This is the accepted manuscript. Find the published article at

https://doi.org/10.1016/j.socnet.2025.07.001

@article{STUHLER2025199,
title = {The cultural construction of personal relationships},
journal = {Social Networks},
volume = {83},
pages = {199-214},
year = {2025},
issn = {0378-8733},

The cultural construction of personal relationships

Network analysis aspires to be "anticategorical," yet its basic units—relationships—are usually readily categorized entities with labels like "friendship," "love," or "patronage." In this way, a nontrivial cultural typification underlies the very building blocks of most network analyses. Despite work showing that a specific "type of tie" often stands in for quite heterogeneous empirical phenomena, this typification is seldom challenged in research practice. This article expands on recent efforts to more adequately theorize ties by further developing and arguing for the concept of *relationship frames*—cultural models that stabilize relational expectations. I suggest that such frames are rooted in regularities in the duality of *dyad* and *content*. Building on this idea, I develop a formal notion of *frame ambiguity*—the extent to which the actions and symbols designating a relationship index a variety of frames rather than just one. Putting these ideas to analytical use, I inductively identify relationship frames from the content of 1.2 million relationships between characters in fiction writing. I conclude with an exploratory investigation of some of the conditions under which ties in fiction writing display variation in frame ambiguity.

Keywords: relationship frames, types of ties, ambiguity, duality, text analysis

Introduction

Formal network analysis typically takes tie meaning as a fait accompli. Relationships between actors come as "types of ties," with attached labels like "friendship," "love," or "patronage," and the primary focus lies on studying the arrangement of these ties.

Meanwhile, qualitatively oriented scholars studying social relationships tease out complexities and ambiguities embedded in them (Small 2017; Swidler 2001); especially, the relational work perspective has shown us how actors navigate and strategically manage heterogeneous tie meanings (see especially Bandelj 2020; Mears 2015; Zelizer 2000).

In an effort to bridge this gap, recent conceptual work has pointed to the potential of "frames" as a concept for theorizing tie meanings more adequately (Fuhse 2021; Lizardo 2024). A central advantage of conceptualizing tie semantics with frames is that it allows us to connect them to the level of practice and cognition and, thus, to the rich theoretical apparatus developed for theorizing these dimensions. Yet, while the frame concept is theoretically useful, it has largely eluded measurement and remained relatively absent from formal empirical studies, with McLean's (1998, also see 2017) work on patronage-seeking letters being a notable exception.

In this paper, I develop a notion of relationship frames, largely following the conceptual lead of Fuhse (2021). Specifically, I consider relationship frames as latent cultural constructs that bind expectations about what practices and emotions bundle together in social relationships. To study these constructs, I take inspiration from the vast literature that builds on the principle of duality to study cultural forms (see esp. Mohr and Duquenne 1997; Breiger 2000; Mohr and White 2008; Martin and Lee 2018; Breiger and Mützel 2020; Basov et al. 2020; Breiger and Wagner-Pacifici 2023). I argue that at the most basic level, relationship frames are grounded in the *duality of dyad and content*. Relationship frames emerge as actors observe practices co-occur in relationships — in both real life and cultural works.

I then build on these foundations to propose both a concept and a measure of *frame ambiguity* — the extent to which the practices and symbols designating a relationship index a variety of frames rather than just one. Ambiguity in relationships has long fascinated network analysts but has proven challenging to measure formally (see, esp. Leifer 1983, 1988; Burt and Schøtt 1985; Padgett & Ansell 1993; Park 2020; White 1992, pp. 85-87). The notion of frame ambiguity put forward here offers a way to think about the ways in which multiple meanings are at play within a relationship.

Translating these theoretical ideas into empirical strategy, I show how relationship frames can be empirically identified and measured in textual data by studying the content of more than a million relationships between characters in a sizable corpus of fiction writing. Using a novel natural language processing approach, I extract "semantic triplets" (Franzosi 1989), that is, composites of [subject][verb][object] from the novels. I then represent every relationship between two characters as sets of transitive verbs ("kiss," "believe-in," "hate," and so forth): those sent from character A to character B and those sent from character B to character A. Instead of defining relationship frames a priori, I demonstrate how mixture models can be leveraged to inductively uncover relationship frames as defined above, that is, as latent variables that incorporate information as to which actions and emotions typically go together in relationships. Through this lens, any dyad can be represented as a

distribution over a set of different relationship frames, while frame ambiguity is operationalized as the uncertainty in this distribution.

My empirical analysis then explores the internal structure and patterning of the relationship frames in fiction writing. Among other things, I find that love- and courtship-related frames are those with the strongest reciprocity norms. Frame ambiguity is in part determined by character importance, with highly central characters not merely drawing on a larger variety of frames in their own actions but also being treated that way by other characters. Furthermore, I show that interactions among male characters display less frame ambiguity than those that involve women.

The contributions of this paper are then threefold. Conceptually, I take up and concretize the notion of relationship frames. I build on this conceptualization to develop a notion of frame ambiguity, a new way of thinking about the ways in which interaction in ties draws on multiple meanings. Methodologically, I propose a mixture modeling approach for operationalizing both frames and frame ambiguity. Empirically, I offer a large-scale analysis of the frame structure of interaction in literary relationships.

Theory

Types of ties

In their seminal paper that introduced blockmodels, White, Boorman, and Breiger (1976) argued that an analysis of social structure should be based on "concrete, observable interactions" (p. 731) and that, therefore, contemporaneous, "largely categorical descriptions of social structure [had] no solid theoretical grounding" (p. 732). This argument was pitted against both theoretical approaches simply "positing categorical aggregates (e.g., 'functional subsystems,' 'classes')" but also contemporaneous empirical work satisfied with "cross-tabulating individuals according to their attributes (e.g., lower-middle-class white Protestants who live in inner city areas and vote Democrat)." (p. 733) Categories that describe social structures, in other words, should not be presupposed; instead, they should emerge inductively from identifying regularities in a lower-level data structure, notably relationships among individuals. As Emirbayer and Goodwin (1994) later eloquently put it: network analysis, if taken seriously not just as a method but in its theoretical program, follows an "anticategorical imperative." While not entirely denying the effectiveness of categories, it should reject "the primacy of attributional categories and other substantives" (Emirbayer 1997, p. 298).

And yet, formal network analysis faces a dilemma. Its core units — relationships — are usually measured in categorized form. In this way, a nontrivial cultural typification of relational meaning underlies the very building blocks of most network analyses (similarly argued by DiMaggio 1993; Mische and White 1998). When eliciting "types of ties" with name

generators, for instance, we effectively presuppose that there is a well-defined and widely shared understanding of such *types*. This may be unproblematic in cases where our goal is to elicit a discrete biological, highly institutionalized, or even legally codified state of a relationship (e.g., marriage, parenthood, business partnerships). However, research has shown that for many of the more informal, culturally contingent relationship types network analysts study, this presuppositions is less justified (see Gondal 2022; Fine & Kleinman 1983): people have different ideas about what constitutes friendship (Fischer 1982a; Kitts and Leal 2022), what it means to discuss important matters (Bearman & Parigi 2004), or to love someone (Yeung 2005). Relationships, it turns out, are not always easy to categorize.

A significant advancement in formal network analysis has been to move below the level of abstraction implied by types of ties toward the study of relational *events* (Butts 2008; Butts et al. 2023; Bianchi et al. 2024). This literature investigates discrete interaction events that take place between actors or other social entities. These events are time-stamped and often directed. Examples include radio communication (Butts 2008), acts of violence (Niezink & Campana 2022), online interaction (Vu, Pattison, and Robins 2015), or financial transactions (Bianchi & Lomi 2023). This research has proven highly successful at capturing what Butts and colleagues call "micro-temporal mechanisms" (2023). For instance, turn-taking in conversation tends to be highly patterned. This approach can also be integrated with more classic network analyses by investigating how such micro-mechanisms depend on specific network configurations (see, e.g., Gibson 2005). Beyond opening paths novel for empirical analysis, this literature also makes an important conceptual advance, for it attunes us to the distinction between relational “states” and relational “events,” which is, as will become evident, in many ways integral to the ideas outlined below. Yet overall, this research paradigm has been less focused on questions concerning the semiotics of relationships and how micro actions or events may be used to index or initialize “states” like friendship or love (for work precisely at this intersection, see McFarland et al. 2013, 2024; McFarland and Wolff 2022; Palotti, Weldon, and Lomi 2022).

More so than the formal paradigm, qualitative research on social relationships has been attuned to the ambiguity and complexity of ties and their meanings. Perhaps most notably, the relational work perspective has shed light on how actors navigate and strategically manage the coexistence of divergent tie meanings as well as the tensions among them (see especially Bandelj 2020; Zelizer 2000). Actors may choose relational practices that are meant to uphold interpretive misalignments of what a relationship is about. Mears (2015), for instance, shows how party promoters’ ability to exploit women depends on their ability to obscure the transactional nature of their relationships by framing them as leisure and friendship. Similarly, in what Rossman calls “obfuscated exchange” (2014, 2018), actors engage in a variety of practices to obscure disreputable economic transactions to one another but also to third parties. At other times, actors may engage in “blurring practices” that “create a zone of indeterminacy” (Bandelj, p. 258) to avoid a relationship from being easily categorizable. In their study on the relational work of Swedish horse farmers, for instance, Cederholm & Akerstrom (2016) find that people make heavy use of “practices that maintain indistinct boundaries between different types of relationships.” On a more general

note, Tavory (2009) reminds us that many relational practices like flirtation or diplomatic talk aim to hold ties in an ambiguous state and prevent role-set categorization. And Small's (2017) qualitative study on whom people share intimate information with directly aims at formal network analysis' most prominent instrument for eliciting social ties — the GSS' "important matters" question. While this question was generally presumed to elicit "reasonably strong ties" (see Marsden 1987, p. 123), Small shows that the practice of sharing intimate information does not neatly map onto our concept of strong ties. In fact, many people share such information with strangers they barely know.

To be sure, the notion of types of ties has been highly productive for formal network analysis. However, its relationship to the empirical reality of people's relationships is at times a precarious one in that many relationships do not conform to these types as we, the researchers, conceptualize them; but also in that our measurement is impaired by the fact that people have heterogeneous ideas about what constitutes a type. Even for seemingly concrete types of ties like "having sex," this may be truer than is convenient for formal network analysis (see Sanders and Reinisch 1999, Sanders et al. 2010); but it is likely to be especially true for types of ties that are relatively abstract cultural constructs like friendship or love.

Relationship frames and the duality of dyad and content

How can we move beyond a categoric, "bare-bones" (Mische and White 1998, p. 695) conception of ties and networks toward one that takes seriously how relational meaning is constructed in interaction and discourse — that recognizes, in other words "that networks and culture are mutually constitutive" (Pachuki and Breiger 2010, p. 209). One recent strand in the theorization of relationships has emphasized the analytical utility of the concept of "frame" (Fuhse 2021; Lizardo 2024). Lizardo's notion of frames draws on cognitive psychology (notably Barsalou 1992). Here, a frame is an attribute-value structure that represents a higher-order concept. Fuhse's notion of relationship frames, on the other hand, is rooted in symbolic interactionism. I primarily follow the latter here, though I do not consider the two perspectives as necessarily incompatible.

According to Fuhse (2021, pp. 191-198; also see Fuhse 2013, 2009), relationship frames are cultural blueprints that bundle and stabilize relational expectations in dyads. This definition borrows from Goffman (1974) who conceptualized frames as cognitive structures that organize experience and help individuals make sense of situations. Frames like "joke," "work," or "play" provide answers to the fundamental question, "What is it that's going on here?" They let people interpret and develop shared understandings of others' actions as well as guide their own actions accordingly. Analogously, Fuhse argues that relationship frames are what helps individuals make sense of "what goes on" in social relationships. "Friendship," "love," "patronage," or "collegiality" all prescribe bundles of expectations as to what is appropriate behavior within a relationship. They help reduce uncertainty between

ego and alter, for they not only guide ego's actions but also facilitate the interpretation of alter's actions. Established relationship frames are "relational institutions" (Fuhse 2021, p. 54) in the sense that they exist relatively independently from any particular social relationship. Norms around them are constructed in movies, novels, self-help books, social media, and so forth.

What is the structure of a relationship frame? To approach this question, I argue that we can build on the concept of *duality*. Originally used to theorize (Simmel 1890[1908]) and formally study (Breiger 1974) the co-constitution of individuals and collectivities, a vast body of work has shown that the concept can be productively generalized for the analysis of cultural structures (for programmatic contributions or reviews of this literature, see Breiger 2000; Martin and Lee 2018; Breiger and Mützel 2020; Basov et al. 2020; Breiger and Wagner-Pacifici 2023). The pioneer of this line of work is John Mohr (see esp. Mohr and Duquenne 1997; Mohr 1998; Mohr and White 2008). As DiMaggio (2022) succinctly put it a reflection on Mohr's work, the core insight behind his use of the duality principle was that "cultural elements at different scale are mutually constitutive, with small meanings combining to constitute larger cultural forms, which in turn shape meaning and interaction at the micro level." (p. 1) In this way, for instance, Mohr showed that moral orders are organized around "discourse roles" — a relatively small set of "structurally equivalent kinds of actors" (1994, p. 355; Mohr and Lee 2000; see Stuhler 2021 for a reconceptualization of this concept). These roles are rooted in regularities in the duality of identities and the practices targeting them.

Analogously, I argue that at the most basic level, relationship frames are also grounded in a duality: the *duality of dyad and content*. There is considerable heterogeneity among social relationships, but the content of any specific dyad is far from a random draw from the set of possible relational practices and emotions available. Instead, the heterogeneity of relationships is highly structured in that the same content items typically bundle together. A relationship that draws on the relationship frame "love," for example, might involve "kissing," "missing," and "hugging" but is less likely to involve "hiring." The broad term "content" is chosen deliberately to subsume the behavioral, cognitive, and affective dimensions of relationships.[1]

Relationship frames may be seen as "content domains" (Burt and Schøtt 1985): they emerge as actors recognize specific practices, emotions, and cognitions co-occur in relationships — in both real life and cultural works.[2] The most institutionalized relationship frames have

[1] The term "event" would have been an alternative here. However, given how it is commonly used, most sociologists will find it unintuitive to subsume cognitive processes under it.

[2] Burt and Schøtt (1985) describe precisely this as "coincidence relations" between contents. Sets of contents that coincide frequently and are thus semantically "substitutable" form "content domains." Their use of the term "substitutable" strikes me as somewhat unintuitive, for contents may be perceived to co-appear frequently without being plausibly describable as "substitutable." This may derive from the fact that the "contents" discussed in their study are not individual practices (as is the case here) but abstractions that come close to the notion of types of ties (friendship, advice, intimacy). The difference in level of abstraction

names like “friendship” or “love,” and once a frame is named, its structuring force is likely to intensify. However, here, too, following the anticategorical imperative and in line with the arguments made above, we should not make the mistake of assuming that these labels necessarily provide a good approximation of the heterogeneity in actual social relationships’ content. Instead, as I will demonstrate in the analyses below, it is worth studying dyad content inductively.

