual network and Burke’s pentad are nothing other than the journalists’ five Ws, nothing other than a story grammar in its most basic representation. Linguistics and sociology meet here in the common grounds of narrative.

Sociologists and narrative

Linguists and literary critics have, no doubt, done most of the hard work on narrative. But, in recent decades, social scientists, from psychologists to sociologists and anthropologists, have paid increasing attention to the study of narrative. Although this work has privileged traditional, qualitative approaches to narrative (e.g., Kohler Riessman 1993), there have also been attempts to introduce more formal and quantitative methodological approaches to text, narrative texts in particular (e.g., Heise 1979, 1989; Corsaro and Heise 1990; Abell 1987, 1993; Carley 1993; Fararo and Skvoretz 1986; Skvoretz and Fararo 1989, 1996; Fararo 1993; Skvoretz 1993). Although drawing from different disciplinary backgrounds, these approaches typically rely on a common understanding of narrative and social action in terms of agents (Who) and actions (What) in time (When) and space (Where), for some reasons (Why) and with certain outcomes and instruments. Abell’s approach comes closest to the story grammar approach proposed in these pages, based as it is on actors performing some actions at least partially ordered in time, on set theoretical mathematics, and on network representations (Abell 1987, particularly Chapter 5, pp. 52–61; Abell 1993).⁶

If the micro structure of narrative centered on actors and actions is central to the work of Abell (but also Heise and Fararo and Skvoretz), the sequential organization of narrative structures has been central to Abbott’s search for patterns of recurrent sequences (1983, 1995; Abbott and Hrycak 1990; Abbott and Barman 1997) and to Bearman and Stovel’s (2000) representation of narrative sequences as networks to shed light on the process of identity formation in an in-depth analysis of one short Nazi autobiographical account.⁷ Smith (2007) uses that same network approach to

⁶ Abell’s work contains precious insights. I do not share Skvoretz’s assessment (1993, 140) that “Abell’s narrative method appears to have limited utility ... [that] the narrative method holds little promise as a means to examine the routinized activity that composes a social institution.” In my view, it is unfortunate that the combination of a mathematical approach and symbols typically unfamiliar to sociologists, lack of substantive applications (see Heise’s critique on this point, 1993), not to mention the lack of software that would allow scholars to implement such complex approach, may have prevented Abell’s work to gain a wider audience.

⁷ Interesting work by Gibson (2000, 2003, 2005a,b), although not necessarily dealing with narrative, but with conversations and turn taking in conversation, is also similarly highly formalized and based on sequences and networks.

highlight the different structures of Italians' and Slavs' conflicting narratives of their past, in the border region of Istria.

A grammar of stories

In my own work on narrative, over the last twenty-five years, I have attempted to formalize the properties of narrative in a “story grammar,” keeping close to the linguists' approach to narrative texts. With the help of “rewrite rules” typically used in syntax, we can express Halliday's clause as a combination of three elements (hence my use of the term “semantic triplet”):

$$\langle \text{semantic triplet} \rangle \rightarrow \{ \langle \text{participant} \rangle \}, \{ \langle \text{process} \rangle \}, [\{ \langle \text{participant} \rangle \}]$$

where the symbol \rightarrow refers to a rewrite rule (or production), whereby an element to the left of the rule can be rewritten in terms of the elements to its right. Thus, the element $\langle \text{semantic triplet} \rangle$ can be rewritten as the combination of three elements: $\{ \langle \text{participant} \rangle \}$, $\{ \langle \text{process} \rangle \}$, and $[\{ \langle \text{participant} \rangle \}]$. The angular brackets $\langle \rangle$ denote elements that can be further rewritten; while “terminal elements,” i.e., the words or linguistic expressions found in the text, have no $\langle \rangle$. Curly brackets $\{ \}$ denote elements that can occur more than one time; while square brackets $[]$ denote optional elements. Thus, in the clause “workers strike” there is only one participant (the agent), while the clause “employer fires workers” has two participants (the agent, employer, and the goal or patient, workers). As a result, the grammar requires only the first participant; the second is optional.⁸ Each element of the triplet can then be further rewritten as follows:

$\langle \text{participant} \rangle$	\rightarrow	$\{ \text{actor} \} [\{ \langle \text{actor characteristics} \rangle \}]$
$\langle \text{actor} \rangle$	\rightarrow	crowd workers student police ...
$\langle \text{actor characteristics} \rangle$	\rightarrow	$\langle \text{type} \rangle \langle \text{number} \rangle \langle \text{organization} \rangle \langle \text{space} \rangle \dots$
$\langle \text{type} \rangle$	\rightarrow	female male immigrant
$\langle \text{organization} \rangle$	\rightarrow	$\langle \text{firm} \rangle \langle \text{trade union} \rangle \dots$
$\langle \text{process} \rangle$	\rightarrow	$\{ \langle \text{verb} \rangle \} \{ \langle \text{circumstances} \rangle \}$
$\langle \text{verb} \rangle$	\rightarrow	strike rally layoff charge ...
$\langle \text{circumstances} \rangle$	\rightarrow	$\langle \text{time} \rangle \langle \text{space} \rangle \langle \text{type} \rangle \langle \text{reason} \rangle \langle \text{instrument} \rangle$ $\langle \text{outcome} \rangle$
$\langle \text{type} \rangle$	\rightarrow	wildcat general ...
$\langle \text{reason} \rangle$	\rightarrow	$\langle \text{name of reason} \rangle [\langle \text{semantic triplet} \rangle]$
$\langle \text{name of reason} \rangle$	\rightarrow	wage increases layoffs ...
$\langle \text{instrument} \rangle$	\rightarrow	bomb stick ...
$\langle \text{outcome} \rangle$	\rightarrow	$\langle \text{name of outcome} \rangle [\langle \text{semantic triplet} \rangle]$
$\langle \text{name of outcome} \rangle$	\rightarrow	positive negative disruption ...

Semantic triplets can be aggregated hierarchically into higher-level objects, like events, and events into macro-events:

$$\begin{aligned} \langle \text{macro} - \text{event} \rangle &\rightarrow \{ \langle \text{event} \rangle \} \\ \langle \text{event} \rangle &\rightarrow \{ \langle \text{semantic triplet} \rangle \} \end{aligned}$$

⁸ An example of a fully specified grammar, developed to capture narratives of historical protest events, can be found in www.pc-ace.com.

Furthermore, to allow for the possibility of having different narratives for the same event or triplet (when using different text sources), the event⁹ can be rewritten as

$$\begin{array}{ll} \langle \text{event} \rangle & \rightarrow \{ \langle \text{semantic triplet} \rangle \} [\langle \text{alternative event} \rangle] \\ \langle \text{alternative event} \rangle & \rightarrow \langle \text{event} \rangle \end{array}$$

How can this help social scientists? An application to the rise of Italian fascism (1919–1922)

A story grammar has desirable linguistic and mathematical properties (Franzosi 1989, 1994, 2004a, 2010, pp. 33–41). In particular, the relational properties of the grammar (e.g., with actors related to actions, actions to time and space and objects) make a story grammar a far more superior tool than content analysis, the traditional quantitative social science approach to texts. In content analysis, coding categories are listed seriatim with no relationship to one another (on content analysis, see Franzosi 2004b, 2008; for a comparison of content analysis and quantitative narrative analysis, see Franzosi 2004a, pp. 59–61, 91–92; 2010, pp. 34–36). But, can this methodological tool also help social scientists in their pursuit of explanations for socio-historical problems? To understand the answer to that question, let's code the following newspaper article taken from *Avanti!* of 19/07/1921 (page 2, column 2):

Unheard-of, fascist monstrosities around Pavia (Pavia 18/07/1921)

Republicans plunged in Bissone di S. Cristina around 10pm of this month at the pub Prati. A guy, who went by the name of “captain,” took out a list of names and did the roll call loudly. Five of the workers in the pub whose name was called out were sequestered and taken outside the village. The others were taken out of the pub flanked by two wings of fascists, and, one by one, savagely beaten up with retorted cowhide with led inside. Among the sequestered workers, there was the comrade Fiorenzo Scala at whose house the fascists had carried out earlier a careful search. Thanks to the arrival of carabinieri, the prisoners could flee. But so did the republicans. ...¹⁰

This article is a good example of narrative text, made up of chronologically ordered clauses characterized mostly by Halliday's processes of doing (*plunge*,

⁹ The same, can be done for a semantic triplet to allow for an alternative sentence construction (e.g., “fascists attack socialists,” according to *Avanti!* and “socialists attack fascists,” according to *Il Popolo d'Italia*).

¹⁰ Inaudite mostruosità fasciste nel pavese (Pavia 18/07/1921). A Bissone di S. Cristina i repubblicani piombarono verso le 22 del 5 c.m. nell'osteria Prati. Un tale—che si faceva chiamare capitano—sfoderò una lista di nomi e fece la chiama. Cinque fra i lavoratori presenti, il cui Name era scritto sulla nota, furono sequestrati e portati fuori dal paese. Gli altri furono fatti uscire dall'osteria fra due ali di fascisti, e, uno alla volta, percossi selvaggiamente con un nervo di bue piombato. Fu sequestrato anche il compagno Fiorenzo Scala, nella cui abitazione i fascisti avevano compiuto una minuta perquisizione. Grazie al sopraggiungere dei carabinieri, i prigionieri poterono darsela a gambe. Ma anche i repubblicani se ne andarono liberamente. ...

sequester, beat up, search). Within the categories of a story grammar, this narrative can be structured as follows:

- [Semantic triplet 1: [Participant: [Actor: *republicans*]] [[Process: [[Verb: *plunge*] [Circumstances: [Space: [City: *Bissone di S. Cristina*] [[Location: *pub*] [Name: *Prati*]]]] [[Time: [Date: *05/07/1921*] [Hour: *10pm*]]]]]
- [Semantic triplet 2: [Participant: [Actor: *captain*]] [Process: [[Verb: *does roll call*] [Circumstances: [Type of action: *loudly*] [Instrument: *list*]]] [Participant: [Actor: *workers*]]]
- [Semantic triplet 3: [Participant: [Actor: *republicans*]] [Process: [Verb: *sequester*] [[Verb: *take*] [Circumstances: [Space: [Space direction: *outside*] [City: *Bissone di S. Cristina*]]]]] [Participant: [[Actor: *workers*] [Actor characteristics: [Number: *5*]]]]]
- [Semantic triplet 4: [Participant: [Actor: *republicans*]] [Process: [[Verb: *take*] [Circumstances: [Space: [Space direction: *out of*] [Location: *pub*]]]]] [[Verb: *beat up*] [Circumstances: [Instrument: *cowhide with lead inside*]]]]] [Participant: [[Actor: *workers*] [Actor characteristics: [Number: *those left*]]]]]
- [Semantic triplet 5: [Participant: [Actor: *fascists*]] [Process: [[Verb: *search*] [Circumstances: [Type of action: *carefully*] [Space: [City: *Bissone di S. Cristina*] [Location: *house*] [Property: [Actor: *individual*] [Actor characteristics: [Name: *Fiorenzo*] [Lastname: *Scala*] [Political affiliation: *comrade*]]]]]]]]]
- [Semantic triplet 6: [Participant: [Actor: *fascists*]] [Process: [Verb: *sequester*]] [Participant: [[Actor: *individual*] [[Actor characteristics: [Name: *Fiorenzo*] [Lastname: *Scala*] [Political affiliation: *comrade*]]]]]]]
- [Semantic triplet 7: [Participant: [Actor: *carabinieri*] [Process: [Verb: *arrive*]]]
- [Semantic triplet 8: [Participant: [Actor: *prisoners*] [Process: [Verb: *flee*]]]
- [Semantic triplet 9: [Participant: [Actor: *republicans*]] [Process: [Verb: *flee*]]]

This example of text coding shows that, by-and-large, most narrative elements found in the original input text also appear in the coded output, but appropriately organized in the coding categories of a complex story grammar. Why should this matter? It matters because: 1. for each category you know exactly what it contains (e.g., an actor, a reason, a city); thus, you know where to look if you are interested in actors; 2. within each coding category you find one, and only one, textual element (e.g., the specific actor “workers” or “police”); 3. for each textual element you also know its exact relation to other textual elements (e.g., an actor related to a specific reason of a specific action; e.g., “workers” “strike”).