Frame ambiguity

The core advantage of this conceptualization is that it allows us to theorize more adequately the relationship between different kinds of relationships and *practice*. While people can and do, of course, label their relationships, frames are ultimately latent cultural constructs that cannot be observed in toto in a given tie. Instead, actions and symbols *cue* relationship frames (cf. McLean 1998). The notion of relationship frames, defined via the duality of dyad and content, then allows us to think more systematically about a quality that has long fascinated network analysts but has proven challenging to track formally: the *ambiguity* of social ties.

As discussed above, social relationships are typically more ambiguous than the notion of types of ties would have us believe, and qualitative research has pointed us to various ways in which people deal with multiple tie meanings. However, it would be false to say that this has gone unnoticed by formally oriented scholars (Pachuki and Breiger 2010). Most notably, Padgett and Ansell’s (1993) famous study of Cosimo de Medici theorized the strategic advantages of maintaining ambiguity in ties. A central factor of his rise to power was his ability to act credibly in a highly multivocal, sphinxlike fashion. Their analysis drew on and extended Leifer’s concept of “robust action” (1983), which he later termed “local action” (1988). Leifer argued that, in certain settings, desirable roles are attained not by laying claims via unilateral role behavior but by engaging in an ambiguous, playful mode of action that involves testing and waiting. If both ego and alter engaged in “local action,” this effectively suppresses role differentiation (relatedly, see Blau 1964 on restraint in courtship, pp. 76-87). White later transferred this idea into a concept of structure under the term “Leifer tie” (1992, pp. 85-87): ties with ambiguous, unestablished role differentiation. Another important line of thinking concerns the realization that uncertainty in tie meanings can stem from different sources. White and colleagues distinguish “ambiguity” from what they call “ambage” (White 2008, pp. 57-59; White, Godart, and Thiemann 2013): whereas the former, according to their definition, is about uncertainty in cultural rules, ambage concerns uncertainty regarding whether the enactment in a specific instance will correspond to these rules (see esp. White 2008, p. 58; White, Godart, and Thiemann 2013, pp. 138-139; for a discussion of the centrality of this distinction for White’s work, also see Schmitt and Fuhse 2015, pp. 152-153).

---

between “contents” and “content domains” is thus relatively small. Besides this, however, many of the arguments in the article align closely with those put forward here (also see Burt 1982).

While these works are considered classic readings by network scholars, there have, to date, been few attempts to directly operationalize or formally define tie ambiguity (though Park's 2020 study on role ambiguity). This presumably has much to do with the fact that measuring relationships via types of ties, as we usually do, effectively eradicates ambiguity — at least at the level of the individual tie itself. Ambiguity can then still be identified at the level of *network position*. For instance, Padgett and Ansell (1993) show how the Medici occupied a position that was structurally ambiguous in the sense it was situated between different factions. The paper's brilliance, of course, lies in that it succeeds at arguing that the Medici kept their ties ambiguous, but this is achieved via qualitative, historical description, rather than a formally defined notion of ambiguity.

Building on the previously established notion of relationship frames, I propose thinking of ambiguity in relational meaning as *frame ambiguity*. I define frame ambiguity as the extent to which the practices and symbols designating a relationship index a variety of frames rather than just one. This idea becomes formally tractable through the notion of relationship frames and their indexical, non-deterministic relationship to practices. Specifically, we can formally assess whether the actions in a relationship unambiguously cue a specific frame, or whether they point to multiple possible frames.

To illustrate this, consider an account of a relationship that involves the following: kissing in public, verbal confession of love, binge-watching TV shows together, sexual intimacy, cohabitation, making long-term plans, celebrating anniversaries, and posting photos with romantic captions. Most people would recognize these elements as collectively indexing a "romantic love" relationship frame. Meanwhile, another relationship might involve sexual intimacy, binge-watching TV shows together, sharing memes, making out publicly at parties, but lack most of the elements referenced above. In this case, we see more frame ambiguity, with some practices (sexual intimacy) indexing "romantic love" but others (sharing memes, binge-watching TV) might point more to other frames.

What is important here is that frame ambiguity doesn't just arise from a relationship having practices that index different frames. Practices don't just differ in *what* frames they index but in *how strongly* they index particular frames and, hence, the extent to which they resolve or maintain ambiguity in a relationship (on this point, see Burt and Schøtt 1985, pp. 298-299). We might call this indexicality. For instance, posting photos with romantic captions is by no means a necessary constituent of a romantic love relationship, but it is unambiguously associated with romantic love. Meanwhile, making long-term plans together is something one might also associate with a friendship, and binge-watching TV shows together is something we might do with a romantic partner but also with a roommate or even an acquaintance. This also illustrates how exclusivity with which a practice is associated with a specific frame is not necessarily the same thing as the importance of a practice for a relationship.

Frame ambiguity as defined here comes close to White's notion of "ambage" (2008). As he writes, "one test [for ambage] is correct prediction of the appearance of a tie of specified sort in some action context." (p. 58) Meanwhile, this conceptualization, for now, does not incorporate uncertainty in cultural rules about the relationship between practices and frames — a limitation I will take up upon up in my discussion. It is also worth noting that frame ambiguity is related to but distinct from multiplexity. Multiplex relationships are generally understood to contain multiple meanings in the sense that they contain multiple types of ties (see, e.g., Verbrugge 1979; Fischer 1982b; Gondal 2022). For example, two people might be friends as well as coworkers. This common definition is premised on a binary conceptualization of the ties on each layer (though statistical associations between different types of ties can be explored to study the meanings of a specific type of tie, see e.g., Burt and Schøtt 1985; Burt 1997; Cross, Borgatti, and Parker 2001; Gondal 2022; Yeung 2005). When a tie is multiplex, multiple meanings are at play, but the conventional interpretation is additive: for instance, ego and alter are both friends and coworkers. Hence, multiplexity is often interpreted as a measure of involvement in a tie (Gondal 2022, p. 211), rather than ambiguity. By contrast, the concept of frame ambiguity focuses on the indexical nature of relational practices — that is, how specific actions cue multiple possible interpretations of a relationship. This allows for a more dynamic, practice-oriented perspective on relational meaning, especially in cases where relationships are unstable, contested, or emergent.

Relationships in fiction writing

As I argued above, relationship frames are cultural constructs. They are performed in actual practice between people and learned by observing and interpreting such practice. At the same time, relationship frames are constructed and communicated through cultural works. For example, romance novels play a key role in shaping and disseminating cultural expectations about appropriate behavior and communication in romantic relationships (Radway 1991; Illouz 2012).

In this paper, I focus on depictions of ties in cultural works. Specifically, I study the construction of relationships in fiction writing. My analyses are based on 97,632 distinct character networks that were extracted from a corpus of novels published between 1850 and 2010. By their nature, these data do not reflect *real* relationships among people. Instead, they capture how relationships have been *described* historically. In other words, they constitute "public stories" that provide cultural scripts for social relationships (Jamieson 1988; Fuhse 2021, p. 174-175; de Nooy 2005).

The textual basis for these data is the NovelTM corpus (described in detail by Underwood 2020), which is, to my knowledge, the largest and most comprehensive collection of English-language fiction writing to date. It has served as a resource for various "distant reading" projects (Underwood 2019, Piper 2018; on distant reading, also see Moretti 2013). The NovelTM is based on data collection by the HathiTrust Digital Library. It constitutes an

Anglocentric representation of the literary field. Although the collection includes works originally written in languages other than English, these works must have met the somewhat higher hurdle of being deemed worthy of translation by a publisher and considered valuable enough for purchase by a U.S. library in order to become a part of the corpus. Additionally, it's important to note that the data reflect the book-buying decisions of university libraries, which tend to align with the reading preferences of a more educated audience.

The networks were extracted from the raw text through a combination of methods. The process has been described extensively by Stuhler (2024). Here, I will give a summarization of the workflow, which is illustrated in Figure 1. First (Panel A), the data were processed using the spaCy language pipeline (Honnibal et al. 2023). A language pipeline is a model that processes and annotates raw text in a series of common steps that are useful for further analysis: the text is *tokenized* (i.e., split into "words") and segmented into sentences. *Lemmatization* reduces each word to its base form (e.g., "asked" becomes "ask"). In *part-of-speech tagging*, each token is annotated with its part of speech (POS; e.g., "noun," "verb," "adjective," etc.). Finally spaCy also performs *dependency parsing*. Dependency parsers are models that predict the syntactic relationships between the lexical units within a sentence. Each sentence is annotated with a "dependency tree" — a directed, labeled graph that specifies which tokens are related to which other tokens through which syntactic relationships. For instance in the sentence "Then she killed him." the token "she" is the nominal subject of "killed" while "him" is the direct object of "killed." This labeling happens according to a specific dependency grammar. Dependency grammars specify the set of possible relations as well as when these are applicable, but also define a series of constraining properties for the graph. In the grammar used by spaCy (ClearNLP, see Choi and Palmer 2012), for instance, each sentence has one "root" node with in-degree 0 (usually the main verb of the sentence), all other nodes have an in-degree of 1, the dependency tree is fully connected, and there are no cycles. For comprehensive introductions to this technique, I refer to the work of de Marneffe and Nivre (2019) as well as Jurafsky and Martin (2024).

These tagged and parsed data provide the basis for a series of subsequent tasks. The first is extracting actor–action–recipient triplets that link pairs of entities in each text (Figure 1, panel B). Social scientists have long engaged in the extraction of semantic triplets from textual data (see, e.g., Franzosi 1989, Carley 1993, Roberts 1997, Tilly 1997). Traditionally, this extraction was done by hand. Only more recently have researchers begun automating the process. Thus far, most social-science studies in this area have relied on querying dependency trees (see, e.g., Mohr et al. 2013, Sudhahar et al. 2013, van Atteveldt et al. 2017, Goldenstein and Poschmann 2019, Knight 2022, Bellodi 2025). It is worth noting that NLP holds a variety of related approaches for achieving similar ends (e.g., semantic role labeling or SRL, see Gildea and Jurasky 2002; open information extraction or OIE, see Etzioni et al. 2008; frame semantic parsing or FSP, see Das et al 2014).

Here, I use the semgram R package (Stuhler 2022a; version 0.1.1) which implements an extensive set of query rules for texts labeled with the ClearNLP dependency scheme. The

queries consider both the assigned parts of speech as well as the labeled syntactic relations between the tokens (see Stuhler 2022b for details). In this way, for instance, the sentence "She killed him." is decomposed into a [she]-[kill]-[him] triplet. Yet, the advantage of the package is that it facilitates extraction for sentence structures that are syntactically more complex: for instance, passive constructions like "He was killed by her" are also parsed to a [she]-[kill]-[him] triplet, and compounds between verb and preposition are considered, such that "She talks to him." becomes [she]-[talk-to]-[him] (for the full set of accounted for syntactic patterns, see Appendix A of Stuhler 2024). I will refer to transitive verbs that relate characters as "actions," yet it must be noted that such verbs do not exclusively connote social actions in the Weberian sense. They can also stand for cognitive processes ([She]-[think-of]-[him]) or emotions ([She]-[hate]-[him]). An alternative term could have been "relational events," but this might have been at least as equally unintuitive in the context of emotions and cognitive processes. Manual evaluation suggests that 95.3% of the triplets that were extracted are correct (see Stuhler 2024, Appendix A, p.6 for further details).

Finally, in order to aggregate the triplets into meaningful characters networks, another major task must be completed: coreference resolution (Panel C of Figure 1). Consider that many of the time a character is referenced, this will not happen via a mention of their name but via pronouns or other phrases ("our father"). Therefore a model is needed that can detect which tokens or phrases refer to the same entities. For this, I build on the BookNLP python pipeline, which contains a coreference module specifically designed for completing this task with fiction texts (for details, see Bamman 2023). Additional details and an extensive set of validation analyses on this methodology are provided in a recent contribution by Stuhler (2024).[3]

Once this step is completed, the information from the triplets and the coreference resolution can be combined to form a character network for each book (Panel D in Figure 1). Each relationship consists of two action sets: those in which character A was the actor and character B was the recipient (i.e., the triplets [character A]-[action]-[character B]) and those in which character B was the actor and character A was the recipient (i.e., the triplets [character B]-[action]-[character A]). We can consider these sets as a formal representation of the idea of relational "action profiles" (Martin 2009, pp. 11-12, p. 21). To illustrate this, consider Panel E of Figure 1, which shows the relationship between the characters John and Gerda in *The Hollow* by Agatha Christie, by many accounts the best-selling fiction author ever. The two are husband and wife. John is a successful and physically attractive doctor while she is an insecure homemaker. Their marriage is strained by John's emotional distance. John is at times paternalistic towards his wife. Gerda admires John and is desperate to please him, while John is involved with a mistress and appears to have interest in a former lover. Partway through the novel, John is fatally shot and the story revolves

[3] This includes, among other things, validation of action extraction (Appendix A), validation of character gender prediction and coreference resolution (Appendix M), details on measures of character importance (Appendix N), and an analysis of reference forms by character gender (Appendix O). The results of these analyses are applicable for the analyses conducted here.

around determining who committed the crime. Gerda is a prime suspect at first but is quickly exonerated. Eventually — spoiler — she is revealed as the murderess. The novel builds its main tension out of the question of whether Gerda could have murdered her husband and therefore contains plenty of descriptions of the interaction dynamics and sentiments in the relationship — a deliberate balancing of ambage and motive by the author, we might say in White's terms (see 2008, p. 296). The example is also illustrative in that most relationship descriptions in the corpus derive from everyday genre fiction rather than from the kind of critically acclaimed works of the literary canon that might first come to one's mind.

Operationalized through the process described above, the relationship is represented as a set of 19 Gerda → John and 30 John → Gerda actions, with 35 unique features. In analogy to text analysis approaches that model documents as "bags of words," we might say that relationships here are "bags of actions." An alternative way of thinking about the structure of these data that network scholars are likely more familiar with is to consider them as weighted, highly multiplex networks. Multiplexity is at play here not merely because the relationship consists out of 35 unique elements, but because we simultaneously learn that it does *not* include the several thousand other possible verbial descriptors that it could have contained. Furthermore, each layer is weighted because relationships can contain the same element several times (e.g., "love" occurs four times).

There are several important aspects that this illustration highlights. Perhaps the first thing to notice is that this is a considerable reduction of the complexity of the relationship. Many things that one might consider important for the relationship, such as, for instance, John's infidelity, are not part of this representation. This kind of information is simply too complex to be represented in this format. At the same time, many of the aspects of the high-level summary given in the previous paragraph are indexed by the items we see. For instance, we see that it involves "marrying" (John → Gerda), "loving" (both directions), but also "killing" and "shooting" (Gerda → John). Furthermore, a series of verbs index the power imbalance in the relationship, with Gerda "adoring" John, who, in turn "teases," "interrupts," and "enjoys-without" Gerda, but also "looks-after" and "protects" her. The representations modeled here should not be misunderstood as a factual accounts of relationships for several important reasons: negation is not accounted for ("he did not love her" results in [he]-[love]-[her]), characters lie (especially in whodunnits like *The Hollow*), there is no differentiation among perspectives (a statement by a character is treated equally as on by an omniscient narrator), and there is no distinction between aspects presented as facts or possibilities ("she killed him" and "she could kill him" results in [she]-[kill]-[him]). Finally, the analysis offered in this paper focuses on action, but abstracts away narrative context or what we might call "scene," in reference to Burke (1945; also see Mohr et al. 2013). This is an important gap that I will return to in my discussion further below. Additionally, in Appendix E, I offer a qualitative analysis of the relationship between John and Gerda that is placed alongside the extracted actions.

Despite these limitations, however, what the verbs do provide is a relatively rich tapestry of what kinds of descriptors typically go together in narrativized social relationships. The

sociological insights we can gain from these representations don't concern the nuances of particular relationships. Instead they concern the larger association structure of the symbols designating social relationships — for example, do we generally "adore" those who "tease" us, when did "marrying" someone become related to "loving" them, and how common is it for romantic relationships to involve violence? These representations, while clearly a simplification, are nonetheless rich in that they offer a relatively free, inductive depiction of how social relationships are narrated.

When using a cutoff of at least five exchanged actions between two characters, there are a total of 1.2 million relationships, an average of 12.3 per book. These relationships usually contain two "action profiles" (Martin 2009, pp. 11-12, p. 21), one set of actions directed from A to B and one in the other direction. In a few relationships, all actions are contained in one profile, so that there are 2.3 million action profiles. The average relationship contains 13.8 actions, and the average profile comprises 7.1 actions.

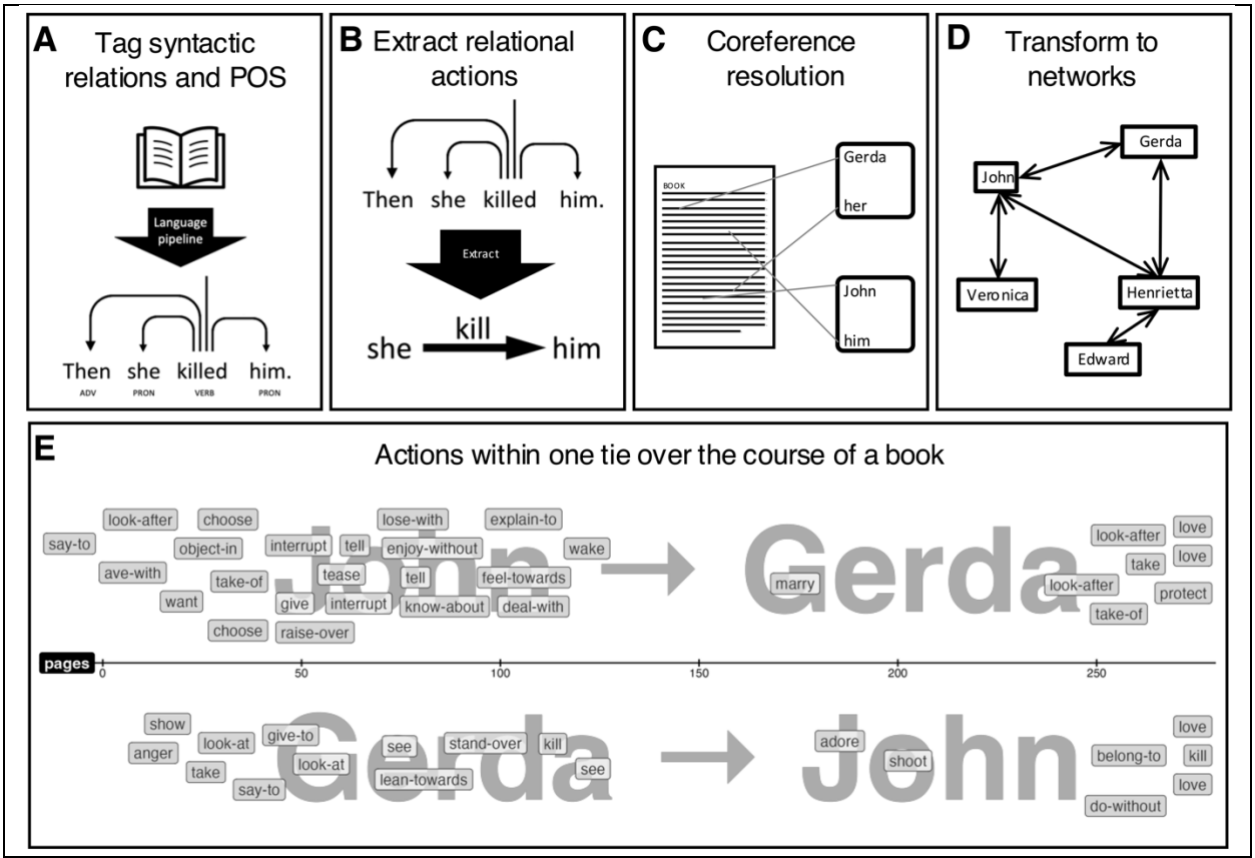

Figure 1. Workflow of relationship extraction
The figure illustrates the process used for extracting relationships from text used by Stuhler (2024). First all books were annotated with the spaCy language pipeline (A). Subsequently (B), actions between characters were extracted via the semgram R package (version 0.1.1). Coreference resolution (C) was done using the BookNLP python library. Finally (D), the actions exchanged between characters of the books were aggregated to relationships and networks. The panel shows the network between a

subset of the characters in Agatha Christie's "The Hollow." Panel E shows the relationship between the characters "John" and "Gerda" in said novel, which consists of two distinct action profiles. Actions are placed along the x-axis according to when they appear in the text — with slight adjustments via the ggrepel R package (version 0.9.4) to avoid overlap. Page numbers were estimated using the token position of the respective verb.

Analysis

Measuring relationship frames

Following the conceptual arguments above, the identification of relationship frames must involve finding out which relational items typically go together. There are a variety of clustering approaches one could use for this task. An option that aligns well with the stated theoretical assumptions are mixed membership models. In sociology, such models have been popularized in the context of text analysis, where they serve as "topic models." In Latent Dirichlet Allocation (LDA, see Blei, Ng, Jordan 2003), for instance, each document is represented as a “bag of words” and assumed to be generated by all topics. Each topic, on the other hand, is associated with a probability distribution over all possible words in the vocabulary. Topics are estimated probabilistically so that words that frequently co-occur in the corpus tend to be assigned high probability under the same topic. For instance, when modeling newspaper articles, one might get the topics *sport* and *economics*. The former is likely to have high probabilities for generating the tokens “football,” “referee,” and “team” while the second has high probabilities for the tokens “corporation,” “CEO,” and “stock.” Using the topic variables, one can then infer a membership vector for each document that specifies the relative importance of all topics for that document.

I suggest that we apply a similar logic to the modeling or action profiles. As I noted earlier, much like text analysis represents documents as “bags of words,” we can consider action profiles as “bags of actions.” Formally, this means we have a matrix in which each row is an action profile and each column corresponds to a particular action. The entries capture the number of times a specific action occurred in a given action profile. For instance, the row vector for the action profile Gerda → John contains has an entry of 2 for the “love” column but and entry of 0 for the “fight-with” column and so forth. Now, while LDA and related models are typically used to identify the topical composition of a set of documents, I suggest we can apply the same logic to identify the frame composition of action profiles. Whereas topic variables are probability distributions over the set of possible words, we can use the same approach to generate frame variables that correspond to a probability distribution over the set of all possible actions. While a topic variable will assign high probabilities to words that often come together within documents, a frame variable estimated under the same regime will assign high probabilities to actions that frequently co-occur in action profiles. Similarly, once we have estimated frame variables, we can use them to infer the composition of frames most likely to have generated the actions within a given action profile.

Unfortunately, the action profiles described above only contain a relatively small number of actions, and models like LDA typically do not perform well on highly sparse data structures. Therefore, I turn to the Biterm Topic Model (BTM, Yan et al. 2013), which was developed as a topic model for short texts such as headlines and tweets and is therefore likely to work better for modeling action profiles. Similar to LDA, the BTM is typically used to produce latent

variables that represent probability distributions across the feature vocabulary. However, unlike LDA, the BTM infers these variables by directly modeling the overall feature co-occurrence structure. Specifically, to overcome the extreme sparsity, rather than assuming partial membership of each document (or, in our case, action profile), it provides a generative process for "biterms," that is, the co-occurrence of two features (in our case actions) within a document (in our case action profile).[4] The latent variables can then be used to estimate the topical composition of every document, so that, like in LDA, a document can be depicted as a distribution over topics. In our case, this means we can use the frame variables to determine the frame composition of each action profile (see Appendix for formal definition of the model's generative process).

Note that instead of the action profiles, one could also model entire relationships as bags of actions, but this would effectively presuppose reciprocity or at least unnecessarily deprive us of the chance to study reciprocity (see Martin 2009, p. 21 on this point). Frames correspond to the latent variables that are estimated by the model. In this way, each frame is a probability distribution over the set of possible actions. This also allows me to represent every action profile, that is, every set of actions directed from one character to another as a composition of all frames, which cumulatively adds to 1. For this procedure, I limit myself to the 500 most frequent actions. I then fit a model on the 2.3 million unique action profiles, which stem from 1.2 million unique relationships. The unit of books does not play a role in this modeling step, nor do any textual features other than the extracted actions. I found that a solution with 25 frames provided a good tradeoff between model complexity and frame coherence, and produced highly interpretable results. Three variables were omitted as they were deemed uninterpretable. Further details on model selection are provided in the Appendix.

Relationship frame	**Actions**
Observe	glance-at, look-at, smile-at, stare-at, turn-to, grin-at, study, gaze-at, glare-at, eye, turn-toward, face, regard, watch
Know & think-of	see, remember, think-of, know, think-about, meet, imagine, forget, look-for, tell-about, recognize, find, hear-of, know-about
Converse-with I	say-to, turn-to, whisper-to, look-at, explain-to, call-to, ask, say, go-to, answer, smile-at
Converse-with II	ask, call, talk-to, remind, invite, look-at, hug, think-about, listen-to, smile-at, thank
Give-to	give, send, pay, show, teach, get, keep, buy, want, owe, make, offer, treat, put, do
Marriage & courtship	marry, like, care-for, love, refuse, accept, meet, hate, treat, dance-with, admire, think-of
Distant communication	write-to, write, hear-from, visit, ring, send, think-of, go-to, meet, receive-from, telephone, spend-with, persuade

[4] The BTM assumes each biterm to be generated by a single topic, which technically makes the BTM a mixture model, rather than a mixed membership model. Additional details on the generative process assumed by the model are provided in the Appendix.

Love & romance	love, want, lose, kiss, miss, forget, think-of, need, hurt, adore, marry, care-for
Beat & kick	hit, grab, push, slap, beat, drag, pull, kick, shake, shove, glare-at, throw
Tell & assure I	remind, assure, inform, warn, point-to, urge, ask, explain-to, advise, order, persuade
Tell & assure II	tell, call, talk-to, explain-to, show, lie-to, assure, trust, warn, believe, hear
Hold authority over	send, command, teach, order, spare, summon, call, bid, call-to, instruct, bless
Lead & guide	lead, show, hand, teach, introduce, pull, hand-to, guide, help, put-around, hug, wink-at
Physical affection	kiss, hold, pull, put-around, touch, draw, lift, hug, release, want, hurt, carry, grab
Physical care	carry, take, lift, pick, put, lay, feed, hold, set, bring, place, leave
Shoot & kill	kill, shoot, murder, hate, hit, do-to, beat, strike
Dislike & forgive	hate, forgive, believe, despise, understand, trust, listen-to, hurt
Search	find, look-for, spot, ride-with, go-for, get, come-for, save
Long-for	call, miss, need, think-about, want, talk-to, sleep-with, phone
Motion-with & motion-to	follow, watch, join, glance-at, call-to, catch, hear, stand-beside, catch-with
Thank, beg, & promise	thank, beg, assure, pray, cry, beg-of, entreat, wish, promise, serve
Be-with	come-to, leave, be-with, go-to, stay-with, come-with, go-with, marry, be-to

Table 1. The frame system of fiction: relationship frames and associated actions.
Note: For each relationship frame, this table shows the most strongly associated actions. These were selected according to a criterion that balances frequency of an action under a frame and exclusivity (FREX). Specifically, I compute a score that is the mean of an actions rank by probability and rank by exclusivity. The table contains all actions with scores smaller than 30 for each frame, resulting in 11.1 actions per frame. Actions are ordered such that the respectively first actions have the highest FREX score. The Appendix contains an expanded version of this table with more terms.

In Table 1, I present the identified frames together with their most indicative actions. These frames can be seen as dominant structures of dyadic action in literary relationships between 1850 and 2010. Actions were selected and ordered according to a score (FREX) that balances two aspects: the rank based on the probability of an action according to a specific frame variable and, on the other hand, the rank based on exclusivity of that action to that specific frame (i.e., if we observe the action, what is the likelihood of it stemming from this frame). This means that the respectively first words are the most central to the frame. FREX scoring is a standard approach to generating outputs of mixed membership or related models. To select the actions for Table 1, I used a threshold of 30 for the FREX score. Nonetheless, because BTM is a soft clustering approach, every action is assigned a probability (albeit often minimal) under each frame.

The labels of the frames reflect the author's interpretations and aim to give a high-level summary of the associated actions. For instance, the frame labeled *Love & romance* is most strongly associated with the actions “love,” “want,” “lose,” “kiss,” “miss,” and so forth. Substantively, this means that these actions often occur together in action profiles, that is, in the set of actions that one character directs at a particular other character. Meanwhile, the frame *Shoot & kill* is strongly associated with “kill,” “shoot,” “murder,” “hate,” “hit,” and “do-to,” among others. As previously discussed, actions can be associated with multiple frames. For instance, one frame is about the *Longing for* an alter. Besides the actions “call,” “need,” “think-about,” and “want,” this also involves “miss,” which is also associated with the previously mentioned *Love & romance* frame. Allowing for this is also important to deal with the polysemy of some highly frequent verbs. For instance “send” is associated with the frame *Holding authority over* (which also involves the actions “command,” “teach,” “order,” “spare,” “summon”), but also with *Distant communication* (“write-to,” “write,” “hear-from,” “ring,” “telephone,” etc.), and the frame *Giving to* (“give,” “pay,” “get,” “buy,” “owe,” etc.). Other notable frames include *Physical care* (“carry,” “take,” “lift,” “pick,” “put”), *Marriage & courtship* (“marry,” “like,” “care-for,” “love,” “refuse,” “accept,” “dance-with,” “admire,” etc.), *Lead & guide* (“lead,” “show,” “hand,” “teach,” “introduce,” etc.), *Holding authority over* (“send,” “command,” “teach,” “order,” “spare,” “summon”), *Thank, beg, & promise* (“thank,” “beg,” “assure,” “beg-of,” “entreat,” “promise,” etc.), beat & kick (“hit,” “grab,” “push,” “slap,” “beat,” “kick,” etc.), and physical affection (“kiss,” “hold,” “put-around,” “touch,” “hug,” “want,” etc.).