Data in this example were coded within the categories of a complex story grammar using software I specifically developed for the analysis of narrative texts: PC-ACE (Program for Computer-Assisted Coding of Events¹¹). PC-ACE takes advantage of the relational properties of the grammar (e.g., actors related to actions, actor characteristics to actors, actions to their circumstances) and organizes data in relational format, with different text elements organized and stored in different computer tables, and where links between related information across different tables are established via common, overlapping fields (e.g., a common ID field between an actor_table that stores verbal

¹¹ See Franzosi (1990a,b, 1995). The current Windows release of PC-ACE—available in the public domain for free download at www.pc-ace.com—has been partly developed with a grant from the Nuffield Foundation.

expressions for actors—workers, students, employers, police—and an action_table that stores actions—strike, rally, demonstrate, layoff, charge). No doubt, a computerized tool like PC-ACE is indispensable for large projects.¹² In many ways, no computer software, no substantive applications of a story grammar approach narrative on a large scale (Franzosi 2010, pp. 59–60). The sheer complexity and sophistication of such coding schemes as story grammars would limit their use to trivial, illustrative examples.

I used PC-ACE to collect data on the rise of Italian fascism (1919–22), on the processes of working-class mobilization (1919–20, the “red years”) and fascist counter-mobilization (1921–22, the “black years”). I coded over 15,000 and over 17,000 newspaper articles from the socialist newspapers *Il lavoro* and *Avanti!*, and nearly 20,000 from the fascist newspaper *Il Popolo d'Italia*. The three databases resulted in over 17,000 semantic triplets from *Il Lavoro*, nearly 140,000 from *Avanti!*, and nearly 100,000 from *Il Popolo d'Italia*.

Despite the size of such datasets and the nature of the data (basically, words), the relational properties of the database make data queries general via a simple query language: SQL (Structured Query Language). SQL is based on a handful of commands, in particular: *select* (a specific field), *from* (a specific table), *where* (condition). The function *count*, within a select clause, further allows us to compute frequency counts of selected objects (for a lengthier treatment of RDBMS and SQL, see Franzosi 2004a, pp. 66–89; 2010, pp. 68–81). Thus, the following query (based on the *Avanti!* database) extracts the frequency distribution of all names of actors (actor_name) from the actor_table:

```
select actor_name, frequency=count(ID)
from actor_table
```

yielding the list of most frequent actors in the database¹³ of Table 1.

A query aimed at extracting the participant’s characteristics (e.g., type of actor or political affiliation) would reveal that of the 38,873 “individuals,” 10,708 are “fascist,” 3,627 “workers,” 3,542 “socialist” or “communist,” 1,174 “police,” 947 “entrepreneurs,” 425 “landowners,” ...

A query operating on the process_table would yield a frequency distribution of all actions registered in the database (see Table 2).

```
select process_name, frequency=count(ID)
from process_table
```

This simple frequency distribution of actions gives a clear sense of the turbulent nature of those years. Indeed, the most frequent types of action are actions of violence (e.g., shoot, wound, assault, beat up, kill). The power of SQL, however, lies in its ability to operate across multiple tables, thus allowing us to match actions to actors. Which actors, for instance, committed such actions as “shoot,” “wound,”

¹² PC-ACE, unfortunately, does not do the hard work of coding the text for you. You have to do it. Artificial Intelligence has yet to deliver on early promises of finding a solution to the problem of computer understanding of natural languages (and it may be a while yet before it happens). All PC-ACE does (or any other currently available software of textual analysis, for that matter, except for basic word counts) is provide a computerized tool that will make the coding task easier and more reliable.

¹³ Most items found in the database for any coding category occur with a frequency of 1 or, in any case, less than 3 (for similar findings, see Franzosi 2004a, p. 293).

Table 1 Frequency distribution of actors in the *Avanti!* database

Actor	Frequency
Individuals	38,873
Workers	16,428
Fascists	9,763
Socialists	5,504
Police	5,254
...	
Industrialists	1,553
Land owners	850
Land workers	833
...	

“kill?” Do the fascists in this period typically engage in violent actions? The following query addresses that question:

```
select process_table.actor_name, frequency=count(process_table.ID)
from actor_table and process_table
where actor_table.actor_name="Fascists" and
actor_table.triplet_ID=process_table.triplet_ID
```

Indeed, most actions the fascists perform are actions of violence (see Table 3).

Thus, by counting, words can be transformed into numbers (e.g., Franzosi 2004a, pp. 116–118; 2010, p. 81). Out of these numbers variables can be constructed and models can be tested (e.g., Franzosi 1994a, 2004a, pp. 97–118; 2010, pp. 107–127).

Yet, turning words into numbers and numbers into variables for use in regression models is somewhat of a missed opportunity. The relational nature of narrative data, with actors related to other actors via their actions, and the fundamental characteristics of narrative (namely, time and space) allow us to move away from a variable-centered approach to an actor-centered one. Several statistical techniques do just

Table 2 Frequency distribution of actions in the *Avanti!* database

Action	Frequency
Rally with Speech Making	2,953
Strike	2,588
Beat up	2,165
Shoot	1,812
Arrest	1,764
Wound	1,564
Threaten	1,265
Assault	1,165
Kill	951
...	

Table 3 Frequency distribution of actions performed by the fascists in the *Avanti!* database

Action	Frequency
Beat up	1,686
Shoot	1,054
Assault	1,045
Wound	989
Destroy	770
Threaten	672
Clash	635
Set fire	530
Kill	504
...	

that: from network models (based on the relations among social actors around specific spheres of action, e.g., violence), time series models (e.g., event history analysis, based on time and duration), sequence analysis (based on the temporal order of actions and events), and GIS-Geographic Information System-tools (based on the geographic location of actions) (see Franzosi 2010, pp. 109–127).

Let's look at network models (on network models applied to narrative data, see Franzosi 1998, 1999, 2004a, pp. 97–109; 2010, pp. 109–114). Network graphs provide a visual representation of social interaction among social actors (where, in network jargon, actors are known as nodes and the action/interaction as relations). To produce a network model, we first need to construct a network matrix. We start from a basic table of our participants and their actions by each occurrence during the 1919–1922 period (for the sake of illustration, I focus here on all actions of violence, those based on disaggregated verb phrases such as “beat up,” “shoot,” “wound,” “assault,” “kill” of Table 2 above and “destroy,” “burn” of Table 3), in their respective roles of subjects and objects (or agents and patients to use the language of semantic roles). Some 5,000 such daily interactions aggregated as violent actions are recorded in my *Avanti!* database (see Table 4).

Then, we compute same-string frequencies (e.g., fascists-violence-workers) to produce a frequency matrix of social interactions (these frequencies can be computed over the entire period, as in this case, or over specific sub-periods, e.g., 1919–1920, the so-called “red years” on a yearly, quarterly, monthly, weekly, or even daily basis). The matrix results (see Table 5) show that the fascists and the police were the typical agents of violence against the Left: workers, socialists, communists (e.g., 871 actions of violence by the fascists against workers).

Such a matrix can then be used to produce the network graph of Fig. 1, where the thickness of the line is roughly proportional to the number of violent actions between any two actors (the nodes); the arrows in the graph measure the direction of a relation between any two nodes (e.g., between fascists and workers); and the numbers refer to the frequency of actions of violence found in the database (the numbers closer to a node/actor refer to the actions of violence against that actor by the other actor in the relation). The network graph of Fig. 1 visually illustrates the social relations of the 1919–1922 period in Italy, with two centers of violence, focused on the police and the

Table 4 List of daily occurrences of triplets of violence in the *Avanti!* database

Date	Subject	Action	Object
19-Jan-1919	Police	Violence	Workers
19-Jan-1919	Police	Violence	Workers
...			
20-Feb-1919	Police	Violence	Socialists
...			
26-May-1919	Protesters	Violence	Police
26-May-1919	Protesters	Violence	Police
26-May-1919	Protesters	Violence	Police
...			
07-Nov-1919	Fascists	Violence	Workers
16-Oct-1920	Fascists	Violence	Socialists
...			
30-Oct-1920	Fascists	Violence	Socialists
...			
30-Oct-1920	Fascists	Violence	Workers
30-Oct-1920	Fascists	Violence	Workers

fascists. They are the main agents of violence, while working-class actors (e.g., workers, socialists) are the patients of the fascists' actions.

When compared to traditional variable-based statistical techniques, network models allow us to bring out patterns in the data without losing touch with the

Table 5 Frequency distribution of triplets of violence in the *Avanti!* database

Subject	Action	Object	Frequency
Fascists	violence	Workers	871
Fascists	violence	People	677
Fascists	violence	Socialists	482
Fascists	violence	Individuals	324
Police	violence	Workers	207
Individuals	violence	Workers	174
Police	violence	People	129
Fascists	violence	workers (agricultural)	126
Police	violence	Protesters	92
Police	violence	workers (agricultural)	88
Fascists	violence	Communists	87
Police	violence	Socialists	84
Police	Violence	Individuals	81
Individuals	Violence	Individuals	69
Workers	Violence	Fascists	56

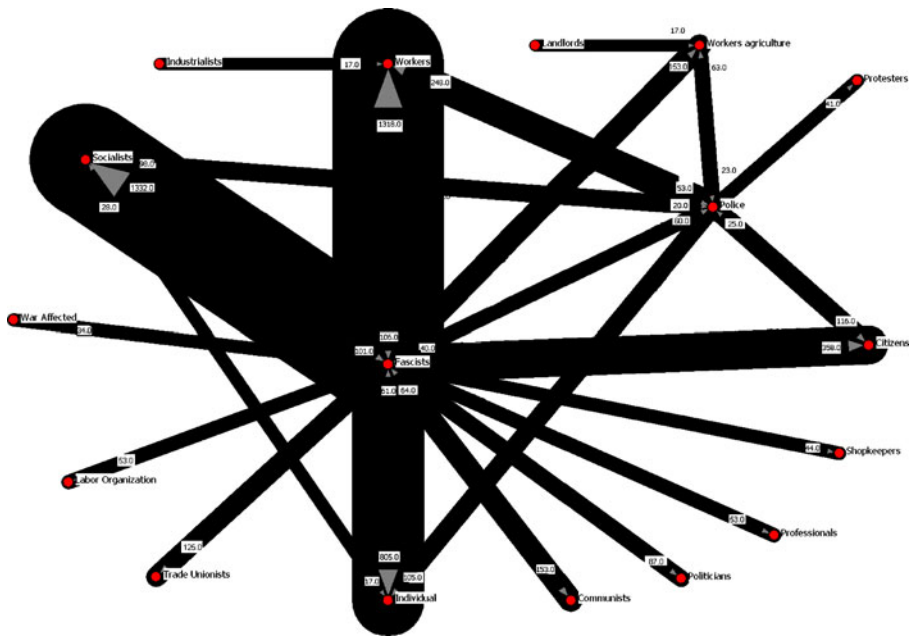


Fig. 1 Network graph of violence (1919–1922)

actors, no longer hidden behind variables. They exploit the fundamental properties of actors related to other actors in specific spheres of action (e.g., violence). They deal fundamentally with Simmel's and von Wiese's *specificum sociologicum*: social relations.