It is worth noting that a focus on the top 500 verbs implies a deliberately coarse description. For reference, actions at this mark are “worry,” “see-with,” and “coax,” with around 3,500 occurrences. The frequency distribution of actions is highly skewed, and the top 500 actions account for 84% of all actions. Meanwhile, for instance, actions ranked around 2000 are “happen-to,” “plead-for,” and “smirk-at,” which all occur 415 times. Including more actions would likely make it necessary to increase the number of frames necessary to account for heterogeneity.

The system of frames identified through the BTM approach provides a broad, high-level structure. It summarizes the co-occurrence structure of action within action profiles. Essentially, it answers the question: If we had to break down literary relationships into 20 frames, what would these frames look like? This bears conceptual resemblance to Propp’s notion of “spheres of action.” In a classic study ([1928] 1968), he argued that folktales are typically structured around archetypical roles engaged in certain kinds of actions, such as the hero, the helper, or the villain. The frame system here extends this logic: rather than assuming homogeneity in the action of a character (i.e., the duality of character and action; a specific character does similar things), it rests on relative homogeneity in the actions a character directs at a specific other character (i.e., the duality of action profile and dyad; a specific character does similar things to a specific other character). Propp also discussed the possibility that single characters (or “dramatis personae”) are engaged in several spheres of action (see p. 80). The mixture logic used for modeling action profiles here can

be seen as a formal implementation of this general idea in the sense that it account for the fact that any specific relationship will draw on not just one but multiple frames.

Note that aiming to identify the 20 main frames that account for variation in the content of action profile this is considerably different then asking what frames exist. People's cognition likely includes more nuanced relationship frames, such as "abusive love relationship" or "frenemieship." Like the frames identified here, these more intricate frames also emerge from regularities in the duality of dyad and content. However, as I will discuss in more detail below, capturing them would likely require a more sophisticated measurement approach that extends beyond verb-based features and accounts for functional interactions, and ordering of actions.

Panel B of Figure 2 shows the relative prevalence of the different frames. The most dominant frames concern *Conversing-with* with 10% ("say-to," "whisper-to," "explain-to," "ask," "talk-to," etc.), and *Telling & assuring* with 9% ("remind," "assure," "inform," "warn," "tell," "call," etc.). For each of these, the model identified two independent frames that had a lot of overlap in the actions associated with them. I document these distinctly in Table 1 but combine them for further analysis. Counting them together suggests that frames denoting communication make up for around a fifth of the actions in relationships. In other words, communication is the primary form of interaction among characters in fiction writing. Most other frames account for 5 to 3% of the actions. At the tail of the distribution, we find frames that are rare but highly distinct in terms of their action content, like Shoot & kill or Hold authority over, respectively, at around 2%. As I show in the Appendix, there are some shifts in the prevalence of these frames over time. The most drastic increase can be seen in frames relating to conversation and observation, resonating some of the findings by Piper (2018). Meanwhile, there is a steep decline in the frames *Thank, beg, & promise* as well as *Distant communication*.

As noted in the theory section, actions cue frames, with some actions being unambiguous cues but others pointing to various possible frames. The modeling approach reflects this: for instance, previously discussed "send" has a 44% probability of stemming from *Holding authority over*, 30% probability of stemming from *Giving-to*, and 14% of stemming from *Distant communication*, all other frames being below 3%, respectively. Meanwhile, "murder" cues Shoot & kill at 95%, with all other frames below 1%. Panel A of Figure 2 illustrates how this feature can be used to predict the frames for specific sets of actions. For instance, in an action profile comprising "dance-with," "love," "kiss," "marry," and "sleep-with," the primary frames are *Marriage & Courtship* (32%) and *Love & romance* (17%), and *Physical affection* (14%). I also show the frame structure of the two action profiles in the relationship between Gerda and John. I find that Gerda's actions toward John load most strongly on the frames *Shoot & kill* (19%) and *Love & Romance* (11%). This aligns well with the core themes of her relationship towards her husband as outlined priorly. Meanwhile, John's action profile uses the frames *Tell & assure* (9%), and *Converse-with* (9%). Note that frames relating to communication typically make up for a large share of the action in all relationships. When comparing the composition of John's actions towards Gerda to the

average frame prevalence, the most distinctive features are *Physical care* (+4% versus global average), *Love & romance* (+4%), and *Marriage & courtship* (+4%). In the next sections, I will systematically examine the frame distributions in relationships.

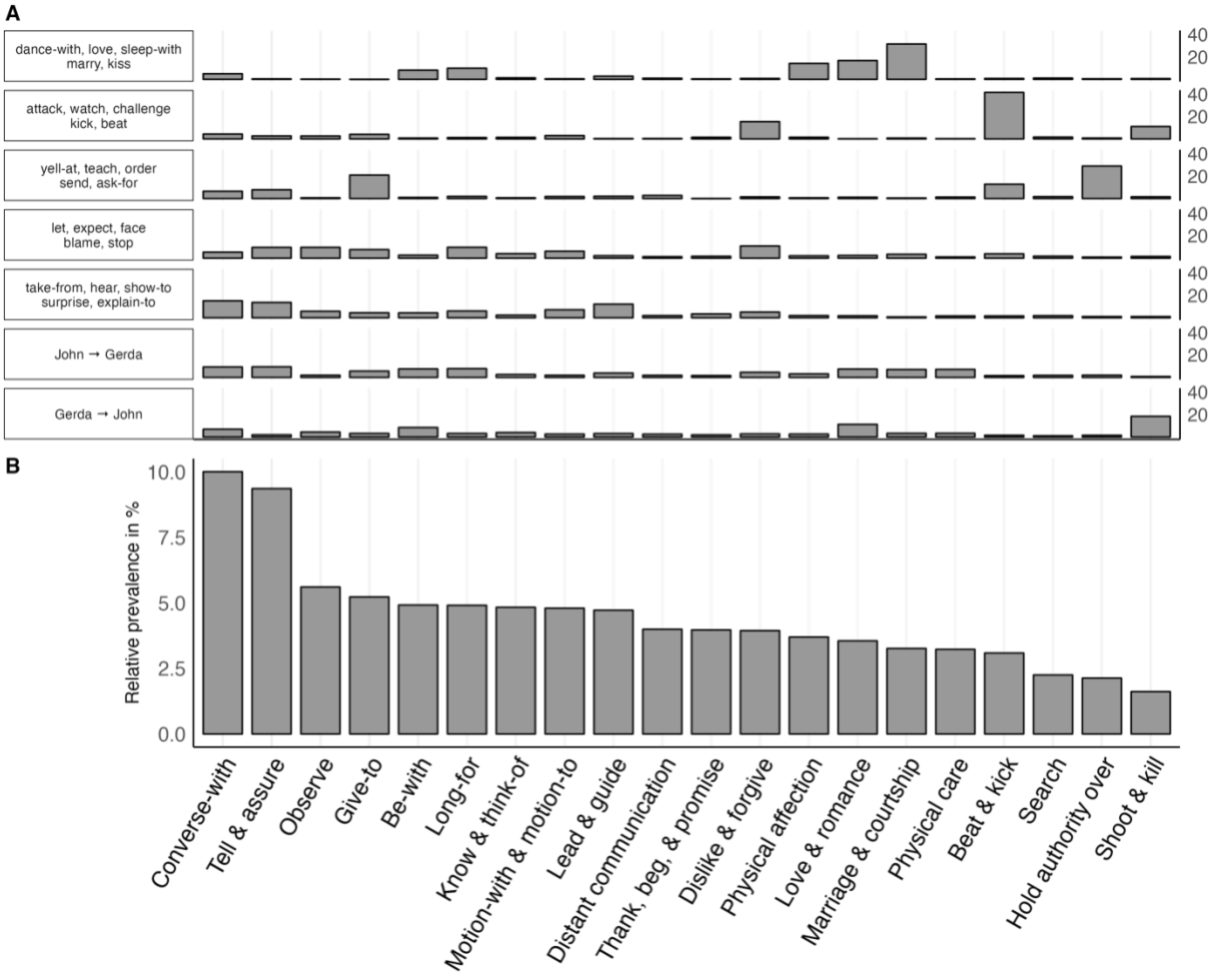

Figure 2. Prevalence of relationship frames.
Panel A shows the frame predictions for exemplary action profiles. The bottom two rows show the predicted frame composition of the two action sets between the characters "Gerda" and "John" in Agatha Christie's *The Hollow* (see Figure 1 for the action sets). Panel B shows the prevalence of relationship frames averaged over 2.3 million action profiles.

The frame structure of action profiles

While there has been work that inductively investigates the copresence of tie types, these studies have usually been confined to particular contexts like school classes (Vörös & Snijders 2017) or urban communes (Yeung 2005; an exception is Gondal 2022, who uses the GSS). Moreover, they were focused on a limited number of ties, which had to be defined a

priori by researchers. Here, I present a comprehensive, high-level analysis of how action is structured within action profiles in fiction writing.

Figure 3 shows the associations between all action frames within action profiles, that is, within the set of actions directed by one character at another. Cells are colored according to the correlation, with associations above the 90th percentile and below the 10th highlighted via annotations. To ease interpretation and description, the rows and columns were grouped using a hierarchical clustering algorithm (see figure note for details). I find that the largest densely connected block of frames (1) concerns intimacy and strong emotions, including the frames *Love & romance*, *Marriage & courtship*, *Being-with*, and *Dislike & Forgive*. All four frames correlate with each other, indicating that they frequently go together in action profiles. Another densely connected set of frames concerns communication (6), and includes *Converse-with*, *Tell & assure*, and *Lead & Guide*. Cluster 6 includes frames that imply a spatial separation of ego and alter: *Knowing- and thinking-of* often goes with *Distant communication*, but also with *Searching* and with *Longing for*. Interestingly, *Holding authority over* is not just associated with *Giving-to*, but also with *Thank, beg, & promise* (7). If we take these two frames as indicative of power distributions, this may suggest that for literary relationships, the primary question is *whether* a relationship involves power — an issue I will return to in the next section. Beyond these four frames, there are three sets that include only two strongly associated frames: *Physical care* and *Physical affection* are the most highly correlated frames (4) at .35. Similarly, a character *Observing* another character is strongly associated with *Motioning-to or -with* that character (3). And the frame *Shoot & kill* is most closely associated with *Beat & kick* (5).

Just as important as identifying which relationship frames are compatible is understanding which relationship frames are highly unlikely to be found in the same action profile. Most notably, there is a clear disassociation between intimacy and strong emotions (1) and communication (6), where all frames in one cluster correlate negatively with the frames in the respectively other cluster. For instance, *Being-with* an alter rarely coincides with *Conversing-with* or *Telling & assuring* said alter. Other strong disassociations are less aligned with the clusters: *Distant communication*, somewhat unsurprisingly, is unlikely to coincide with action frames that connote direct interaction (*Physical affection, Observe, Beat & Kick*). Similarly, the cognitive frame *Knowing & thinking-of* is unlikely to coincide.

In addition to these analyses, the Appendix reports the results of a multidimensional scaling of the frames' association within action profiles. This analysis suggests that the most variation may lie in the distinction between frames that relate to emotions and cognitive processes, on the one hand, and frames that connote forms of physical action or co-presence. However, its more general takeaway is that further reducing the variation in frames to a lower-dimensional space is of limited use. A two-dimensional reduction has a relatively poor fit. This, of course, is not too surprising, as it reflects the BTM algorithm's objective to identify distinct clusters, but also the more general fact that action is simply more complex.

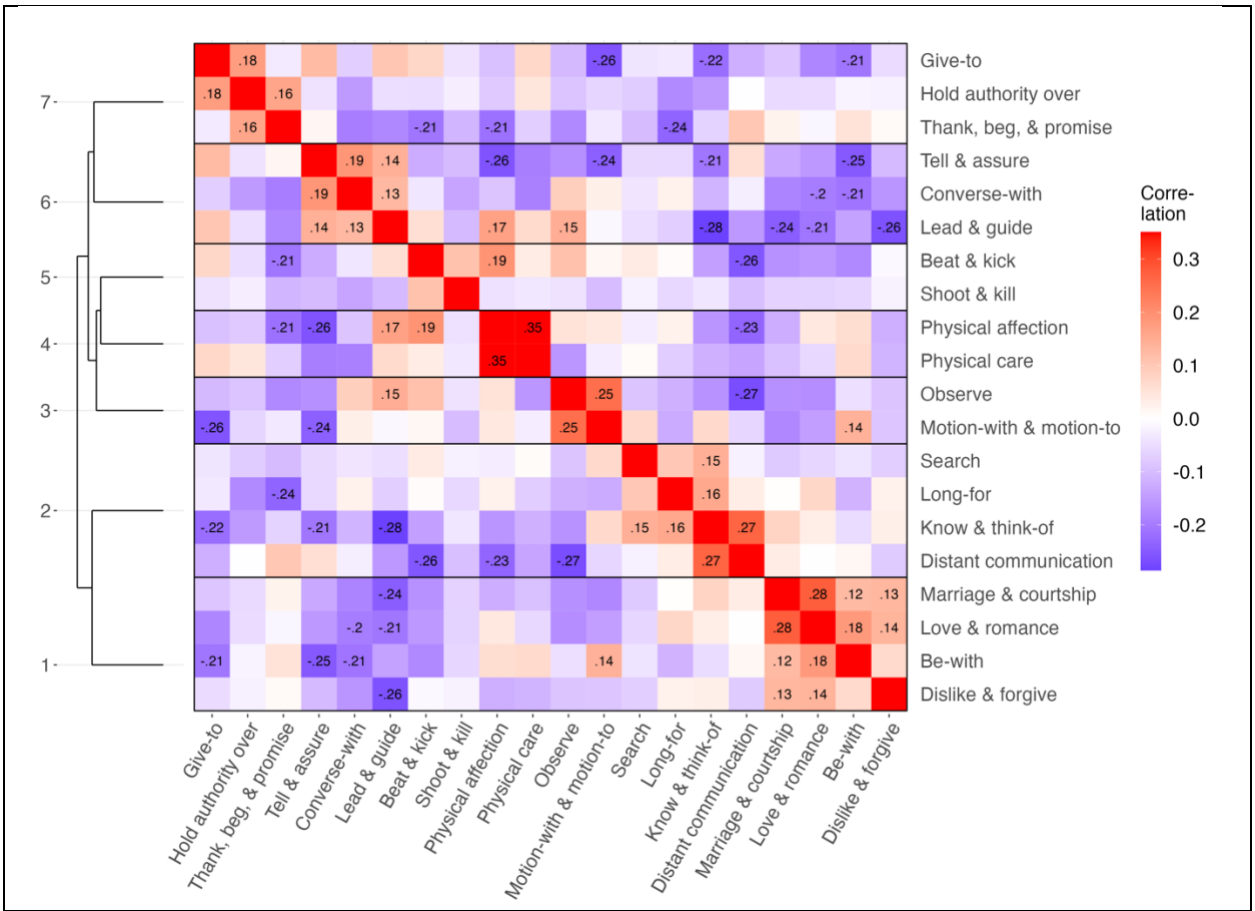

Figure 3. Association of relationship frames within action profiles.
Note: The figure shows the correlation of relationship frames within actions sets. Only action profiles with five or more actions were considered, resulting in 1.2 million action profiles. Correlations above the 90th percentile and below the 10th were highlighted with annotations. The diagonal is 1 and was colored to match that of the highest observed correlation. The left panel shows groupings generated by hierarchical clustering using the Ward algorithm. The distance between relationship frames was respectively measured as $d_{ij} = \sqrt{2(1 - c_{ij})}$, where c_{ij} is the correlation between relationship frames i and j. The tree was cut at seven branches to ease interpretation. The full tree is reported in the Appendix.

Reciprocity and alter's action profile

Whereas the previous analyses looked at associations within action profiles, I now turn to examining how the frames in ego's action profile relate to those in alter's action profile. In other words, what kinds of actions by ego correspond to what kinds of actions by alter? In Figure 4, I create a correlation matrix like before, applying the same clustering approach.

Most relationship frames involve *reciprocity*, that is, actions of a particular kind by ego correspond to similar actions by alter, as indicated by the positive correlations in the diagonal.[5] The strongest reciprocity norms concern *Love & romance* (*r* = .33) as well as *Marriage & courtship* (*r* = .25), and both are also strongly connected with one another. Notable exceptions to the reciprocity norm are *Know & think-of* and *Physical care*. In fact, together with *Search*, these frames are distinct from all others in that they are not associated with any specific kinds of alter behavior.

The strongest cross-frame associations exist between *Hold authority over* and *Thank, beg, & promise* (r = .17). This is not surprising, as from their content, both frames appear like two sides of the same coin. Yet, interestingly, both frames are also reciprocated. This confirms a suspicion briefly discussed above: power in literary relationships is not a zero-sum game where one character has it, and the other doesn't. Rather, there are relationships that involve frames related to power and ones that don't. In the former, it is likely that both characters have some authority over the other, albeit possibly at different times throughout the narrative. This observation also points to a limitation of the present analysis: its diachronic nature.

[5] While it might seem plausible that this could have something to do with errors in attributing actions within a dyad (i.e., conflating subject and object), errors of this type are actually very rare and cannot account for the correlations observed in Figure 3 (see Stuhler 2024, Appendices A and M).

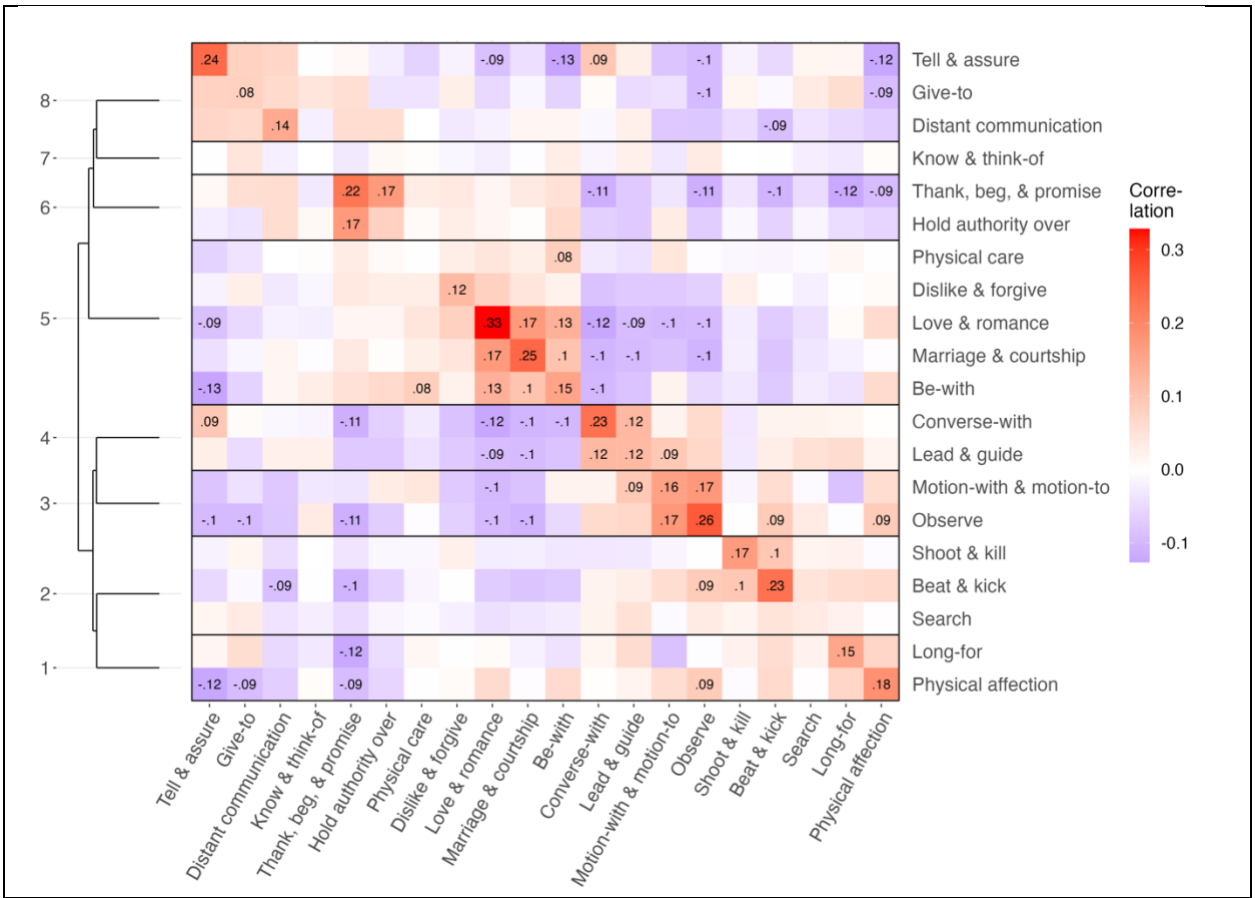

Figure 4. Association of relationship frames with alter's action profile.
Note: The figure shows the correlation of relationship frames of ego's action profile with the relationship frames in alter's action profile. For instance, the relationship frame *Hold authority over* in ego's action profile correlates with the relationship frame *Thank, beg, & promise* in alter's action profile at .17. Only dyads where both action profiles had five or more actions were considered, leaving 360,861 dyads. The procedure and representation choices match those described for Figure 3. The full tree is reported in the Appendix.

Measuring frame ambiguity

Above, I proposed that we think of ambiguity in relational meaning as *frame ambiguity* — the extent to which the actions and symbols designating a relationship index a variety of frames rather than just one or a few. This idea now becomes formally tractable through the notion of relationship frames and their probabilistic relationship to actions. As discussed above, each action is associated with a generative probability under each frame. To measure the frame ambiguity of an action profile, we can measure whether the actions contained in it unambiguously cue specific frames, or whether they point to various possible frames.

To operationalize, I draw on the concept of entropy. Entropy is a measure from information theory that quantifies uncertainty or unpredictability. Simply put, it describes how evenly the

probabilities are spread across possible outcomes. For instance, in the case of a coin flip, entropy (uncertainty) is maximized when the coin is fair (i.e., $p(head) = p(tail) = .5$). A biased coin (e.g., $p(head) = .9; p(tail) = .1$) has a more predictable outcome and hence less entropy. I use this concept to measure the extent to which an action profile draws on many frames, rather than just a few. Formally, the ambiguity A of an action profile is measured as $A = -\sum_{i=1}^{20} p_i \log_2(p_i)$ where p_i is the probability of the i-th frame.

To illustrate this, compare the top three action profiles in Panel A of Figure 1 with the two at the bottom. The upper three action profiles all point to a relatively small set of frames and have entropy of 3.1, 3.1, and 3.4, respectively. Meanwhile, the fourth and the fifth action profiles appear fairly ambiguous and are not uniquely attributable to any specific frames, leading to an entropy value of 4.3 for both. We can also apply the concept to the two actions profiles in the relationship between Gerda and John. I find that the action profile John → Gerda (4.4) is slightly more ambiguous than that of Gerda → John (4.2). This puts them in the 96th and 62th percentile, respectively. The distribution of the measure of frame ambiguity is documented in the appendix.

Predictors of frame ambiguity

What predicts the extent to which an action profile draws on many different frames? In this section, I present an exploratory investigation of some of the conditions under which actions from ego towards alter — measured synchronically across a book — index a heterogeneous set of frames. I use the following set of predictors:

- *The prevalence of frames in alter's action profile*. This tests whether *frame ambiguity* of ego's actions towards alter are a function of the kinds of actions that alter directs towards ego. Put differently, we might say that we assess whether some kinds of action by alter "constrain" the kinds of action ego engages in. Of course, "constrain" cannot be understood here in any causal sense, for we are dealing with fictional, synchronic relationships.
- *The gender composition of the dyad*. This tests whether frame ambiguity differs by whether it is directed from a male or a female character at a male or a female character, respectively. Character gender was predicted using the BookNLP library (Bamman 2023), which uses information information on first names, pronominal coreference, and honorifics. Accuracy for this particular corpus was evaluated manually and is estimated to be at .96 and is robust across time (see Appendix M of Stuhler 2024 for additional details).
- *The "strength" of the relationship*. This is measured as the number of actions in ego's and alter's action set, respectively and tests whether relationships that are described more extensively in the book tend to have more or less frame ambiguity.
- The importance of both ego and alter for the book. Character important is measured as the betweenness centrality. This measure is based on the global network structure

of each book, taking into account relationships with at least five actions. It implements the idea that a character is more relevant if it is placed at the center of a book's social structure. While other measures of character importance such as the degree centrality or the effective network size are equally plausible, these are highly correlated in the corpus (see Stuhler 2024 Appendix N for details).

To test which of these variables predict frame ambiguity within action profiles, I fit a linear mixed model on the 712,056 action profiles. Figure 5, present the results of the model. The dependent variable was z-transformed and the model uses random intercepts for dyads, books, and authors, as well as fixed effects for each decade. Bars denote 95% confidence intervals based on cluster-robust standard errors on the author level (the full model is documented in the Appendix; also see figure description).

In the upper part, I explore the effects of alter's actions toward ego on ambiguity in ego's action profile. All frames were scaled such that the coefficients represent the effect of a 10-percentage point increase in the prevalence of the relationship frame in alter's action profile. I find that some kinds of actions by alter increase ego's frame ambiguity. Notably, when ego is *Observed* by alter, the frame ambiguity of is action profile increases by .03 standard deviations. Other actions by alter that correspond to more heterogeneous actions by ego are *Distant communication*, *Leading & Guiding*, as well as *Thanking, begging, & promising ego*. Put differently, alter drawing on these frames tends to increase the set of probable actions by ego.

On the opposite end, I find that the starkest reduction in frame ambiguity is triggered by alter *Holding authority over* ego, as well as by alter *Beating & kicking* ego. In other words, these are frames that typically correspond to ego engaging in very specific kinds of behavior; they constrain the realm of probable actions by ego. As we saw in the previous section (cf. Figure 4), *Authority* is usually responded to by *Thanking, begging, & promising*, whereas *Beating & kicking* is typically likely to be reciprocated, responded to by *Shooting & Killing*, and making especially intimacy-related frames less likely. Other frames associated with a reduction in the frame ambiguity of ego's actions are *Motioning with & -to*, *Physical care*, *Giving-to*, and *Love & romance*.

Ambiguity is also affected by network position and action volume. I find that ego's betweenness centrality has the strongest positive effect on frame ambiguity. This supports the idea main characters generally draw on a wider variety of frames in their actions. Perhaps more surprising, I find that that the difference between the effects of ego's and alter's betweenness centrality is relatively small. This suggests that important characters don't merely act in more heterogeneous ways, but they are also treated in that way by other characters. In other words, these patterns suggest that action profiles *by* and directed *at* main characters tend to have more frame ambiguity. Importantly, these effects are net of the volume of action within the relationship. So frame ambiguity is heightened in relationships with important characters — even when these relationships are only comprised of a relatively small number of actions.

Finally, I also tested whether the gender composition of the dyads has an effect on frame ambiguity. The reference category for the coefficients presented in Figure 5 is female-female dyads. The model suggests a curious result: interaction among male characters has the lowest frame ambiguity, suggesting that male interaction draws on fewer frames than is the case for all other constellations. Prior work has shown that fiction prioritizes male characters in a various ways, for instance by devoting more space on the page to them and presenting them as more agentic vis-à-vis female characters (Underwood, Bamman, and Lee 2018; Jockers and Kirilloff 2016; Stuhler 2024). An important nuance here is that this finding is net of character importance, with male characters typically having higher degree (Stuhler 2024). The results presented here suggest that net of character importance, male-male interaction is typically focuses on a smaller number of frames. Meanwhile, male characters' actions towards female characters have the most frame ambiguity, suggesting that male characters' actions are more varied when directing them at female characters. Finally, female characters' actions index more frames when acting toward other female characters than when acting toward male characters. Investigating the scope conditions of this finding is beyond the purview of the present study. However, it complicates the prevalent idea that women are portrayed as more unidimensional characters.