But more can be done with narrative data. Figure 2 provides a hot spots map of the distribution in time and space of fascist actions of violence and of the founding of new *fasci* (the organizational cells of the fascist party at the local level) (Franzosi 2010, pp. 124–127). The map exploits another fundamental property of narrative: that social action, as represented in narrative, is located in time and space (see also Abbott on this point, 1997). And although this map is used here for purely illustrative purposes, it does show an overlap between fascist violence and organizational drive, between mobilization as action and mobilization as institution building.

Illustrative as they are, the analyses highlight patterns of interaction among social actors and the strategic use of different courses of action. The isomorphic relationship between narrative and story grammars centered on actors and their actions transforms the syntactical/semantic approach to narrative adopted here into a “syntax of social life,” in Peter Abell's words. Quantitative narrative analysis moves away from variable-centered explanations and toward narrative explanations (Abell 2004; Abbott 1992; Franzosi 2004a, pp. 238–247) “Our normal methods—Abbott wrote (1992, p. 428)—parse social reality into fixed entities with variable qualities. They attribute causality to the variables ... rather than to agents; variables do things, not social actors. Stories disappear.” Quantitative narrative analysis remains true to social action and stories.

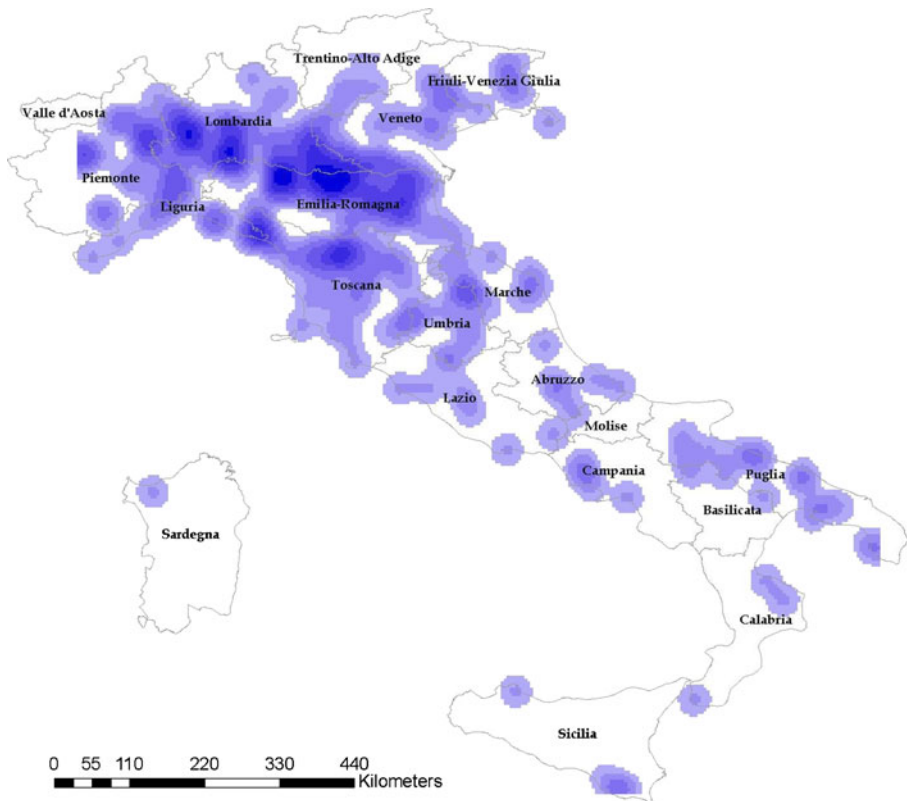


Fig. 2 Hot spots map of fascist actions of violence and fascist institution building (founding of local fasci) (1921)

Confessions of an author

Like the entire world, I was convinced that all colors are contained in light. No one had ever told me otherwise ... I quickly wanted to see through a prism [a la Newton] ... I was in a totally white room. As I held the prism before my eyes, I expected, in keeping with Newtonian theory in mind, that the entire white wall would be separated into different colors ... But I was quite amazed that the white wall showing through the prism remained as white as before. Only when there was something dark did a more or less distinct color show ... It required little thought to recognize that an edge was necessary to bring about colors. I immediately spoke out loud to myself, through instinct, that Newtonian theory was erroneous.

(in Matthaai 1971, p. 199)

It is an unlikely scientist who left these words in his *Confessions of an Author*, the last part of a three-part *Theory of Colours*: Johann Wolfgang von Goethe (1749–1832), an author better remembered for his poetic work, for his *Faust*, *The Sorrows of Young Werther*, or *Elective Affinities*. But Goethe himself pegged his hope for immortality on his scientific work—on plants, color, geology (see the collection of

scientific writings, von Goethe 1988; on Goethe's position in the history of science, see Amrine 1996a, 1996b, 1999; see also Amrine et al. 1987, Seamon and Zajonc 1998). Just on color, Goethe left behind some 2,000 pages and several different pieces of work that, starting in 1791, kept him occupied till the end of a long life in 1829 (Matthaei 1971, pp. 5, 205–206). Indeed, with Goethe, Halliday's color circle, becomes a commonplace representation ("three or six colours, which may be conveniently included in a circle," 1970, p. lvii; see also Matthaei 1971, pp. 4, 41, 50). True to his words "that Newtonian theory was erroneous," Goethe mounted a virulent attack against Newton and Newtonianism—something he would later come to regret. In the "Preface" to his *Theory of Colour* he wrote:

We examine the Newtonian theory; a theory which by its ascendancy and consideration has hitherto impeded a free inquiry into the phenomenon of colours. We combat that hypothesis, for although it is no longer found available, it still retains a traditional authority in the world (von Goethe 1970, p. xli).

He goes on to compare Newton's theory to an old castle, "constructed by its architect with youthful precipitation," forever enlarged, fortified, and worked upon, but at present "uninhabitable" and "a deserted piece of antiquity." He sets himself the task of "begin[ning] at once, without further ceremony, to dismantle it from gable and roof downwards" (von Goethe 1970, p. xliii). Without directly naming Newton, and simply referring to him as a "great mathematician," Goethe accuses Newton of not only "having been mistaken in his observations as an experimentalist," but also of having "employed the whole force of his talent to give consistency to this mistake" (von Goethe 1970, p. lx). He would reiterate that point, almost word by word, when discussing the relation between his theory of color and mathematics:

A great mathematician was possessed with an entirely false notion on the physical origin of colours; yet, owing to his great authority as a geometer, the mistakes which he committed as an experimentalist long became sanctioned in the eyes of a world ever fettered in prejudices. (von Goethe 1970, p. 287)

"How could a mind so great, so all encompassing," Goethe asks, "be so wrong?" (cited in Böhme 1987, p. 153; see also Matthaei 1971, p. 14). He thunders against the "incredible arrogance ... [of] the Newtonian school" (von Goethe 1970, p. xlv).

That battle is over

"Goethe was a master salesman of his own ideas." Judd wrote in his "Introduction" to the MIT 1970 facsimile edition of the first 1840 English translation of *The Theory of Colours* by Charles Lock Eastlake (von Goethe 1970, p. xiii). On page after page of his 400-page *Theory of Colours*, Goethe convincingly describes one experiment after the other, seemingly delivering a fatal blow to the Newtonian conception of colors (if not to science, altogether). He directly appeals to the lay reader, to the amateur, refusing a view of science for the specialist (von Goethe 1970, pp. 353–355). In many of his writings, he makes the point emphatically that it is not his "intention to write solely for the expert;" "it is my desire to interest the general public" (in Matthaei 1971, pp. 13, 15). "I have gone to great pains to enable

anyone to undertake the practical experiences under discussion easily” (in Matthaei 1971, p. 16)

For as confident as Goethe was to have put the study of colors on a new footing though, he surely overstated his case. “Goethe’s explanation of color makes no physical sense at all,” Judd wrote (von Goethe 1970, p. xi). When 1932 German physics Nobel laureate Werner Heisenberg visited the controversy in a lecture delivered in Budapest on May 5th 1941, he proclaimed the following verdict: “That battle is over.” And Goethe lost it. “The decision on ‘right’ and ‘wrong’ in all questions of detail has long since been taken. ... [V]ictory, and hence influence on the research of the following century, has been Newton’s. ... [His] theory forms to this day the basis of all Physical Optics” (Heisenberg 1952, pp. 60, 62). And for sure, hardly anyone followed on Goethe’s footsteps. When Heisenberg revisited the issue of Goethe’s science in a lecture delivered to the Goethe Society in Weimar, on May 21, 1967, he stated: “[I]n spite of the enormous effect Goethe’s poetry exercised in the nineteenth century, his thoughts upon science have become known and fruitful only within a relatively small circle” (Heisenberg 1974, p. 132). Holdrege similarly writes of Goethe the scientist: “Goethe is a perennial outsider ... within the broader contemporary scientific community his efforts are virtually unknown or deemed irrelevant to the advancement of science” (Holdrege 2005, p. 12).

Yet, Heisenberg concedes, “Goethe’s colour theory has ... borne fruit in art, physiology and aesthetics” (Heisenberg 1952, p. 60; on Goethe’s physiological contributions, see Böhme 1987). “Newton’s theory makes possible a certain control over the phenomena of light and their practical use but it is plainly of no assistance to a better appreciation of the world of colour surrounding us” (Heisenberg 1952, pp. 60, 63). It may well be that “Newton’s scientific method is superior to Goethe’s intuitive power, but basically the two theories simply deal with different things,” (Heisenberg 1952, p. 64) the result of a separation of subjective and objective spheres that constitutes “the first presupposition of all research” (Heisenberg 1952, p. 64). Newton was a champion of the objective, Goethe of the subjective (but “no less powerful for that,” p. 68). Goethe himself wrote that “the world divides itself into two parts, and the human being as *subject* stands opposed to the *object*” (von Goethe 1970, p. 75).

Goethe rejects the idea of an objective phenomenon, separate from the subjectivity of the observer. “The phenomenon is not detached from the observer, but intertwined and involved with him” (von Goethe 1998, p. 155, Maxim n. 1224). No doubt speaking against Newton’s *experimentum crucis* on light, for Goethe, “an experiment, even connected experiments, prove nothing” (in Matthaei 1971, p. 60). *Anschauung* (or “intuitive perception”) provides the basis of Goethe’s science, of what he called “*delicate empiricism* (*zarte Empirie*)—the effort to understand a thing’s meaning through prolonged empathetic looking and seeing grounded in direct experience” (Seamon 1998, p. 2; see also Ebach 2005, and the other articles in Robbins and Holdrege 2005). No doubt, the 1993 Nobel laureate Barbara McClintock would be Goethe’s champion, with her unusual approach to science so empathetically portrayed in Keller’s book, *A Feeling for the Organism. The Life and Work of Barbara McClintock* (Keller 1983).

And *Anschauung* is fundamentally relational. “Our intuitive perception is nothing more than relating our experiences of the phenomenon to itself and to other

phenomena. The *Anschauung* is thinking by way of relationship” (Ebach 2005, p. 7). As Goethe wrote: “Throughout nature, as presented to the senses, everything depends on the relation which things bear to each other, but especially on the relation which man, the most important of these, bears to the rest” (von Goethe 1970, p. 75). Furthermore, thinking relational means thinking of nature as wholeness, rather than isolated facts and phenomena brought under experiment (on Goethe’s conception of nature as a whole, see Bortoft 1996). “Nature, however manifold it may appear, is nevertheless always a single entity, a unity ...” (von Goethe 1996, p. 92).