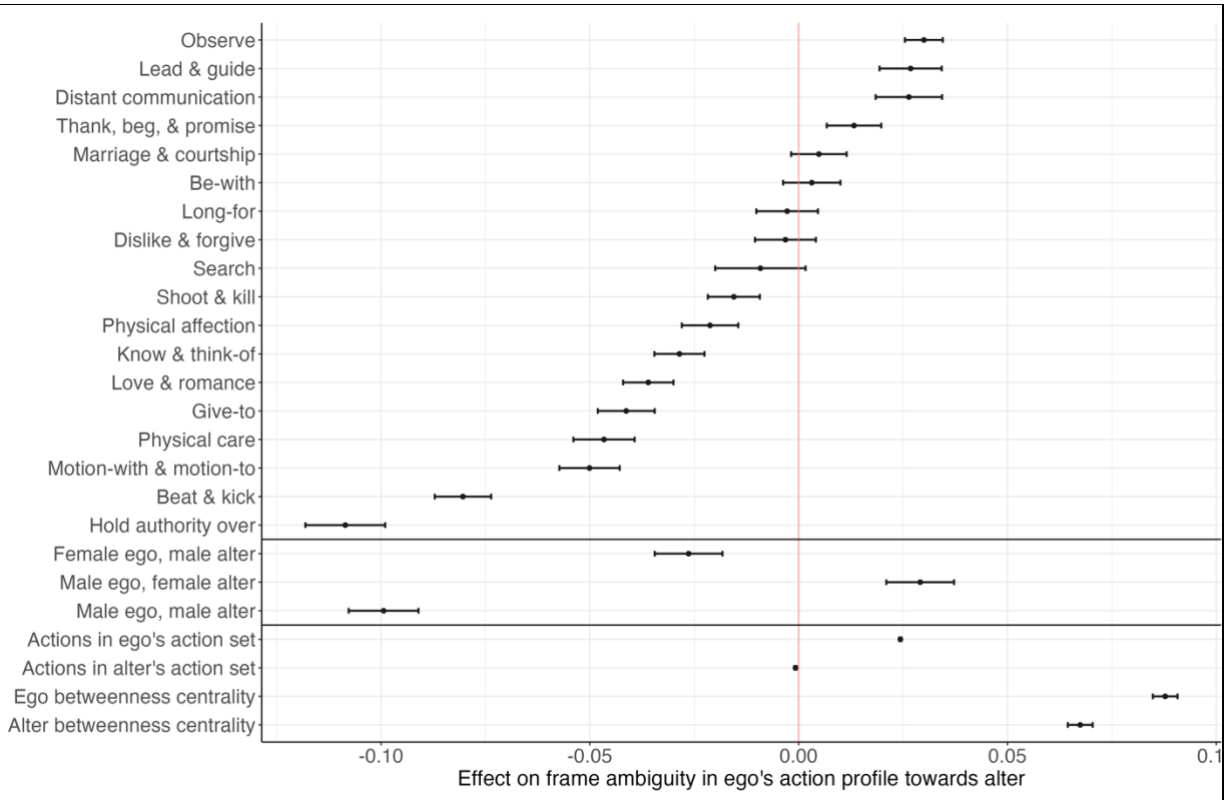

Figure 5. Predictors of frame ambiguity.
Note: The figure shows the effect different variables on frame ambiguity of ago's action profile. The coefficients are based on a linear mixed model fit with the 712,056 action profiles. Only dyads with at least five actions in both action profiles were considered. The

dependent variable was z-transformed. All relationship frames were scaled such that the coefficients represent the effect of a 10-percentage point increase in the prevalence of the relationship frame in alters action profile. Betweenness centrality was standardized within each book. The reference category for the gender constellations are female-female dyads. The model uses random intercepts for dyads, books, and authors, as well as fixed effects for each decade. Bars denote 95% confidence intervals based on cluster-robust standard errors on the author level. The full model is documented in the Appendix.

Discussion: limitations and opportunities

There are several limitations to the analysis put forward here, which point to opportunities for future work. In what follows, I will discuss these systematically, addressing issues concerning feature selection and modeling, context, and temporality and institutionalization.

Feature selection and modeling

A reasonable challenge to the above is that while I argue in favor of dissolving tie categories, verbs or what I call actions are themselves categorizations. If a novel states that character A "loves" character "B," isn't this precisely the kind of typification I'm criticizing? This is true to an extent, but "love" here serves less as a typification but rather as one of many possible symbols (500, to be precise) tagging a tie. Therefore, it is worth having our anticategorical ambitions push us down to the level of verbs.

This points to another critical question: what features should go into what I have subsumed here under the term "content"? For ties described in literary works, verbs are a plausible option, for they are the symbols that most directly designate the relationship between two characters. However, information about the relationship between two characters will often be given in more subtle ways. Verbs offer an effective but coarse choice for featurizing the content of ties. Relationships between people are a more complex matter, and the question of what useful lower-level features are will strongly depend on the context and available data.

This choice will have major implications for the kinds of relationship frames we are able to uncover. A serious limitation of the present approach is that the 20 frames I was able to uncover here are rather rudimentary. As one reviewer aptly pointed out, people likely draw on more nuanced relationship frames when making sense of what goes on in social relationships. For instance, most people will have an implicit understanding of the frame "abusive love relationship," which will entail both aspects of romance and violence. Similarly, people may be aware of the expectations attaches to the constructs "frenemieship," "sugar arrangements," or "friends with benefits." These more nuanced

frames don't seem to challenge the key argument that any such frame is ultimately rooted in regularities in the duality of dyad and content — meaning, that there are many relationships with a specific set of features. However, they would require a more sophisticated approach to detecting these regularities and possibly more data.

There are different reasons for this. For one thing, at the level of measurement, the proposed approach can get at verbs and thereby capture some of the constitutive practices of relationship frames. However, other aspects may be too complex to be measured in this way. For instance, signaling to others that one is unavailable may be considered an important part of the relationship frame "monogamous love relationship." Yet this is a complex action that cannot be captured on the level of verbs. Second, the relationship between frames and their constitutive elements may take on functional forms that cannot be well captured with a mixture modeling approach such as the BTM. This involves interactions (e.g., the copresence of "transferring money" determines whether "providing accommodation" indexes of a "host-guest" frame or a "landlord-tenant" frame) but also non-linearities (e.g., the frame "good neighborship" may require *some* "chatting" but above a certain level, things might cross over into "friendship"). Third, there may be criteria relating to the order of steps that are not accounted for in the present modeling approach. For instance, in the context of the artworld, "creates artwork for" and "provides material resources to" might co-occur in relationships, but the order in which they do determines whether the frame "patronage" and "buyer-contractor" is applicable.

Overall, then, the frames identified via the BTM approach — based on the 500 most common verbs — should be viewed as a coarse superstructure rather than a complete taxonomy. They by no means capture the full range of interpretive frames people might use to make sense of social relationships, which likely includes more nuanced frames with more complex features. Inductively uncovering these frames would require a more refined analytic strategy.

Context

Another limitation of the present study is that it focuses strictly on the actions exchanged between characters in fiction writing. This leaves out several aspects concerning the context.

First, it is worth noting that my analyses are strictly dyadic and do not take into account *network context*. This relates both to the identification of relationship frames as well as to the statistical analyses on frame ambiguity. While relationship frames are conceptually rooted at the level of the single tie, there are culturally established network formations. One example of this would the "love triangle," where one person maintains romantic relationships to two (competing) partners. Similarly, in a "third wheel" friendship triad, two people share a close friendship and a third person is attached only weakly. A "wingman" helps one person romantically pursue a third person. Just like relationship frames, the

concepts are cultural forms that people recognize and enact or construct in writing. Identifying these inductively would likely require a more sophisticated methodological approach than the BTM but may present a promising direction for future research.

When it comes to the analyses of frame ambiguity, a natural next step would be to investigate how such ambiguity is situated within larger network structures. Here, I approaches this only in a rudimentary way by looking at whether central characters maintain more ambiguous ties. However, the analyses do not account for common alters or incorporate broader network structure. On the one hand, this implies that we cannot exclude the possibility of biases in some of the estimates presented in this paper; on the other hand, this also points to several open research questions for the growing literature on literary networks (see, e.g., Moretti, 2011; Piper et al., 2017; Kraicer and Piper 2019; Sims and Bamman 2020). How to authors embed ambiguous ties into literary networks? For instance, does frame ambiguity cluster in networks or distribute evenly? For the case of fiction writing, the former would suggest that some areas of the social structure in a book are treated as stable and aligned with cultural norms while others are not. Meanwhile, the latter would imply that ambiguous ties are embedded within culturally stable network context.

Second, the analysis presented above focus on action, but largely abstracts away *narrative context*. A well-known approach to systematically thinking through what we mean by narrative context is Burke's Dramatic Pentad. In *A Grammar of Motives* (1945), Burke famously famously distinguished five elements for analyzing human actions: act (what?), scene (when and where?), agent (by whom?), agency (how?), and purpose (why?). Getting at the other elements of this pentad is difficult for a purely formal approach like the one pursued here. In Appendix E, I offer a qualitative analysis of the relationship between John and Gerda that I used as a running example throughout. While verbs and verb phrases that directly connect characters with one another are important descriptors of a relationship, this analysis also illustrates that the meaning of these elements is to a good extend context-dependent. That said, some of the other elements of the pentad do lends themselves to formal measurement. Mohr and colleagues (2013) operationalized "scene" in the NSSR with via topic modeling. Meanwhile, other scholars have started to formally investigate how attributes of characters, that is, "agents" are related to their actions and the characteristics attributed to them. Thus far, these analyses have largely been limited to character gender (see Jockers, M., & Kirilloff 2016; Underwood, Bamman, and Lee 2018; Cheng 2020; Stuhler 2024), which is likely to be at least partially a reflection of the fact that character gender is easily inferable via names and pronouns. However, generative language models have recently shown enormous potential for extracting structured information from unstructured text (Stuhler, Ton, Ollion 2025; Lee et al. 2025). This could open up the possibility of formalizing other character attributes, roles, or motives. This could then allow for a more comprehensive analysis of the how character relationships are embedded in narrative context.

Finally, third, another opportunity for further inquiry lies in expanding the analysis to the *social context* of the works studied here. In my analysis, I collapsed 160 years of fiction writing into one analysis, in order to identify dominant frames. However, books and hence the relationship portrayed in them constitute "local meaning structures" (Basov, de Nooy, and Nenko 2021), that is, they are ensembles of semantic associations embedded in very specific social contexts. Many writers are embedded in specific literary circles. Not only do such idiocultures (Fine 1979) influence the kinds of cultural works artists produce, but so too are their positions within these communities (Basov, Lee, and Antoniuk 2016; also see Basov 2020). Data on the networks among authors could facilitate a socio-semantic analysis of how relationship frames are embedded within and shaped by social structures (Basov & Brennecke 2017; Basov & Roth 2025). This is especially important in our case, for we know that cultural models of relationship vary across contexts and fields (Yeung 2005; Fuhse and Gondal 2024; Burt and Schøtt 1985, pp. 296-298). Zooming out further, literary works belong to genres, respond to market or field demands, reflect the culture of a specific time or a specific sociodemographic group. Hence, cognitive conceptions of relationship frames are likely to vary across these contexts. Do authors from different class, racial, or national backgrounds portray relationships differently? How do authors' positions in the literary field (e.g., dominant vs. non-dominant) shape how relationships are portrayed? These questions point to a comparative research agenda that tracks how relationship frames vary across contexts and groups that aligns well with recent advances in network ecology (Doehne, McFarl, and Moody 2024).

Temporality and institutionalization

Another exciting possibility lies in expanding the framework offered here with a dynamic perspective. This applies in at least two important ways.

First, the analyses above are also synchronic in the sense of *historical time*. However, a crucial question concerns whether and how models for social relationships evolve over time. As I have indicated above, having a way of discovering regularity in relationship content independent of labels will be key to studying processes of institutionalization, which may involve the emergence of but also the weakening of existing relational categories (Breiger and Wagner-Pacifici, 2023; Fuhse 2020, esp. pp. 146 ff). This is chiefly because it allows us to uncover heterogeneity that lies below or across existing labels. In other words, uncovering regularity in dyadic content can allow us to uncover processes by which relationship categories emerge. Network analytic methods have proven useful in studying the institutionalization of categories (see e.g., Mohr 1994; Mützel 2022; Jones et al. 2012). However, thus far, this has rarely been applied to relationship categories themselves. While I did not include relationship categories in my analysis, it is especially here that narrative accounts of the type investigated in this will be highly relevant (White 2008), for "linguistic objectification" (Berger and Luckmann 1991[1966], esp. pp. 53-56, 89-92; Schütz, 1932 p. 40) plays a crucial role in institutionalization processes.

This can be illustrated with the example of "situationships," a relatively new label for what is described in the media as a supposedly new form of relationship prevalent among Zoomers and late Millenials. The attributes of this supposed new relationship form are varied and, at times, involve promiscuity, low levels of commitment, romantic affection, sexual activity, and significant amounts of time spent together (see, e.g., Chicago Daily Herald 2024). Academic research on "situationships" is scarce (though see Langlais et al. 2024, Armstrong et al. 2024), and it is unclear whether they constitute a genuinely new form of romantic engagement. Furthermore, while some people might have internalized a genuinely novel set of behavioral expectations, they may be unaware of the label "situationship," making situationships hard to study. If we accept the argument outlined in this paper that frames arise from the duality of dyad and content, however, the validity of this new concept can be defined. Specifically, we would be able to determine whether there is indeed a new cluster in relationship-attribute space, when it emerged, what attributes characterize it, and so forth.

What this example also illustrates is that there may be times when cultural models about relationship frames are fundamentally uncertain. This gets back at White and colleagues' (2008; White, Godart, and Thiemann 2008) distinction between uncertainty concerning cultural rules ('ambiguity'), and uncertainty as to whether the enactment in a specific instance will correspond to these rules ('ambage'). The present study effectively treats relationship frames as invariant cultural structures, and models uncertainty in concrete instantiations of ties premised on this stability. A diachronic perspective on relationship frames could allow us to distinguish historical contexts in which the relationship between content and frames is relatively weaker or stronger.

Second, the measurement formalization of frame ambiguity proposed above is strictly synchronic one. In other words, it does not account for *narrative time*, let alone the possibility of multiple coexisting temporal registers in a work of fiction (Padgett 2018; also see Padgett et al. 2020). It is adequate insofar as we look at relationships in their entirety. Yet, if we consider their narrative quality, we might find that frame ambiguity changes throughout the story. A time-variant formalization of ambiguity exceeds the focus of this study and may be an interesting path for future work. It might, for instance, allow us to observe "switchings" (White 1995; Mische and White 1998; Godart and White 2010; or "stitchings," see Fuhse 2023), in which identities shift from actions that cue one frame to ones cueing another. Furthermore, such a measure would come closer to the notions of ambiguity and multivocality put forward by Leifer (1983 1988), Padgett, and Ansell (1993), which strongly depend on temporality, on suspense, and holding open future options. Such a formalization could be developed from the same basic concepts offered here and would not necessitate challenging the notion of relation frames or frame ambiguity.