In the end, Netwotn’s and Goethe’s “theories ... deal with two entirely different levels of reality” (Heisenberg 1952, p. 67). Goethe’s theory does not count but weighs events. As for past events, it interprets them, rather than explains them. Heisenberg, the modern physicist and Nobel laureate, regards this separation of an objective and subjective as dangerous, as detrimental to the advancement of science (pp. 70–71), since it leads to “an over-simplification of reality” (p. 74). But, if “the history of science is science itself,” as Goethe claims (von Goethe 1970, p. xliv), there may be hope yet. ... For Heisenberg (p. 68), modern science is bridging this dangerous gap between these two seemingly opposing views of science. In any case, Goethe, too, in the end, resorted to abstraction, albeit not mathematically, but through the notions of *Urphänomen* and archetype, as original phenomenon, an Idea, an Ideal that underlies the many different naturally occurring phenomena, the different living forms (Heisenberg 1974, pp. 138–139; see also Ebach 2005, on the archetype).

Signs and systems of rules

Winner or loser in his battle with Newton, what would the great German Master think of my attempt at measuring meaning, at turning words into numbers, of my “formal” approach to narrative, with its mathematical foundations in set theory and its graphical representations (on the set theoretic foundations of QNA, see Franzosi 1994a, 2004a, pp. 97–100)? We do know that Goethe, in his polemic with Newton, thought that mathematics can be “of the greatest use,” but “a false application of its methods can also be quite “prejudicial” (von Goethe 1970, pp. 286–287). As for the diagrams of Newtonian physics, they “represent mere notions; they are symbolic resources, hieroglyphic modes of communication, which by degrees assume the place of the phenomena and of Nature herself, and thus rather hinder than promote true knowledge” (von Goethe 1970, p. xlix). “Mathematics ... in practice ... is an art like rhetoric” (von Goethe 1996, p. 68).

The mathematician relies on the element of quantity, on all that is defined by number and size, and thus to some degree on the universe in its external form. But if we set out to apply the full measure of mind and all its powers to this universe, we will realize that *quantity and quality* must be viewed as two poles of material existence. This is why the mathematician refines his language of formula so highly; as far as possible he wants to incorporate the incalculable world into the realm of measure and number. Everything will then seem

graspable, comprehensible, and mechanical ...” (von Goethe 1996, p. 65; emphasis added).

Mathematics, for Goethe, runs into a double danger. First, it risks reducing the complexity of nature to a handful of signs. Modern science proceeds by simplifying the complexity of nature and focusing on a handful of relations, represented as signs in mathematical models (according to the principle of parsimony). One can only hope that other relations, not included in the models, play no role in the understanding of the phenomenon! And that goes for the social sciences as well. Abbott (1990) has argued that two contrasting views of social phenomena (e.g., careers) are dominant in sociology: stochastic (or statistical) and “whole.” “In this view, the career is a reality, a whole, not simply the list of successive realizations of an underlying stochastic process.” Underlying these different approaches to the social scientific enterprise, even if implicit, are different assumptions about causality, time, relations, and explanations—just as Goethe would have it. Second, mathematics also risks reducing the phenomenon to *just* signs.

The investigator of nature should take heed not to reduce observation to mere notion, to substitute words for this notion, and to use and deal with these words as if they were things. ... Yet, how difficult it is to avoid substituting the sign for the thing; how difficult to keep the essential quality still living before us, and not to kill it with a word (von Goethe 1996, pp. 283, 302).

Goethe has a point. Each of the steps involved in the production of knowledge has its signs (e.g., actors into variables, social relations into coefficients). More problematically, each sign then obeys its specific system of rules.

Step 1 Take the first few “triplets” of the event coded above. When I first read the event in the PC-ACE database, it had been coded as:

[Semantic triplet: [Participant: [Actor: *republicans*]] [[Process: [[Verb: *arrive*] [Circumstances: [Space: [City: *Bissone di S. Cristina*] [[Location: *pub*] [Name: *Prati*]]]] [[Time: [Date: *05/07/1921*] [Hour: *10pm*]]]]]]

[Semantic triplet: [Participant: [Actor: *republicans*]] [Process: [Verb: *sequester*] [[Verb: *take*] [Circumstances: [Space: [Space direction: *outside*] [City: *Bissone di S. Cristina*]]]] [Participant: [[Actor: *workers*] [Actor characteristics: [Number: *5*]]]]]]

[Semantic triplet: [Participant: [Actor: *republicans*]] [Process: [[Verb: *take*] [Circumstances: [Space: [Space direction: *out of*] [Location: *pub*]]]] [[Verb: *beat up*] [Circumstances: [Instrument: *cowhide with lead inside*]]]] [Participant: [[Actor: *workers*] [Actor characteristics: [Number: *those left*]]]]]]

The code was “correct,” by the coding rules of the research project: 1. each coding category contained the expected textual element (e.g., each instance of the category <verb> contained verbs and not nouns); 2. each coding category contained only one element (e.g., each <verb> correctly contained “sequester,” “take,” and “beat up” separately, rather than “sequester, take outside and beat up” together); 3.

the relationship between the various coding categories and their content made sense in terms of their interrelated meaning and of the overall meaning of the story (the coded story enjoyed both local and global “semantic coherence”). The coded story “passed” the hurdles imposed by the system of coding rules (indeed, it had “passed” the “semantic coherence” test of a verifier who had read it). *Yet*, the story *was* different from the original newspaper article (as denoted by missing information in the code, such as the pub and what happened there, the use of the verb “arrive” instead of the stronger “plunge,” or the end of the story, with both workers and fascists being able to flee at the carabinieri’s arrival). As Heisenberg put it in his 1941 Budapest lecture (Heisenberg 1952, p. 71): “our experiments are not nature itself, but a nature changed and transformed by our activity in the course of research.” It is against that violence on nature that Goethe had thundered: “Let the observer of nature suffer the primordial phenomenon to remain undisturbed in its beauty” (von Goethe 1970, p. 73).

Step 0 Of course, the original newspaper article was no doubt even more selective about the information that finally went into the article from “social reality” (indeed, Step 0, taking a step back to the sources of our data). How many were the fascists that plunged or arrived in Bissone di Santa Cristina? How did they arrive, and from where? Who were they? Were they young? How were they dressed? How did each of them behave? What did they say? Certainly, much more went missing from reality to the article, than from the article to the code. And in any case, how do we know that “what the papers say” actually happened? Newspaper “signs” themselves obey their own set of rules: the rules of the language in which the article was written (the syntax and semantics of Italian language), but also, newsworthiness of stories and, more generally, the rules of selection of news, the format of presentation of news (those journalistic rules of the 5 Ws that so closely match a “story grammar”) (on newspapers as sources of socio-historical data and on the social scientists’s justifications for their use, see Franzosi 1987, 2004a, pp. 167–172, 180; Olzak 1989; Earl et al. 2004; Ortiz et al. 2005).

Step 2 The process of creation of numbers obeys its own set of rules. First, SQL queries, the first step in the sequence, must be properly formed, not just syntactically, but, most importantly, semantically (after all, the SQL jet engine would automatically signal any syntax error, not so for semantic errors). Thus, if a user is extracting from the database all instances of violent actions, the SQL query would need to include all verbs denoting violence, or the count (the “number” to be used in the next step in the sequence of scientific production) would be wrong:

```
select frequency_fascist_violence=count(process_table.ID)
from actor_table and process_table
where actor_table.actor_name="Fascists" and
      actor_table.triplet_ID=process_table.triplet_ID and
      (process_table.processs_name="sequester" or
       process_table.processs_name="beat up" or ...)
```


The numbers you get as fascist violent actions depend upon which individual verbs you put in the list. Would you put “threaten” there? If so, would you extend that to verbal threats? Your numbers will reflect your decision rules (even when the rules are fuzzy, such as threaten as a violent verb) (on these issues, see Franzosi 2004a, pp. 287–295, 2010, p. 82; on the use of fuzzy logic, see Franzosi 2004a, pp. 293–95, 2010, pp. 95–96).

Step 3. The next step: statistical/mathematical modeling. I give an example of one such model in *From Words to Numbers: Narrative, Data, and Social Science* (Franzosi 2004a, pp. 116–118) with a dependent variable measured by the level of fascist violence in 1921 in a province¹⁴ and two independent variables measured by 1) the level of social conflict in that province during the “red years” and 2) the percentage of communes in the province that obtained a socialist majority during the administrative elections of November 1920. I write:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i \quad (3.1)$$

where:

the subscript i indexes the set of provinces upon which the regression is run ($0 < i < 57$);

Y_i refers to the number of fascist violent actions in province i in 1921;

X_{1i} refers to the number of conflict events on the Left in province i during the Red Years of 1919–20;

X_{2i} refers to percentage of communes in province i that obtained a socialist majority during the administrative elections of November 1920;

and where the expected signs of the parameters are: $\beta_1 > 0$ and $\beta_2 > 0$. β_0 is a constant measuring the intercept. If the “red menace” hypothesis were correct, we should expect a violent fascist reaction in those provinces where the middle and/or upper classes may have felt threatened by the level of working-class mobilization in their provinces.

I summarize the results of the least-squares estimate of Eq. 3.1 in a table and conclude: “The coefficients of both independent variables have the expected sign and are highly significant.” The R-squared explains a good proportion of the variance (.573). The statistical model behaves “well,” according to the rules of Ordinary Least Squares estimation. I conclude: “The regression results thus seem to confirm an interpretation of fascist violence as a response to a threat from below” (Franzosi 2004a p. 116).

¹⁴ Thirty-six provinces were excluded from the analyses, because of missing measurements on two of the variables in the model (nineteen missing on fascist violence; twenty-nine missing on protest from below). The fifty-seven provinces on which the regression model was run are: Alessandria, Ancona, Arezzo, Bari, Benevento, Bergamo, Bologna, Brescia, Cagliari, Caltanissetta, Catanzaro, Cremona, Cuneo, Ferrara, Firenze, Foggia, Forlì, Frosinone, Genova, Grosseto, Imperia, L’Aquila, La Spezia, Lecce, Livorno, Lucca, Mantova, Massa Carrara, Milano, Modena, Napoli, Novara, Padova, Palermo, Parma, Pavia, Perugia, Piacenza, Pisa, Pistoia, Pordenone, Potenza, Roma, Rovigo, Savona, Siena, Siracusa, Taranto, Terni, Torino, Treviso, Trieste, Udine, Varese, Venezia, Verona, Vicenza.

Goethe would no doubt view as problematic this approach to science as hypothesis testing. He wrote:

Hypotheses are lullabies with which the teacher soothes his pupils to sleep. The thoughtful and faithful observer grows increasingly conscious of his limitations, for he perceives that the more knowledge extends the more numerous are the problems that emerge. ... All hypotheses get in the way of the *anatheorismos*—the urge to look again, to contemplate the objects, the phenomena in question, from all angles (von Goethe 1996, pp. 85–86).

Goethe is right. The newspaper stories, the linguistics-based story grammar, PC-ACE, the SQL queries, the regression models climax at the point of hypothesis testing. If our instruments confirm our hypothesis, there is no reason to look further, from other angles. And emphasis on verification tends to brush aside all idiosyncratic behavior, all outliers (indeed, dummied out or excluded from the analyses). Yet, idiosyncratic behavior may potentially lead to a new understanding of the phenomenon.¹⁵ Exceptions may well be the rule (Levine 1993).

At this step of knowledge production, the system of rules for the signs (mathematical and statistical signs) imposes a large number of constraints (many more are at hand, beyond outliers and influence—multicollinearity, normality, heteroschedasticity, autocorrelation—each problem with its battery of tests and solutions). Depending upon the type of statistical/mathematical model you use, you may need to impose even more rules to deal properly with your signs. Thus, if you use network models to map relationships among social actors (for which you would need an actor as subject, a process, and an actor as object, as in: fascists kill worker), you may need to add implicit social actors as objects for such intransitive processes as “strike,” “rally,” or “approve” (after all, “workers strike” against an “employer” as the implicit object). Otherwise, the network models would underestimate the amount of conflict from workers toward employers for such sentences as “workers strike” or “workers rally” (Franzosi 2004a, pp. 292–293, 2010, pp. 136–139). Similarly, “nominalized processes,” where a noun takes the role of a verb (e.g., “the beating of five workers,” where someone beats five workers), would need to be converted back into processes, to avoid underestimating the role of certain social actors and their actions, if known (Franzosi 2004a, pp. 291–292).

In any case, the network graph of Fig. 1, depicting not only the amount of violence among social actors but the direction of violence between these actors (who is violent towards whom) may well depend upon who tells the story. In this case, the story behind the graph is told by the official newspaper of the Socialist Party, the *Avanti!* The direction of the relation (if not the size of the line) may change if we used *Il Popolo d'Italia*, as source, the official newspaper of the Fascist Party. After all, story telling depends upon the tellers’ different points of view and ways of seeing, as we saw above in the section “Narrative and Sequences (and Ways of Seeing)”. Preliminary comparative findings from *Avanti!* and *Il Popolo d'Italia* data on violence show that the two papers fundamentally agree on the spatio-temporal diffusion of violence, but differ on who is to blame, as epideictic rhetoric would lead us to expect: for *Avanti!* the socialists are innocent victims of fascist aggression

¹⁵ For a critique of this sociological approach based on theory verification rather than theory generation, see Glaser and Strauss (1967, pp. 10, 40, 185–186).

while the opposite is true for *Il Popolo d'Italia*—a finding that affects the direction of the lines in the graph of Fig. 1 but not their thickness.

Similarly, the hot spots map of fascist violence and fascist institution building of Fig. 2 depends upon SQL queries that automatically extract the location of fascist actions of violence and of the founding of new *fasci*. Yet, PC-ACE coded output for many triplets show that the location, “Bissone di S. Cristina,” is coded in only 3 semantic triplets out of 9. The system of rules of good writing imposes that we do not repeat the space or time for every verb of a text. But the system of rules for that same text structured in a relational database imposes that time and space are repeated for every verb within every semantic triplet or an SQL query would miss the fact that the fascists beat up workers in Bissone di S. Cristina (semantic triplet no. 4 in PC-ACE coded output, indeed, without space) and our maps would be biased (on these points, see Franzosi 2010, pp. 134–136).

In estimating Eq. 3.1, I subject the estimates to a battery of tests. I write (Franzosi 2004a,b, p. 116):

Unfortunately, the values of leverage, Cook’s D Statistics, and Mahalanobis distance all single out the observation for Genoa as a highly influential observation (with values ten times as high or higher than the next most influential observation). Given the potentially distorting effect of influential observations on parameter estimates, Model 2 reports the estimates of Eq. 3.1, with the addition of a dummy variable for Genoa (on outliers and influence, see Belsley et al. 1980; Franzosi 1994b). The results show that the coefficient of the dummy variable for Genoa is, indeed, the most highly significant among the set of independent variables. Nonetheless, the coefficients of both variables of Eq. 3.1 remain significant and with the expected sign. The adjusted R-squared increases from .554 to .686. The values of leverage, Cook’s D Statistics, and Mahalanobis distance for the new model estimate now point to Milan as a potentially troublesome data point (a problem known as “masking”; see Atkinson 1988). Model 3 of Table 3.3 reports the estimates of Eq. 3.1 with the addition of a second dummy variable for Milan. Model 4 includes a dummy variable for Rome, “unmasked” by the estimates of Model 3 (after which, no more significant influential observations are brought out). Again, the estimates confirm the significance of the dummy variables; but the data values of Genoa, Milan, and Rome do not seem to affect the signs and significance of the substantive coefficients (percent of communes with a socialist majority in the 1920 administrative elections and number of working-class actions of conflict and violence during 1919 and 1920).

Scientists (including us, social *scientists* with a small *s*) place extra care in their work, in the painstaking fending of potential problems. Indeed, I discovered the error in coded output for the Bissone di Santa Cristina story, despite its overall semantic coherence, precisely because, alongside verification of coded output for semantic coherence, I was carrying out input versus output verification (Franzosi 1990, 2004a, b, pp. 78–79). And to guard against the eventuality of one-sided newspaper stories, I relied on different sources for data collection: the socialist *Il Lavoro* and *Avanti!* and the fascist *Il Popolo d'Italia*. Peter Medawar, 1960 Nobel laureate in Medicine, in his *Advice to a Young Scientist*, makes it clear that “In the performance of what is intended to be a critical experiment, clarity of design and *fastidiousness of execution* are the qualities to be aimed at” (Medawar 1979, p. 73; emphasis added). That term

“fastidiousness” leaves no doubt about what Medawar has in mind: “Possessing or displaying *careful, meticulous* attention to *detail; difficult to please; exacting; excessively scrupulous* or *sensitive*.” These are the dictionary’s definition (with emphasis added). In reality, how much care scientists put in their work, few really care to find out: most scientific experiments are never replicated. And the ones that are replicated, despite all this care, are often difficult to reproduce.¹⁶

In light of “the painstaking fending of potential problems” at this last step in the production of knowledge, perhaps, Randall Collins is correct when he writes (1984, p. 339): “The community of researchers is less concerned with whether a given finding is true than with whether it can pass the hurdle of a very high level of ritual distrust imposed on it.” Goethe would have agreed. After all, he wrote (von Goethe 1996, p. 34):

The contemplation of the world, with all these theorists, has lost its innocence, the objects no longer appear in their natural purity. If these learned people, then, give an account of their observations, we obtain, notwithstanding their love of truth as individuals, no actual truth with reference to the objects; we always get the taste of a strong subjective mixture.

No doubt, in this long sequence of steps, Goethe’s phenomenon risks being the innocent victim, forgotten behind a cumulative system of signs. At each step in the sequence, as we tend to look forward, rather than backward, there is a real danger that the phenomenon recedes to the background when compared to the more pressing concerns of dealing with the problems posed by each type of sign. For the journalists it is the entrapment of good writing or the pressure of the newsroom that puts a distance between the word and the world (that goes for the traditional narrativist historian as well, if we are to believe White, 1987, pp. 26–57). For the social scientists (and the quantitative historians, Bailyn 1982, p. 6) it is the entrapment of technical issues. I look at newspaper articles but wonder how I can fit the information contained in them into the categories of a story grammar. And once fitted, I extract numbers via SQL queries, numbers that I can further fit in an equation or a network model.

Indeed, where is the phenomenon behind Eq. 3.1? Have we ended up “substituting the sign for the thing?” Where are the social actors, the landed elite, industrialists, young middle-class army officers during the war, socialist working class, all strategically interacting in the pursuit of their interests? Those actors, and their interests, have disappeared behind variables— Y , X_1 , X_2 —and coefficients— β_0 , β_1 , β_2 . After all, the instrument that I developed—the computerized story grammars that are at the methodological heart of this article—reflects my original quest for the role of social actors (“*in search of the actor*,” as I have put it; 2004a, pp. 3–4). That concern for social actors and their actions pushed me to look in particular directions: in action, for instance, rather than meaning. It focused on certain questions: Who does what, pro or against whom? As such, it embraces a behavioral approach to social scientific explanation, an approach based on the belief that social reality can be measured and that what is worth measuring is behavior, i.e., what people do.

Yet, individuals do not just do (i.e., act). They think. They feel. They sense. Behind the behavioral questions—who did what, when, and where—lie other questions: Why did the fascists behave the way they did? What did they believe?

¹⁶ On replication in science, see Collins (1985), King (1995), Herrnson (1995), McCullough and Vinod (1999, 2003a,b).

What was their world view?¹⁷ If what is important is the meaning that actors attach to their actions, rather than the action per se, the instrument of measurement I developed is inadequate (see the ironic conclusion of *From Words to Numbers*, 2004a, in particular, pp. 312–315, 332–334). In search for meaning, I must look elsewhere and with a different tool. Halliday’s mental and verbal verbs, of being and sensing that I so carefully excluded from the process of measurement would have been helpful. “Seeing, knowing, sensing, believing,” is Goethe’s path to science, as he himself put it in a letter to C.D. von Buttel, May 1827 (in Heisenberg 1974, p. 133). Alas, every instrument of measurement answers certain questions and prevents looking at other questions at the same time (Franzosi 2010, p. 146).

Every instrument of measurement is deeply embedded in larger beliefs about the world (e.g., a behavioral versus a cultural approach). No doubt, “in science the way in which a question is put and the method of research employed already singles out a finite and limited field from the abundance of physical phenomena” (Heisenberg 1952, p. 74). In the end, if lucky, we find what we set out to find. Furthermore, “every scientific theory arises in a certain mental climate which implies some idea as to how the projected theory might later be applied. This background is often conditioned by the historical development of the science concerned and the author of the theory may be only vaguely conscious of it” (Heisenberg 1952, pp. 65–66). My theory (more modestly, my approach) is rooted in my doctoral training in mathematics and statistics, in the disappointment with what a variable-centered approach delivered in my dissertation on Italian strikes for the period 1950–1980, in my subsequent “search for the actor,” but in light and within the limits of a quantitative background, and, finally, in a postdoctoral year at the Center for Research on Social Organization at the University of Michigan where Tilly, Paige, Gamson, and many others were busy measuring what social actors do along behavioral models. That personal and cultural milieu deeply marked the path I was to follow for the next few decades (for that story, see Franzosi 1996, 2004a, pp. 4–5, 35–40).

Goethe himself, by making the history of colors an integral part of his approach to a science of colors, took a fundamentally historical approach to science and knowledge. And that approach led him to view science with relativistic eyes: “The conclusions of men are very different according to the mode in which they approach a science or branch of knowledge; from which side, through which door they enter” (von Goethe 1970, p. lxi). “Goethe sought meaning,” Matthaei writes (1971, p. 44); he sought what “the observer does not see ... with his eyes, but more with his soul” (in Matthaei 1971, p. 65).

Coming full circle

In 1960, Claude Lévi-Strauss published a review of the first English edition of Propp’s *Morphology of the Folktale* (1958). The Russian original had been published thirty years earlier in Moscow (Propp 1928). Propp himself tells us that story (1984, p. 67):

Morphology of the Folktale was published in Russian in 1928. ... The book, like so many others, would probably have been forgotten or

¹⁷ See, for instance, Baldassini’s work on fascists’ diaries (2002).

remembered occasionally only by specialists, but a few years after the war it emerged again. It was frequently mentioned at congresses and in articles, and it was translated into English. The cause of this renewed interest should be sought in the revolutionary discoveries made in the exact sciences through the use of much more advanced and reliable methods of research and computation. Attempts to apply similar methods extended to the humanities as well. Structural and mathematical linguistics sprang up, and other disciplines followed, poetics among them. Then it appeared that the concept of art as a system of signs, the procedure of formalization and modeling, and the possibility of using computation had been anticipated in Morphology, although at the time it was written the concepts and the terminology with which poetics operates today did not exist.

In his review of the book, Lévi-Strauss pays tribute to Propp's originality:

The most striking aspect of Propp's work is the power with which it anticipates further developments. Those among us who first approached the structural analysis of oral literature around 1950, without direct knowledge of Propp's attempts a quarter of a century earlier, recognize there, to their amazement, formulae—sometimes even whole sentences—that they know well enough they have not borrowed from him. ... These are so many intuitions, whose perspicacity and prophetic character arouse our admiration. They earn for Propp the devotion of all those who, unknown to themselves, were his followers. ... my ... reservations and ... objections ... can neither diminish Propp's tremendous merit nor contest the priority of his discoveries.¹⁸

And in science, priority of discovery is the most coveted prize, if we are to believe Merton (1973, pp. 286–289, *passim*). But Lévi-Strauss also does not pull his punches. Propp has a “lack of knowledge of the true relationship between myth and the folktale,” “appears to be torn between his formalist vision and the obsession with historical explanations,” “is the victim of a subjective illusion.” “It is not the past that he lacks, it is context.”¹⁹

If Propp does not fare well, formalism, as a school of thought, does not do better (with a self-serving comparison between formalism and structuralism; needless to say, Lévi-Strauss's structuralism emerges victorious; 1984a, pp. 179, 183). “The limitation, which we believe to be inherent in formalism, is particularly striking in the main chapter of Propp's work ...” “The inconsistency is so flagrant that Propp desperately seeks ...” “*Formalism destroys its object.*” “There is a more serious matter still ... inadequacy of formalism ...” “The error of formalism is thus twofold.” The very last paragraph of the review starts with the words: “We have denounced the error of formalism ...”²⁰

Propp did not take the critiques lightly. When the Italian translation of the *Morphology* appeared in 1966, as Propp tells us, he “gladly accepted the invitation of the Einaudi Publishers to write a rejoinder.” And if the language of argumentation is often implicitly based on a war metaphor (e.g., win or lose, defeat or kill arguments;

¹⁸ Lévi-Strauss (1984a, pp. 175–176; see also 179, 184).