Conclusion

Recent work has pointed to the potential of "frames" (Goffman 1974) as a concept for theorizing social relationships (Fuhse 2021; Lizardo 2024). Thinking about tie meanings through frames rather than types allows us to connect them to the level of practice and cognition and, thus, to the rich theoretical architectures developed for thinking about these dimensions. In this paper, I make conceptual, methodological, and empirical contributions to this line of work. *Conceptually*, I expanded on previous arguments by further developing the notion of relationship frames (Fuhse 2021). Specifically, I argued that such frames are rooted in regularities in the duality of dyad and content. This gives the notion of relationship frames a relatively simple and empirically tractable concretization. I then built on this conceptualization to develop a notion of frame ambiguity, a new way of thinking about the ways in which interaction in ties draws on multiple meanings that allows for a more dynamic, practice-oriented perspective on relational meaning, especially in cases where relationships are unstable, contested, or emergent. *Methodologically*, I demonstrate how to operationalize frames and frame ambiguity using mixture modeling techniques. Finally, I conducted an *empirical* analysis of the frame structure of interaction in literary relationships, identifying the dominant frames, investigating their prevalences, and examining norms around reciprocity. I further showed that frame ambiguity is related to character importance and gender. Specifically, more important characters both act and are acted towards in more diverse ways. Simultaneously, interactions among male characters display less frame ambiguity than those that involve women.

What is the practical relevance of the case for relationship frames and for studying the duality of dyad and content? Is this perhaps but an impractical call for nuance? The notion of types of ties has been highly productive for sociology, and the reality is that nearly all network data we have was collected in this form, that is, with a relatively limited set of potential attributes for each tie. However, with the increasing digitization of social spheres and the advance of "large-scale, unobtrusive data collection techniques" (Lizardo 2024, p. 138), this might change. So far, these developments have largely led network analysts to study larger networks. "Big data" has typically meant "big networks." Going forward, it might mean "high-def ties," that is, having rich data about the content of social relationships. Whereas in-depth analyses of relationships have thus far been the domain of qualitative researchers, this development could herald the emergence of a new *formal study of relationship content*.

Examples pointing in this direction can be found in recent studies that examine the content of written communication within organizations (Goldberg et al. 2016; Bhatt et al. 2022; Aral & Van Alstyne 2011). This kind of data allows for a direct examination of the symbols in play within a given tie, which makes it possible to study at scale what has previously only been studied qualitatively or in smaller contexts. For instance, what are the structural conditions under which people draw on particular frames in their communication (McLean 1998), what communication strategies are effective at retaining flexibility (Padgett and Ansell 1993), how does something like an idiosyncratic "relationship culture" (Fuhse 2009) that goes beyond culturally shared frames emerge in social ties, or how do people communicate to achieve or respond to tie change (McFarland & Wolff 2022)?

Another avenue for this kind of work is to use rich data recorded from social interaction. McFarland and colleagues (2013, 2024), for instance, study the emergence of romantic relationships in the context of first dates. Their studies leverage a rich set of features derived from both the full spoken content of the dialogue as well as its prosodic qualities. In a similar vein, Voigt et al. (2017) use video and footage from police body cameras to investigate linguistic features of officer-driver interactions. Palotti, Weldon, and Lomi (2022) study audio-visual recordings from surgeries in a London hospital involving interactions and communication between surgeons, nurses, anesthetists, and others. Meanwhile, Fuhse and colleagues (2020) study the relationships among parties in the Weimar Reichstag by examining the content of all utterances made during their respective opponents' speeches. An older yet remarkable example of this kind of work is Gibson's investigation of turn-taking behavior in business meetings, as it connects the level of interaction with that of higher-order structural data (Gibson 2005).

Almost 30 years ago, Mische and White (1998) noted an unfortunate lack of engagement between network analysts and sociolinguists. "One reason for this," they pointed out, "is the absence of theoretical understanding of the commingling of network relations and discursive processes" (p. 695). While much of the work cited here shows that this call was taken seriously, the intersection has remained a difficult research frontier, in part due to data constraints. There is good reason to expect that the new wealth of information on communication in (or about) social ties will reinvigorate this area of research. Beyond new methods, this will require conceptual work at the intersection of network analysis on the one hand and interaction and cognition on the other; concepts, in other words, that replace abstract notions like "type of tie" and describe the relation between the larger social structures network analysis typically studies and the level of practice and conversation. The present paper contributes to this paradigm by advancing the theorization of relationship frames and introducing frame ambiguity as a concept for studying complex tie meanings.

References

Aral, S., & Van Alstyne, M. (2011). The diversity-bandwidth trade-off. *American Journal of Sociology, 117*(1), 90-171.

Armstrong, E. A., Hamilton, L. T., Garrison, S. A., Giles, K. N., Hoffman, C. M., & Perone, A. K. (2024). "It's Complicated": How Black and White Women Innovate with Situationships at Midlife. Social Problems. https://doi.org/10.1093/socpro/spae021

Basov, N., Breiger, R., & Hellsten, I. (2020). Socio-semantic and other dualities. Poetics, 78, 101433.

Basov, N., & Brennecke, J. (2017). Duality beyond dyads: Multiplex patterning of social ties and cultural meanings. In Structure, Content and Meaning of Organizational Networks: Extending Network Thinking (pp. 87-112). Leeds: Emerald Publishing Limited.

Basov, N., de Nooy, W., & Nenko, A. (2021). Local meaning structures: mixed-method sociosemantic network analysis. American journal of cultural sociology, 9, 376-417.

Basov, N., Lee, J. S., & Antoniuk, A. (2016, November). Social networks and construction of culture: A socio-semantic analysis of art groups. In International Workshop on Complex Networks and their Applications (pp. 785-796). Cham: Springer International Publishing.

Basov, N., Lee, J.-S. & Antoniuk, A. (2016). Complex Networks & Their Applications V: Proceedings of the 5th International Workshop on Complex Networks and their Applications (COMPLEX NETWORKS 2016). Cherifi, H., Gaito, S., Quattrociocchi, W. & Sala, A. (eds.). Cham: Springer Nature, p. 785-796 12 p. (Studies in Computational Intelligence; vol. 693).

Basov, N., & Roth, C. (2025). Socio-semantic networks. In Handbook of Culture and Social Networks (pp. 251-270). Cheltenham: Edward Elgar Publishing.

Bandelj, N. (2020). Relational Work in the Economy. Annual Review of Sociology, 46, 251–272.

Bamman, D., Lewke, O. Lewke, & Mansoor, A. (2020). An annotated dataset of coreference in English literature. in N. Calzolari et al (Eds.) Proceedings of the 12th Conference on Language Resources and Evaluation (pp. 44-54). Marseille, 11-16 May 2020.

Bamman, D. (2023). BookNLP. Version 1.0.7 https://github.com/booknlp/booknlp. (Accessed July 23, 2024)

Barsalou, L. W. (1992). Frames, concepts, and conceptual fields. In A. Lehrer, E. F. Kittay, & R. Lehrer (Eds.), Frames, fields, and contrasts: New essays in semantic and lexical organization (pp. 21–74). New York: Routledge.

Bearman, P., & Parigi, P. (2004). Cloning Headless Frogs and Other Important Matters: Conversation Topics and Network Structure. Social Forces, 83(2), 535–557.

Bellodi, L. (2025). Bureaucratic information in Congress. Retrieved from https://lucabellodi.com/material/Bureaucratic_Information_in_Congress.pdf

Berger, P. L., & Luckmann, T. (1991[1966]). The Social Construction of Reality. A Treatise in the Sociology of Knowledge. London, England: Penguin.

Bhatt, A. M., Goldberg, A., & Srivastava, S. B. (2022). A language-based method for assessing symbolic boundary maintenance between social groups. *Sociological Methods & Research, 51*(4), 1681-1720.

Bianchi, F., Filippi-Mazzola, E., Lomi, A., & Wit, E. C. (2024). Relational event modeling. *Annual Review of Statistics and Its Application*, 11.

Bianchi, F., & Lomi, A. (2023). From ties to events in the analysis of interorganizational exchange relations. Organizational Research Methods, 26(3), 524-565.

Blau, P. (1964). Exchange and power in social life. New York: Wiley and Sons.

Breiger, R. L. (2000). A tool kit for practice theory. Poetics, 27(2-3), 91-115.

Breiger, R., Mützel, S. (2020). Duality beyond Persons and Groups: Culture and Affiliation. In R. Light, J. Moody (Eds.) *The Oxford Handbook of Social Networks* (Chapter 21). Oxford: Oxford University Press.

Breiger, R., Wagner-Pacifici, R. (2023). Social Networks and Social Categories. In J. McLevey, J. Scott, & P. J. Carrington (Eds.) The Sage Handbook of Social Network Analysis (pp. 32-43). Thousand Oaks, CA: Sage.

Burke, K. (1945). A Grammar of Motives. Berkeley, CA: University of California Press.

Burt, R. S. (1997). A note on social capital and network content. Social networks, 19(4), 355-373.

Burt, R. S. (1982). Chapter 2. Distinguishing relational contents. In R. Burt and M. J. Minor (Eds.) Applied network analysis: a methodological introduction (pp. 35-74). Beverly Hills: Sage.

Burt, R. S., & Schøtt, T. (1985). Relation contents in multiple networks. Social Science Research, 14(4), 287-308.

Butts, C. T. (2008). A relational event framework for social action. *Sociological Methodology*, 38(1), 155-200.

Butts, C. T., Lomi, A., Snijders, T. A., & Stadtfeld, C. (2023). Relational event models in network science. *Network Science*, 11(2), 175-183.

Carley, K. (1993). Coding choices for textual analysis: A comparison of content analysis and map analysis. Sociological methodology, 75-126.

Cheng, J. (2020). Fleshing out models of gender in English-language novels (1850–2000). Journal of cultural analytics, 5(1).

Chicago Daily Herald (2024). The Situationship: teen dating is not what it once was. April 14, 2024. Sunday ML2 Edition.

Choi, J. D., Palmer M. (2012). *Guidelines for the Clear Style Constituent to Dependency Conversion*. https://www.mathcs.emory.edu/~choi/doc/cu-2012-choi.pdf

Cederholm, E., & Åkerström, M. (2017). With a little help from my friends: Relational work in leisure-related enterprising. The Sociological Review 64(4), 748-765.

Cross, R., Borgatti, S. P., & Parker, A. (2001). Beyond answers: Dimensions of the advice network. Social networks, 23(3), 215-235.

Das, D., Chen, D., Martins, A. F., Schneider, N., & Smith, N. A. (2014). Frame-semantic parsing. Computational linguistics, 40(1), 9-56.

DiMaggio, P. (1993). Nadels paradox revisited: Relational and cultural aspects of organizational structure. In N. Nohria, & R. Eccles (Eds.), Networks and Organization (pp. 118–142). Boston, MA: Harvard Business School Press.

DiMaggio, Paul (2020). *Duality and relationality: The cultural matrix of John W. Mohr. Poetics 78*.

De Marneffe, M. C., & Nivre, J. (2019). Dependency grammar. Annual Review of Linguistics, 5(1), 197-218.

De Nooy, W. (2005). Stories, scripts, roles, and networks. Structure and Dynamics 1(2).

Doehne, M., McFarland, D. A., & Moody, J. (2024). Network ecology: introduction to the special issue. Social Networks, 77, 1-4.

Emirbayer, M., & Goodwin, J. (1994). Network analysis, culture, and the problem of agency. American Journal of Sociology, 99(6), 1411–1454.

Emirbayer, M. (1997). Manifesto for a relational sociology. American Journal of Sociology, 103(2), 281-317.

Etzioni, O., Banko, M., Soderland, S., & Weld, D. S. (2008). Open information extraction from the web. Communications of the ACM, 51(12), 68-74.

Fischer, C. S. (1982a). What do we mean by 'friend'? An inductive study. Social Networks, 3(4), 287-306.

Fischer, C.S., (1982b). To Dwell Among Friends: Personal Networks in Town and City. Chicago: University of Chicago Press.

Fine, G. A. (1979). Small groups and culture creation: The idioculture of little league baseball teams. American sociological review, 733-745.

Fine, G. A., & Kleinman, S. (1983). Network and meaning: An interactionist approach to structure. Symbolic interaction, 6(1), 97-110.

Franzosi, R. (1989). From words to numbers: A generalized and linguistics-based coding procedure for collecting textual data. Sociological methodology, 263-298.Fuhse, J. (2009). The meaning structure of social networks. Sociological Theory, 27(1), 51-73.

Fuhse, J. A. (2009). The meaning structure of social networks. Sociological theory, 27(1), 51-73.

Fuhse, J. (2021). Social networks of meaning and communication. Oxford: Oxford University Press.

Fuhse, J. A. (2023). Stitching the social world: The processing of multiple social structures in communication. *Socius, 9*,

Fuhse, J. A. (2013). Social relationships between communication, network structure, and culture. In F. Dépelteau & C. Powell (Eds.), Applying relational sociology: Relations, networks, and society (pp. 181-206). New York, NY: Palgrave Macmillan US.

Fuhse, J. A., & Gondal, N. (2024). Networks from culture: Mechanisms of tie-formation follow institutionalized rules in social fields. Social Networks, 77, 43-54.

Fuhse, J., Stuhler, O., Riebling, J., & Martin, J. L. (2020). Relating social and symbolic relations in quantitative text analysis: A study of parliamentary discourse in the Weimar Republic. *Poetics, 78*, 101363.

Gibson, D. R. (2005). Taking turns and talking ties: Networks and conversational interaction. *American Journal of Sociology, 110*(6), 1561-1597.

Gildea, D., & Jurafsky, D. (2002). Automatic labeling of semantic roles. Computational linguistics, 28(3), 245-288.

Godart, F. C., & White, H. C. (2010). Switchings under uncertainty: The coming and becoming of meanings. Poetics, 38(6), 567-586.

Goffman, E. (1974). Frame analysis: An essay on the organization of experience. New York, NY: Harper Colophon Books.

Goldberg, A., Srivastava, S. B., Manian, V. G., Monroe, W., & Potts, C. (2016). Fitting in or standing out? The tradeoffs of structural and cultural embeddedness. *American Sociological Review, 81*(6), 1190-1222.

Goldenstein, J., & Poschmann, P. (2019). Analyzing meaning in big data: Performing a map analysis using grammatical parsing and topic modeling. Sociological Methodology, 49(1), 83-131.

Gondal, N. (2022). Multiplexity as a lens to investigate the cultural meanings of interpersonal ties. Social Networks, 68, 209-217.

Gondal, N., & McLean, P. D. (2013). Linking tie-meaning with network structure: Variable connotations of personal lending in a multiple-network ecology. Poetics, 41(2), 122-150.

Honnibal, M., Montani, I., Van Landeghem, S., & Boyd, A. (2023). SpaCy: Industrial-strength natural language processing in Python (Version 3.4.1). Explosion AI. https://zenodo.org/records/10009823 (Accessed November 1, 2023)

Illouz, E. (2012). Why Love Hurts. Cambridge: Polity.

Jamieson, L. (1998). Intimacy: Personal Relationships in Modern Societies. Cambridge: Polity.

Jockers, M., & Kirilloff, G. (2016). Understanding gender and character agency in the 19th century novel. Journal of Cultural Analytics, 2(2).