¹⁹ For these quotes, see Lévi-Strauss (1984a, pp. 177, 178, 178, 178–179).

²⁰ For these quotes, see Lévi-Strauss (1984a, pp. 179, 179, 180, 181, 186, and 188).

see Lakoff and Johnson 1980, pp. 4–7), in Propp's rejoinder that metaphor is explicit: "When one is attacked, one tries to defend oneself. ... Lévi-Strauss has thrown down the gauntlet, and I am ready to pick it up. Readers of *Morphology* will thus witness our duel and will be able to determine the winner, should there be one." Lévi-Strauss would later write: "Propp responded to my discussion with an offended harangue. ... I do not wish to engage with him in a polemic on this subject."²¹

If Lévi-Strauss wrote of Propp that "he is not an ethnologist" (1984a, p. 177), Propp wrote of Lévi-Strauss that "he is a philosopher, whereas I am an empiricist, indeed an incorruptible empiricist, who first scrutinizes the facts and studies them carefully, checking his premises and looking back at every step in his reasoning."²² Propp continues:

If we are describing a series of facts and their relationships, our description will bring out what is essential in the phenomenon, and, apart from being of interest to the specialist, will invite philosophical meditations. In my book such meditations were present too, but they were hidden in the epigraphs to some chapters. Lévi-Strauss knows my work only in an English translation; the translator, however, has taken an unpardonable liberty. He missed my point and did not understand the function of the epigraphs. At first glance, they do not seem to belong to the text, so he decided that they were useless embellishments and barbarously suppressed them. Yet the epigraphs were from Goethe's works collected under the title of *Morphology* and from his journals; their purpose was to express certain things not stated in the text of the book.

In paying his tribute to Goethe, Propp concludes²³:

The highest goal of every science is to discover laws. ... The word *morphology* ... came from the writings of Goethe, who used this unifying term in the title of his works on botany and osteology. Behind Goethe's term, we can see the prospect of discovering general laws that permeate all nature. It is not by chance that Goethe went on from botany to comparative osteology. I can heartily recommend these works to the structuralists. And if the young Goethe, like his own *Faust* seated in a dusty laboratory among skeletons, bones, and herbaria, saw nothing in them except the mortal dust, the aging Goethe, a master of precise comparisons in the field of natural sciences, saw in individual phenomena the common and general principle that permeates all nature. But two Goethes, the poet and the scholar, do not exist; the Goethe of *Faust*, who longed for knowledge, and Goethe the naturalist, who attained it, are one and the same person. By starting some chapters with epigraphs, I paid homage to him. The epigraphs also emphasized that the realms of nature and human creativity are not separated. Something unites them; laws common to both can be studied by related methods.²⁴

²¹ For these quotes, see Propp (1984, pp. 68, 68) and Lévi-Strauss (1984b, p. 189).

²² For Propp's quotes in these paragraphs, see Propp (1984, pp. 68, 68, 68–69).

²³ On Goethe's influence on Propp and the formalists, more generally, see Steiner and Davydov (1977), Steiner (1980–81; 1986).

²⁴ Indeed, Propp's *Morphology* has epigraphs taken from Goethe in the Preface and in Chapters 1, 2, 8, and 9.

Propp was right. As Fink (1991) has shown, Goethe fundamentally believed that—botany, anatomy, geology (or texts ...)—one fundamental principle was at work in nature: morphology.

Finale (closing the circle)

At the heart of this article is the illustration of an approach to narrative (Quantitative Narrative Analysis or QNA) that, taking advantage of invariant structural linguistic properties of narrative, delivers a research tool of some interest to historians and social scientists for the study of the behavior of historical actors. In particular, the language of narrative is characterized by the simple SVO structure (Subject, Verb, Object), where, in narrative, both Subject and Objects (participants in Halliday's terminology) are typically actors (although Objects can also be physical or abstract objects), and verbs (processes for Halliday) are typically verbs of doing (or behaving), i.e., social actions. Each element of this basic SVO triplet can have specific modifiers, such as the number and type of actor or the time and space of action—a structure also known as story grammar—(e.g., fascists kill socialist mayor in Bissone di S. Cristina; FIAT management fires 2 workers on 2/18/1921). All narratives are characterized by invariant sequences of this simple SVO structure (with their respective modifiers), as first brought out by Propp in his *Morphology of the Folktale* (1958/1928). When implemented in a computer environment in a relational database, this structure would allow historians and social scientists to carry out large-scale socio-historical research. And the relational properties of a computerized story grammar (as implemented in PC-ACE or in any RDBMS, Relational Database Management System) make possible the statistical analysis of what are basically words, despite the complexity of the structure (Franzosi 1999, 2004a, 2010). In the end, the trick behind the alchemic transformation of words into numbers is simple: you count. But having the information to be counted conveniently parsed in appropriate categories explicitly related to one another through rewrite rules, makes counting both easy and meaningful. Indeed, when applied to some 20,000 newspaper narratives for the immediate Italian post-World War years (1919–1922), the tool helped to bring out clear patterns in these narratives, patterns broadly consistent with what we know about those turbulent years.

Yet, a handful of SQL query results and a couple of graphs hardly qualify as evidence of what QNA can deliver substantively (for substantive results based on *Il Lavoro* see Franzosi 1999, and on *Avanti!* see Franzosi 2009). By and large, the question posed in the subtitle of the article “*Can computer-assisted story grammars help us understand the rise of Italian fascism (1919–1922)?*” here goes unanswered. Similarly unanswered go methodological questions the reader may have about the ins and outs of quantitative narrative analysis: how do you actually do it? (for such a methodological treatment, see Franzosi 2010). In this spirit of Goethe's “candid revelations” (Matthaei 1971, p. 200), I must acknowledge further limits not just of the rhetorical approach of the article but of the story grammar approach to narrative the article proposes:

1. A story grammar focused on behavior cannot give a clear picture of the ideological representation of a narrative stemming from syntactic and semantic rhetorical strategies of story telling (as highlighted above in the section “Narrative and

- Sequences (and Ways of Seeing).” Nothing prevents you from tweaking a story grammar to include coding categories that would tap not only the purely descriptive clauses of a story but also description and, especially, evaluation (on using story grammars for non-narrative texts or parts of text, see Franzosi 2010, pp. 46–52). Yet, you do not need to code twenty-thousand documents to gauge the ideological stance of a newspaper ... one article may just suffice (for an example, see Franzosi 2004b). In which case, you may not even need quantitative narrative analysis. A more nuanced tool of in-depth, qualitative textual analysis, deriving from the tradition of literary criticism, would be more appropriate (e.g., discourse analysis). You need to (indeed, you must) quantify only if you are dealing with large volumes of data (Franzosi 2010, pp. 4–5).
2. As for meaning and its measurement, the road I have traveled is only a tiny step in a very long and arduous journey. Proof being Halliday’s frontispiece. When I first looked at it several years back, I did not pay much attention to its meaning. It was nothing more to me than a colorful cover jacket (a colleague, professor of linguistics, who had used Halliday’s book in his courses, confessed the same to me). It is only after reading Gage, Newton, Goethe, and others on colors and theories of colors that the full meaning and implications of Halliday’s representation became clear to me. *Nothing easy about meaning and its measurement*, it seems. “The process of measuring is a coarse one, and extremely imperfect,” as Goethe himself put it (1996, p. 67).
 3. While waiting for AI solutions to computer understanding of natural languages, parsing large quantities of text into the categories of a complex story grammar is a very labor-intensive process—which is perhaps why sociologists have mostly provided illustrative examples of the sociological use of the linguists’ intuitions on narrative, rather than full-fledged applications to explain complex social phenomena.

Yet, for all its errors, really perhaps *because* of its errors, this journey from words to numbers and from narrative to Newton, Goethe, and Propp, this journey in self-realization and self-knowledge would have been worth it. As Goethe wrote: “I wound my way through many a hypothetical error and narrow-mindedness. Yet ... I felt myself able to connect the color theory to many adjacent areas to such axioms and to align much that was distant” (in Matthaei 1971, p. 200) Perhaps more to the point, my life-long project of going from words to numbers in search of the actor, in the painstaking collection, dissection, and now, finally, analysis of thousand of narratives resonates well with Goethe’s own thinking: “No phenomenon can be explained by itself and of itself; only a number of them, when viewed collectively, and arranged methodically, end by yielding something that may pass for theory” (von Goethe 1996, p. 84).

A social scientific approach that stresses wholeness and relations will also resonate well with Goethe’s own thinking. Narrative wholes are what quantitative narrative analysis deals with. As I write in the book *Quantitative Narrative Analysis*, in comparing QNA & QCA (Qualitative Comparative Analysis): “The same mathematical set theoretic “thin thread” thus underlies both QCA and QNA. And like QCA, where “each case is examined as whole” ... QNA deals with “whole” event narratives, however incomplete these narratives may be. Both QCA and QNA, in other words, are holistic in their research strategies” (Franzosi 2010, p. 128). QNA pushes the investigator to look

at both text and context, at the text and at what goes with the text (Franzosi 2004a, p. 273). And QNA is fundamentally relational. Relational is the linguistic tool that provides a way to structure narrative into constitutive components: story grammar. Relational is the computer science data model used to store the narrative information properly structured via a story grammar: Relational Database Management System (RDBMS). Relational is the statistical tool used to analyze the information stored in an RDBMS: network models. And relational is this “way of conceiving knowledge and understanding ‘which consists in seeing connections’ (Wittgenstein, in Monk 1991, p. 451)” (cited in Franzosi 2004a, p. 273).

Indeed, taking cue from Halliday’s representation of a story grammar as participants and processes and Halliday’s pictorial rendering of different processes around a color circle, the article leads the reader through the history of the color circle as a representational device, discovering Newton’s *Opticks* along the way and Goethe’s *Theory of Colour* and the acrimonious argument that the German Master initiated against the “great mathematician,” as he referred to Newton, anticipating some of the themes later found in the human sciences in the quality versus quantity debate, in the model versus whole approaches to science. At the end of this excursus, appropriately closing the circle, we find Propp, who so forcefully acknowledged his indebtedness to Goethe for his concept of morphology; that same Propp who did the seminal work on the invariant linguistic structure illustrated in this article and at the heart of my research agenda.

There is much of current sociological relevance in Goethe’s work: from the emphasis on relations and connections, meaning and wholeness, to the historicity of knowledge, open science and self-reflexivity, the relationship between subject and object, form and ideal type, quality versus quantity—and in spite of his anti-Newtonian aversion to mathematics, Goethe was always careful to make his position clear: “I have heard myself criticized as if I were an opponent, an enemy, of mathematics in general, which in fact no one can value more highly than I” (von Goethe 1970, p. 69). He made it clear where he stood in this opposition of quality versus quantity:

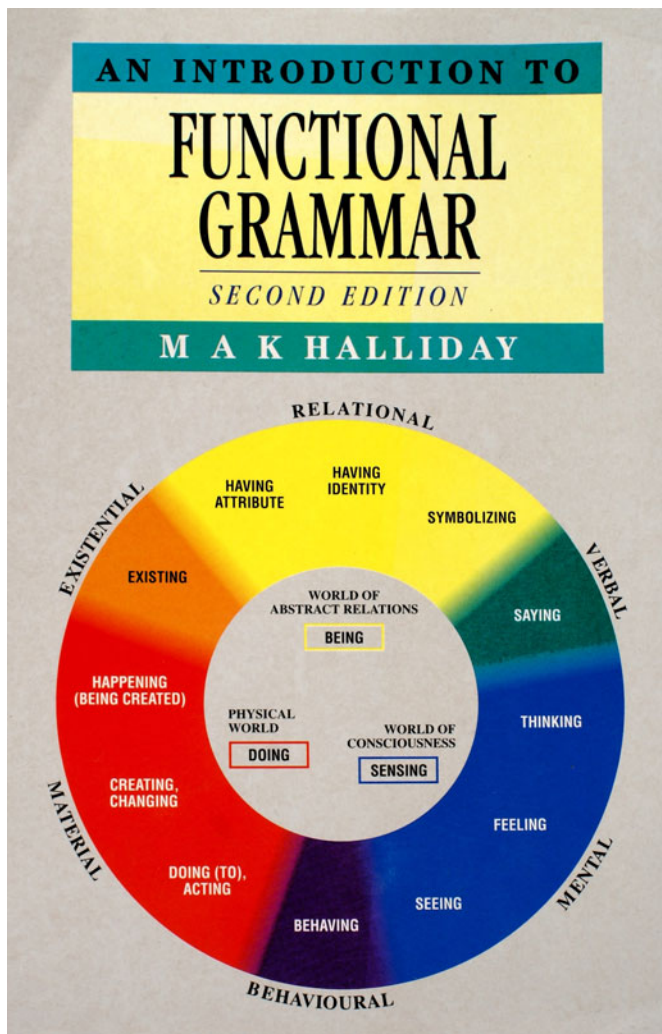
From the mathematician, we must learn the meticulous care required to connect things in unbroken succession, or rather, to derive things step by step. Even where we do not venture to apply mathematics, we must always work as though we had to satisfy the strictest of geometers (von Goethe 1996, p. 67).

All of this should have caught the attention and interest of modern sociologists—as it no doubt did both Simmel’s and Weber’s (Gonzalez García 1992). After all, we find Goethe at the beginning and at the end of Weber’s life. Marianne Weber, in her biography of her husband, writes: “The adolescent did almost no work for school, and only occasionally paid attention in class. In *Tertia*, for example, he secretly read all forty volumes of the Cotta edition of Goethe during class hours.” (1975, pp. 47–48) And in the summer of 1917, towards the end of his life, at Oerlinghausen, at Marianne’s family home, “In the afternoon, when the women were doing needlework, he would read to them poems by Stefan George and excerpts from Gundolf’s book on Goethe. Even though they did not fully understand everything, the music of his voice soothed their souls” (1975, p. 595). Perhaps, the “battle is over” between Goethe and Newton on the color theory, on the subjective and objective view in the natural sciences. But the battle between an interpretive

sociology and a quantitative one still rages, now more furious than ever.²⁵ And Simmel's and Weber's debts to Goethe have largely gone unnoticed. But that's a different story, and for another time and place, a story not for story grammars.

Acknowledgments Data collection for the data presented in this article was made possible by grants from the National Science Foundation, Reading University, University of Trento, Provincia Autonoma di Trento, and Emory University. Development of PC-ACE was made possible by grants from the Nuffield Foundation, University of Reading, University of Trento, Provincia Autonoma di Trento, and Emory University. I am grateful to Peter Bing for suggesting the link between Propp and Goethe, to Roger Friedland, Frank Lechner, Richard Robinson, and two *Theory and Society* anonymous reviewers for their comments on a previous draft, and to Gianluca De Fazio for help with data analysis.

Appendix



²⁵ This is particularly true in the femsist critique of quantitative methodologies (see the introduction to the debate in Franzosi 2004a, pp. 266–269).

References

- Abbott, A. (1983). Sequences of social events: concepts and methods for the analysis of order in social processes. *Historical Methods*, 16, 129–147.
- Abbott, A. (1990). Conceptions of time and events in social science methods. *Historical Methods*, 23(4), 140–150.
- Abbott, A. (1992). From causes to events: notes on narrative positivism. *Sociological Methods and Research*, 20(4), 428–455.
- Abbott, A. (1995). Sequence analysis: new methods for old ideas. *Annual Review of Sociology*, 21, 93–113.
- Abbott, A. (1997). Of time and place: the contemporary relevance of the Chicago School. *Social Forces*, 75(4), 1149–1182.
- Abbott, A., & Hrycak, A. (1990). Measuring resemblance in sequence data. *The American Journal of Sociology*, 96, 144–185.
- Abbott, A., & Barman, E. (1997). Sequence comparison via alignment and Gibbs sampling: A formal analysis of the emergence of the modern sociological article. In A. E. Raftery (Ed.), *Sociological methodology*, Vol. 27 (pp. 47–87). Oxford: Blackwell.
- Abell, P. (1987). *The syntax of social life. The theory and method of comparative narratives*. Oxford: Clarendon.
- Abell, P. (1993). Some aspects of narrative method. In P. Abell (Ed.), *Narrative methods*. Special issue of *The Journal of Mathematical Sociology*, Vol. 18, Nos. 2–3 (pp. 93–134).
- Abell, P. (2004). Narrative explanation: an alternative to variable-centred explanation? *Annual Review of Sociology*, 30, 287–310.
- Agar, M. & Hobbs, J. R. (1982). Interpreting discourse: coherence and the analysis of ethnographic interviews. *Discourse Processes*, 5, 1–32.
- Amrine, F. (1996a). *Goethe in the history of science, Vol.I: Bibliography, 1776–1949*. Studies in Modern German Literature, Vol. 29. Bern, Switzerland: Peter Lang.
- Amrine, F. (1996b). *Goethe in the history of science, Vol.II: Bibliography, 1950–1990*. Studies in Modern German Literature, Vol. 30. Bern, Switzerland: Peter Lang.
- Amrine, F. (1999). *Goethe in the history of science, Vol.III*. Studies in Modern German Literature, Vol. 31. Bern, Switzerland: Peter Lang.
- Amrine, F. R., Zucker, F. J., & Wheeler, H. (Eds.) (1987). *Goethe and the sciences: A reappraisal*. Boston: Reidel Publishing Co.
- Atkinson, A. C. (1988). Transformations unmasked. *Technometrics*, 30(3), 311–18.
- Bal, M. (1977). *Narratologie. Essais sur la signification narrative dans quatre romans modernes*. Paris: Editions Klincksieck.
- Baldassini, C. (2002). Fascismo e memoria. L'autorappresentazione dello squadrisimo. *Contemporanea. Rivista di storia dell'800 e del 900* 5(3), 475–505.
- Barthes, R. (1977[1966]). Introduction to the structural analysis of narratives. In R. Barthes (Ed.), *Image music text. Essays selected and translated by Stephen Heath* (pp. 79–124). London: Fontana Press.
- Bailyn, B. (1982). The challenge of modern historiography. *The American Historical Review*, 87(1), 1–24.
- Bearman, P. S., & Stovel, K. (2000). Becoming a Nazi: a model for narrative networks. *Poetics*, 27(2–3), 69–90.
- Belsley, D., Edwin K., & Roy W. (1980). *Regression diagnostics, identifying influential data and sources of collinearity*. New York: John Wiley & Sons.
- Benveniste, E. (1971[1966]). *Problems in general linguistics*. Coral Gables, FL: University of Miami Press.
- Bortoft, H. (1996). *The wholeness of nature. Goethe's way of science*. Edinburgh: Floris Books.
- Böhme, G. (1987). Is Goethe's theory of color science? In F. R. Amrine, F. J. Zucker, & H. Wheeler (Eds.), *Goethe and the sciences: A reappraisal* (pp. 147–173). Boston: Reidel Publishing Co.
- Bremond, C. (1964). Le message narratif. *Communications*, 4, 4–32.
- Bremond, C. (1966). La logique des possibles narratifs. *Communications*, 8, 60–76.
- Burke, K. (1969[1962]). *A grammar of motives*. Berkeley: University of California Press.
- Carley, K. (1993). Coding choices for textual analysis: a comparison of content analysis and map analysis. In P. Marsden (Ed.), *Sociological methodology*, Vol. 23 (pp. 75–126). Oxford: Blackwell.
- Chafe, W. L. (1970). *Meaning and the structure of language*. Chicago: The University of Chicago Press.
- Chatman, S. (1978). *Story and discourse. Narrative structure in fiction and film*. Ithaca: Cornell University Press.
- Cohan, S., & Shires, L. M. (1988). *Telling stories: A theoretical analysis of narrative fiction*. New York: Routledge.

- Collins, R. (1984). Statistics versus words. In R. Collins (Ed.), *Sociological theory* (pp. 329–62). San Francisco: Jossey-Bass Publishers.
- Collins, H. M. (1985). *Changing order: Replication and induction in scientific practice*. Chicago: University of Chicago Press.
- Corbett, M., & Lightbown, R. (1979). *The comely frontispiece. The emblematic title-page in England 1550–1660*. London: Routledge & Kegan Paul.
- Corsaro, W., & Heise, D. (1990). Event structure models from ethnographic data. In C. C. Clogg (Ed.), *Sociological methodology, Vol. 20* (pp. 1–57). Oxford: Blackwell.
- Earl, J., Martin, A., McCarthy, J. D., & Soule, S. A. (2004). The use of newspaper data in the study of collective action. *Annual Review of Sociology*, 30, 65–80.
- Ebach, M. C. (2005). *Anschaung* and the archetype: the role of Goethe's delicate empiricism in comparative biology. *Janus Head*, 8(1), 254–270.
- Fararo, T. J. (1993). Generating narrative forms. *Journal of Mathematical Sociology*, 18(2–3), 153–181.
- Fararo, T. J., & Skvoretz, J. (1986). Action and institution, network and function: the cybernetic concept of social structure. *Sociological Forum*, 1(2), 219–250.
- Fink, K. J. (1991). *Goethe's history of science*. Cambridge: Cambridge University Press.
- Franzosi, R. (1987). The press as a source of socio-historical data: issues in the methodology of data collection from newspapers. *Historical Methods*, 20(1), 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. (1990). Strategies for the prevention, detection and correction of measurement error in data collected from textual sources. *Sociological Methods and Research*, 18(4), 442–471.
- Franzosi, R. (1994a). 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. (1994b). Outside and inside the regression black box: a new approach to data analysis. *Quality and Quantity*, 28, 21–53.
- Franzosi, R. (1996). A sociologist meets history: critical reflections upon practice. *Journal of Historical Sociology*, 9(3), 354–392.
- Franzosi, R. (1998). Narrative as data. Linguistic and statistical tools for the quantitative study of historical events. In M. van der Linden, & L. Griffin (Eds.), *New methods in historical sociology/social history*. Special issue of *International Review of Social History. New Methods in Historical Sociology/Social History*, Vol. 43 (pp. 81–104).
- Franzosi, R. (1999). The return of the actor. networks of interactions among social actors during periods of high mobilization (Italy, 1919–22). In D. Rucht, R. Koopmans (Eds.), *Protest event analysis*. Special issue of *Mobilization*, Vol. 4, No. 2 (pp. 131–149).
- Franzosi, R. (2004a). *From words to numbers: Narrative, data, and social science*. Cambridge: Cambridge University Press.
- Franzosi, R. (2004b). Content analysis. In A. Bryman & M. Hardy (Eds.), *Handbook of data analysis* (pp. 547–566). Thousand Oaks: Sage.
- Franzosi, R. (2008). *Content analysis, 4 vols*. Thousand Oaks: Sage.
- Franzosi, R. (2009). *Narratives and networks: Mapping social relations during revolutionary periods (Italy, 1919–1922)*. Paper submitted for publication to *Politics & Society*.
- Franzosi, R. (2010). *Quantitative narrative analysis*. Thousand Oaks: Sage.
- Frege, G. (1948). Sense and reference. *The Philosophical Review*, 57(3), 209–230.
- Gage, J. (1993). *Colour and culture. Practice and meaning from antiquity to abstraction*. Little, Brown and Company: Boston.
- Genette, G. (1980[1972]). *Narrative discourse: An essay in method*. Ithaca: Cornell University Press.
- Gibson, D. (2000). Seizing the moment: the problem of conversational agency. *Sociological Theory*, 18(3), 368–382.
- Gibson, D. (2003). Participation shifts: order and differentiation in group conversation. *Social Forces*, 81(4), 1135–1181.
- Gibson, D. (2005a). Taking turns and talking ties: network structure and conversational sequences. *The American Journal of Sociology*, 110(6), 1561–1597.
- Gibson, D. (2005b). Concurrency and commitment: network scheduling and its consequences for diffusion. *Journal of Mathematical Sociology*, 29(4), 295–323.
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. New York: Aldine.

- Gonzalez García, J. M. (1992). *Las huellas de Fausto. La herencia de Goethe en la sociología de Max Weber*. Madrid: Tecnos.
- Halliday, M. A. K. (1994[1985]). *An introduction to functional grammar*. London: Edward Arnold.
- Heise, D. R. (1979). *Understanding events: Affect and the construction of social action*. New York: Cambridge University Press.
- Heise, D. R. (1989). Modeling event structures. *Journal of Mathematical Sociology*, 14, 139–169.
- Hermson, P. S. (1995). Replication, Verification, Secondary Analysis, and Data Collection in Political Science. *Political Science and Politics*, 28(3), 452–455.
- Heisenberg, W. (1974). Goethe's view of nature and the world of science and technology. In: *Across the frontiers* (pp. 122–141). Translated from the German by Peter Heath. New York: Harper & Row. Lecture originally published in German in 1967 in *Goethe-New Series* of the *Goethe Society Yearbook*, ed. Andreas B. Wachsmuth, Vol. 29, Weimar, pp. 27–42.
- Heisenberg, W. (1952). The Teachings of Goethe and Newton on Colour in the Light of Modern Physics. In: *Philosophic problems of nuclear science* (pp. 60–76). Translated by F. C. Hayes. New York: Pantheon Books. Lecture originally published in German in 1941 in *Geist der Zeit*, 19:261–275.
- Keller, E. F. (1983). *A feeling for the organism: The life and work of Barbara McClintock*. New York: Henry Holt.
- King, G. (1995). Replication, replication. *Political Science and Politics*, 28(3), 443–499.
- Kohler Riessman, C. (1993). *Narrative analysis*. Thousand Oaks: Sage.
- Labov, W. (1972). *Language in the Inner City*. Philadelphia: University of Pennsylvania Press.
- Labov, W., & Waletzky, J. (1967). Narrative analysis. In J. Helm (Ed.), *Essays on the verbal and visual arts* (pp. 12–44). Seattle: University of Washington Press.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.
- Levine, J. (1993). *Exceptions are the rule. An inquiry into methods in the social sciences*. Boulder: Westview Press.
- Lévi-Strauss, C. (1960). La structure et la forme. Reflexions sur un ouvrage de Vladimir Propp. *Cahiers de l'Institut de Science économique appliquée*. Series M, No. 7, 1–36. Reprinted in *International Journal of Slavic Linguistics and Poetics* 3 (1960): 122–149, under the title “L'analyse morphologique des contes russes.” Its English translation appeared in Claude Lévi-Strauss, 1976, 115–145, *Structural Anthropology: Volume 2*. Translated from the French by Monique Lay. New York: Basic Books.
- Lévi-Strauss, C. (1984a). Structure and form: Reflections on a work by Vladimir Propp. In: Vladimir Propp. *Theory and History of Folklore* (pp. 167–188). Translated by Ariadna Y. Martin and Richard P. Martin and others. Edited, with an introduction and notes, by Anatoly Liberman. Minneapolis: University of Minnesota Press.
- Lévi-Strauss, C. (1984b). Postscript. In: Vladimir Propp, *Theory and History of Folklore* (p. 189). Translated by Ariadna Y. Martin and Richard P. Martin and others. Edited, with an introduction and notes, by Anatoly Liberman. Minneapolis: University of Minnesota Press.
- Matthaei, R. (1971). *Goethe's colour theory*. Arranged and edited by Rupprecht Matthaei. London: Studio Vista.
- McCullough, B. D., & Vinod, H. D. (1999). The numerical reliability of econometric software. *Journal of Economic Literature*, 37(2), 633–665.
- McCullough, B. D., & Vinod, H. D. (2003a). Verifying the solution from a nonlinear solver: a case study. *The American Economic Review*, 93(3), 873–892.
- McCullough, B. D., & Vinod, H. D. (2003b). Econometrics and software. *The Journal of Economic Perspectives*, 17(1), 223–224.
- Medawar, P. B. (1979). *Advice to a young scientist*. New York: Basic Books.
- Merton, R. K. (1973[1957]). *The sociology of science. Theoretical and empirical investigations*. Edited and with an introduction by Norman W. Storer. Chicago, IL: The University of Chicago Press.
- Monk, R. (1991). *Ludwig Wittgenstein. The duty of genius*. London: Vintage.
- Olzak, S. (1989). Analysis of events in the study of collective action. *Annual Review of Sociology*, 15, 119–141.
- Ortiz, D. G., Myers, D. J., Walls, E. N., & Diaz, M.-E. D. (2005). Where do we stand with newspaper data? *Mobilization: An International Quarterly*, 10(3), 397–419.
- Palmer, F. R. (1994). *Grammatical roles and relations*. Cambridge: Cambridge University Press.
- Prince, G. (1973). *A grammar of stories; an introduction*. De proprietatibus litterarum; Series minor; Vol. 13. Paris: Mouton.
- Propp, V. (1928). *Morfologija skazki* [Morphology of the folktale]. Gosudarstvennyj Institut istorii iskusstv. Voprosy poetiki, vol. 12. Leningrad: Academia.
- Propp, V. (1958[1928]). *Morphology of the folktale*. Edited with an introduction by Svatava Pirkova-Jakobson. Translated by Laurence Scott. Indiana University Research Center in Anthropology, Folklore, and Linguistics, Publication 10. Bloomington, Indiana. *International Journal of American*

- Linguistics* 14 (4). *Bibliographical and Special Series of the American Folklore Society* 9. Philadelphia.
- Propp, V. (1966). Struttura e storia nello studio della favola. In: *Morfologia della fiaba. Con un intervento di Claude Lévi-Strauss e una replica dell' autore*. 201–27. Turin: Einaudi.
- Propp, V. (1984). *Theory and history of folklore*. Translated by Ariadna Y. Martin and Richard P. Martin and others. Edited, with an Introduction and Notes, by Anatoly Liberman. Minneapolis: University of Minnesota Press.
- Ricoeur, P. (1984). *Time and narrative*. Vol. 1. Translated by Kathleen McLaughlin and David Pellauer. Chicago, IL: The University of Chicago Press.
- Rimmon-Kenan, S. (1983). *Narrative fiction: Contemporary poetics*. London: Methuen.
- Robbins, B. D., & C. Holdrege (Eds.). (2005). Goethe's delicate empiricism. *Janus Head*, Vol. 8, No. 1.
- Seamon, D. (1998). Goethe, nature, and phenomenology: An introduction. In D. Seamon & A. Zajonc (Eds.), *Goethe's way of science* (pp. 1–14). Albany: State University of New York Press.
- Seamon, D., & Zajonc, A. (Eds.). (1998). *Goethe's way of science*. Albany: State University of New York Press.
- Skvoretz, J. (1993). Generating narratives from simple action structures. *Journal of Mathematical Sociology*, 18(2–3), 135–140.
- Skvoretz, J., & Fararo, T. J. (1989). Action structures and sociological action theory. *Journal of Mathematical Sociology*, 14(2–3), 111–137.
- Skvoretz, J., & Fararo, T. J. (1996). Generating symbolic interaction: production system models. *Sociological Methods and Research*, 25, 60–102.
- Smith, T. (2007). Narrative boundaries and the dynamics of ethnic conflict and conciliation. *Poetics*, 35, 22–46.
- Steiner, P. (1980/81). Three metaphors of Russian formalism. *Poetics Today* 2(1b), 59–116.
- Steiner, P. (1986). *Russian formalism: A metapoetics*. Ithaca: Cornell University Press.
- Steiner, P., & Davydov, S. (1977). The biological metaphor in Russian formalism: the concept of morphology. *Substance*, 6(16), 149–158.
- Todorov, T. (1977[1971]). *The poetics of prose*. Oxford: Basil Blackwell.
- Todorov, T. (1981[1968]). *Introduction to poetics*. Sussex: The Harvester Press.
- Todorov, T. (1990[1978]). *Genres in discourse*. Cambridge: Cambridge University Press.
- Tomashevski, B. (1965[1925]). Thematics. In L. Lee, & M. Reis (Eds.), *Russian formalist criticism: Four essays* (pp. 61–95). Lincoln: University of Nebraska Press.
- Toolan, M. (1988). *Narrative: A critical linguistic introduction*. London: Routledge.
- Touraine, A. (1984). *Le retour de l'acteur: Essai de sociologie*. Paris: Fayard.
- Trew, A. (1979). Theory and ideology at work. In R. Fowler, R. Hodge, G. Kress, & A. Trew (Eds.), *Language and control* (pp. 94–116). London: Routledge & Kegan Paul.
- van Dijk, T. (1980). Story comprehension: an introduction. *Poetics*, 9, 1–21.
- van Dijk, T. (1983). Discourse analysis: its development and application to the structure of news. *Journal of Communication*, 33, 20–43.
- von Goethe, J. W. (1970 [1840]). *Theory of colours*. Translated from the German with notes by Charles Lock Eastlake. Introduction by Deane B. Judd. Cambridge, Mass.: The M.I.T. Press.
- von Goethe, J. W. (1988). *Scientific studies*. Edited and translated by Douglas Miller. Vol. 12, Suhrkamp Edition in 12 Volumes. New York: Suhrkamp Publishers.
- von Goethe, J. W. (1996). *Goethe on science. An anthology of Goethe's scientific writings*. Selected and introduced by Jeremy Naydler. With a Ffward by Henri Bortoft. Edinburgh: Floris Books.
- von Goethe, J. W. (1998). *Maxims and reflections*. Translation by Elisabeth Stopp and introduction and notes by Peter Hutchinson. London: Penguin Books.
- Weber, M. (1975[1926]). *Max Weber: A biography*. Translated from the German and edited by Harry Zohn. New York: Wiley.
- White, H. V. (1987). *The Content of the form. Narrative discourse and historical representation*. Baltimore: The Johns Hopkins University Press.

Roberto P. Franzosi earned a B.A. in Literature in 1975 from the University of Genoa, Italy, and a Ph.D. in Sociology from Johns Hopkins University in 1982. He is Professor of Sociology and Linguistics at Emory University.. Franzosi's main interests are in the study of conflict and violence (Italian postwar strikes, rise of Italian fascism, lynching in American Jim Crow South). He also has a long-standing interest in language and measurement of meaning. He is the author of several journal articles and books: *The Puzzle of Strikes: Class and State Strategies in Postwar Italy* (Cambridge University Press, 1995), *From Words to Number: Narrative, Data, and Social Science* (Cambridge University Press, 2004), and *Content Analysis* (Sage, 2008), *Quantitative Narrative Analysis* (Sage, 2010).