Jones, C., Maoret, M., Massa, F. G., & Svejenova, S. (2012). Rebels with a cause: Formation, contestation, and expansion of the de novo category "modern architecture," 1870–1975. Organization science, 23(6), 1523-1545.

Jurafsky, D., Martin H. (2024). Speech and Language Processing: Chapter 19. Dependency Parsing. https://web.stanford.edu/~jurafsky/slp3/19.pdf (accessed 03/12/2025)

Kitts, J. A., & Leal, D. F. (2022). What is (n't) a friend? Dimensions of the friendship concept among adolescents. Social Networks, 69, 211-220.

Knight, C. (2022). When corporations are people: Agent talk and the development of organizational actorhood, 1890–1934. Sociological Methods & Research, 51(4), 1634-1680.

Kraicer, E., & Piper, A. (2019). Social characters: The hierarchy of gender in contemporary English-language fiction. Journal of Cultural Analytics 3(2).

Langlais, M. R., Podberesky, A., Toohey, L., & Lee, C. T. (2024). Defining and describing situationships: An exploratory investigation. Sexuality & Culture, 28(1), 1-27.

Lee, M., & Martin, J. L. (2018). Doorway to the dharma of duality. Poetics, 68, 18-30.

Lee, K., Paci, S., Park, J., You, H. Y., & Zheng, S. (2025). Applications of GPT in Political Science Research. PS: Political Science & Politics. Preprint available at https://hyeyoungyou.com/wp-content/uploads/2024/08/gpt.pdf

Leifer, E. (1983). Robust Action: Generating Joint Outcomes in Social Relationships. Dissertation submitted to the Department of Sociology in May 1983, Harvard University, Cambridge, MA.

Leifer, E. (1988). Interaction preludes to role setting: Exploratory local action. American Journal of Sociology, 53(6), 865-878.

Marsden, P. V. (1987). Core discussion networks of Americans. American Sociological Review, 52(1), 122-131.

Martin, J. L. (2009). Social structures. Princeton, NJ: Princeton University Press.

McFarland, D. A., Jurafsky, D., & Rawlings, C. (2013). Making the connection: Social bonding in courtship situations. *American Journal of Sociology, 118*(6), 1596-1649.

McFarland, D. A., & Wolff, T. (2022). Writing into relationships. Social Networks, 71, 96-114.

McFarland, D. A., Broska, D., Prabhakaran, V., & Jurafsky, D. (2024). Coming into relations: How communication reveals and persuades relational decisions. *Social Networks, 79*, 57-75.

McLean, P. (2017). Culture in networks. Polity Press.

McLean, P. (1998). A frame analysis of favor seeking in the Renaissance: Agency, networks, and political culture. American Journal of Sociology, 104(1), 51-91.

Mears, A. (2015). Working for free in the VIP. American Sociological Review, 80(6), 1099-1122.

Mische, A., & White, H. (1998). Between conversation and situation: Public switching dynamics across network domains. Social Research 65(3), 695-724.

Mohr, J. & Duquenne V. (1997). The duality of culture and practice: poverty relief in New York City, 1888-1917. Theory and Society, 26(2/3), 305–356.

Mohr, J. (1998). Measuring Meaning Structures. *Annual Review of Sociology*, 24, 345–370.

Mohr, J. W., & Lee, H. K. (2000). From affirmative action to outreach: Discourse shifts at the University of California. Poetics, 28(1), 47-71.

Mohr, J. W., Wagner-Pacifici, R., Breiger, R. L., & Bogdanov, P. (2013). Graphing the grammar of motives in National Security Strategies: Cultural interpretation, automated text analysis and the drama of global politics. Poetics, 41(6), 670-700.

Mohr, J. W., & White, H. C. (2008). How to model an institution. Theory and Society, 37, 485-512.

Moretti, F. (2011). Network theory, plot analysis. Stanford Literary Lab Pamphlet 2.

Moretti, F. (2013). Distant reading. Verso Books.

Mützel, S. (2022). Making sense: markets from stories in new breast cancer therapeutics. Stanford University Press.Schütz, 1932 p. 40; see also Schütz and Luckmann, 1979, pp. 281-283)

Niezink, N. M., & Campana, P. (2023). When things turn sour: A network event study of organized crime violence. Journal of Quantitative Criminology, 39(3), 655-678.

Padgett, J. F., & Ansell, C. K. (1993). Robust action and the rise of the Medici, 1400-1434. American Journal of Sociology, 98(6), 1259-1319.

Palotti, Francesca, Weldon, S. M., & Lomi, A. (2022). Lost in translation: Collecting and coding data on social relations from audio-visual recordings. Social Networks, 69, 102-112.

Park B. (2020). Chapter 4: Role Ambiguity and (Dis)integration in Naturally Occurring Communities. In Dissertation at New York University.

Pachucki, M. A., & Breiger, R. L. (2010). Cultural holes: Beyond relationality in social networks and culture. Annual review of sociology, 36(1), 205-224.

Piper, A. (2018). Enumerations: data and literary study. University of Chicago Press.

Piper, A., Xu, M., & Ruths, D. (2024). The Social Lives of Literary Characters: Combining citizen science and language models to understand narrative social networks. In Proceedings of the 4th International Conference on Natural Language Processing for Digital Humanities (pp. 472-482).

Roberts, C. W. (1997). A generic semantic grammar for quantitative text analysis: Applications to East and West Berlin radio news content from 1979. Sociological Methodology, 27(1), 89-129.

Padgett, J. (2018). Faulkner's Assembly of Memories into History: Narrative Networks in Multiple Times. American Journal of Sociology, 124(2), 406-478.

Padgett, J. F., Prajda, K., Rohr, B., & Schoots, J. (2020). Political discussion and debate in narrative time: The Florentine Consulte e Pratiche, 1376–1378. Poetics, 78, 101377.

Propp V. ([1928] 1968). *Morphology of the Folktale*. Austin: University of Texas Press.

Radway, I. (1991). Reading the Romance. Women, Patriarchy, and Popular Literature. Chapel Hill and London: The University of North Carolina Press.

Sanders, S. A., & Reinisch, J. M. (1999). Would you say you had sex if...?. JAMA, 281(3), 275-277.

Sanders, S. A., Hill, B. J., Yarber, W. L., Graham, C. A., Crosby, R. A., & Milhausen, R. R. (2010). Misclassification bias: Diversity in conceptualisations about having 'had sex'. Sexual Health, 7(1), 31-34.

Schmitt, M., & Fuhse, J. (2015). Zur Aktualität von Harrison White: Einführung in sein Werk. Springer-Verlag.

Simmel G. (1890[1908]). Die Kreuzung Sozialer Kreise. In G. Simmel: *Soziologie: Untersuchungen über die Formen der Vergesellschaftung* (pp. 403-453). Leipzig: Duncker & Humblot.

Sims, M., & Bamman, D. (2020). Measuring information propagation in literary social networks. arXiv preprint arXiv:2004.13980.

Small, M. L. (2017). Someone To Talk To. Oxford: Oxford University Press.

Stuhler, O. (2021). What's in a category? A new approach to Discourse Role Analysis. Poetics, 88, 101568.

Stuhler, O. (2022a). semgram: Extracting Semantic Motifs from Textual Data. R package Version 0.1.0. 10.32614/CRAN.package.semgram

Stuhler, O. (2022b). Who does what to whom? Making text parsers work for sociological inquiry. Sociological Methods & Research, 51(4), 1580-1633.

Stuhler, O. (2024). The gender agency gap in fiction writing (1850 to 2010). Proceedings of the National Academy of Sciences, 121(29), e2319514121.

Stuhler, O., Ton, C. D., & Ollion, E. (2025). From Codebooks to Promptbooks: Extracting Information from Text with Generative Large Language Models. Sociological Methods & Research, 00491241251336794.

Sudhahar, S., De Fazio, G., Franzosi, R., & Cristianini, N. (2015). Network analysis of narrative content in large corpora. Natural Language Engineering, 21(1), 81-112.

Swidler, A. (2001). Talk of love: How culture matters. Chicago, IL: The University of Chicago Press.

Tilly, C. (1997). Parliamentarization of popular contention in Great Britain, 1758-1834. Theory and Society, 26(2/3), 245-273.

Underwood, T. (2019). Distant horizons: digital evidence and literary change. University of Chicago Press.

Underwood, T., Kimutis, P., & Witte, J. (2020). NovelTM datasets for English-language fiction, 1700-2009. Journal of Cultural Analytics, 5.

Underwood, T., Bamman, D., & Lee, S. (2018). The transformation of gender in English-language fiction. Journal of Cultural Analytics, 3(2).

van Atteveldt, W., Sheafer, T., Shenhav, S. R., & Fogel-Dror, Y. (2017). Clause analysis: Using syntactic information to automatically extract source, subject, and predicate from texts with an application to the 2008–2009 Gaza war. Political Analysis, 25(2), 207–222.

Verbrugge, L. M. (1979). Multiplexity in adult friendships. Social Forces, 57(4), 1286-1309.

Voigt, R., Camp, N. P., Prabhakaran, V., Hamilton, W. L., Hetey, R. C., Griffiths, C. M., ... & Eberhardt, J. L. (2017). Language from police body camera footage shows racial disparities in officer respect. *Proceedings of the National Academy of Sciences, 114*(25), 6521-6526.

Vörös, A., & Snijders, T. A. B. (2017). Cluster analysis of multiplex networks: Defining composite network measures. Social Networks, 49, 93-112.

Vu, D., Pattison, P., & Robins, G. (2015). Relational event models for social learning in MOOCs. *Social Networks*, 43, 121-135.

Wijffels, J. (2020). Biterm Topic Models for Short Text. R Package Version 0.3.7 Available at: 10.32614/CRAN.package.BTM (Accessed July 17 2024).

White, H. (1992). Identity and control: A structural theory of social action. Princeton, NJ: Princeton University Press.

White, H. C. (2008). Identity and control: how social formations emerge. 2nd ed. Princeton University Press.

White, H. C. (1995). Network switchings and Bayesian forks: Reconstructing the social and behavioral sciences. Social Research, 62(4), 1035-1063.

White, H. C. (1998). Between Conversation and Situation: Public Switching Dynamics across Network Domains. Social Research, 65(3), 695-724.

White, H. C., Godart, F. C., & Thiemann, M. (2013). Turning points and the space of possibles: A relational perspective on the different forms of uncertainty. In Applying Relational Sociology: Relations, Networks, and Society (pp. 137-154). New York: Palgrave Macmillan US.

Yan, X., Guo, J., Lan, Y., & Cheng, X. (2013). A biterm topic model for short texts. In Proceedings of the 22nd international conference on World Wide Web (pp. 1445-1456).

Yeung, K.-T. (2005). What does love mean? Exploring network culture in two network settings. Social Forces, 84(1), 391-420.

Zelizer, V. A. (2000). The purchase of intimacy. Law & Social Inquiry, 25(3), 817-848.

Appendix

Table of Contents

Appendix A. BTM estimation, omitted frames, and frame details

A detailed explanation of the BTM model can be found in the original paper by Yan and colleagues (2013). Here I only provide a very concise description of the model's central intuition and formal properties. To be modeled with the BTM, action profiles are first transformed into a multiset of biterms B – unordered action pairs. For instance, an action profile d with three distinct actions a will generate three action-biterms b:

$$d = (a_1, a_2, a_3) \Longrightarrow \{b_1 = (a_1, a_2), b_2 = (a_2, a_3), b_3 = (a_1, a_3)\} \quad (1)$$

After this transformation, the unit of action profiles plays no direct role in the modeling process. As mentioned in the main text, the BTM models the generative process of biterms, not documents (i.e., action profiles), like conventional topic models. To infer latent frame variables (i.e., topic variables), the model assumes that the co-occurrence of two actions in biterms indicates that these features were generated by the same latent frame. The assumed generative process is the following:

1. Draw a frame distribution $\theta \sim \text{Dirichlet}(\alpha)$ for the whole data.
2. For each frame $k \in [1, K]$
    a. draw frame-action distribution $\phi_k \sim \text{Dirichlet}(\beta)$.
3. For each biterm $b_i \in \text{B}$
    a. draw frame assignment $z_i \sim \text{Multinomial}(\theta)$, and
    b. draw two actions $a_{i,1}, a_{i,2} \sim \text{Multinomial}(\phi_{z_i})$.

Parameter $\mathbf{\Phi}$ is a $K \times W$ matrix containing W-dimensional multinomial distributions for each frame with $\phi_{k,w}$ equaling $P(a|z = k)$, that is, the probability to observe action a given frame k. θ is a K-dimensional multinomial distribution for the probabilities of frames in the corpus. Hyperparameters α and β control the sparsity of the distributions of frames in the data θ and the action-frame distributions ϕ, respectively. Hyperparameter K sets the number of frames (see below). Parameters $\mathbf{\Phi}$ and θ are estimated via Gibbs sampling. For estimation, the R implementation of the model by Wijffels (2020) was used (BTM version 0.3.7).

While the BTM does not directly estimate the extent to which each action profile is associated with each frame, one can estimate the distribution of latent role variables most likely to have generated an action profile by leveraging the estimates of the action probabilities under each frame $P(a|z)$. The probability of a frame given an action profile $P(z|d)$ is:

$$P(z|d) = \sum_m P(z|m)P(m|d) \quad (2)$$

$P(z|a)$ can be derived using the conditional action probabilities and frame probabilities contained in $\mathbf{\Phi}$ and θ, respectively:

$$P(z|a) = \frac{P(z)P(a_i|z)}{\sum_z P(z)\,P(a_i|z)} \tag{3}$$

and $P(a|d)$ simply equals the relative frequency of a specific action in an action profile.

To find an appropriate value for K, I rely on both formal criteria and manual evaluation. For the formal part, I first split the data into a training portion (80%) and a held-out portion (20%). I then fit a series of models for different values of K using the training data. For each of these, I specified one frame to serve as "background" frame that equals the empirical word distribution. This has been shown to aid in filtering out common words (see Wijffels 2020 for details). I note that each of these models were estimated using a High Performance Computing cluster and each of them would take several hours to run. Subsequently, the likelihood of the held-out portion was computed under each model. This is a widely used approach to evaluating topic model fit. Figure A1 present the held-out log likelihoods for different values of K. These results do not indicate an obvious choice for K. However, they suggested that models' capacity to predict unseen data increases fairly rapidly from 0 to about 20 but slows down thereafter. In a second step, I examined the frame variables (i.e., $P(a|z)$) for interpretability of the clusters. I found that a solution with 25 frames provided a good tradeoff and produced generally interpretable results. Two of the frames were omitted because they seemed uninterpretable. I report the FREX action for these omitted frames in table A1. Note that the background frame is not contained in the table because, unsurprisingly, none of the actions there pass the FREX threshold.

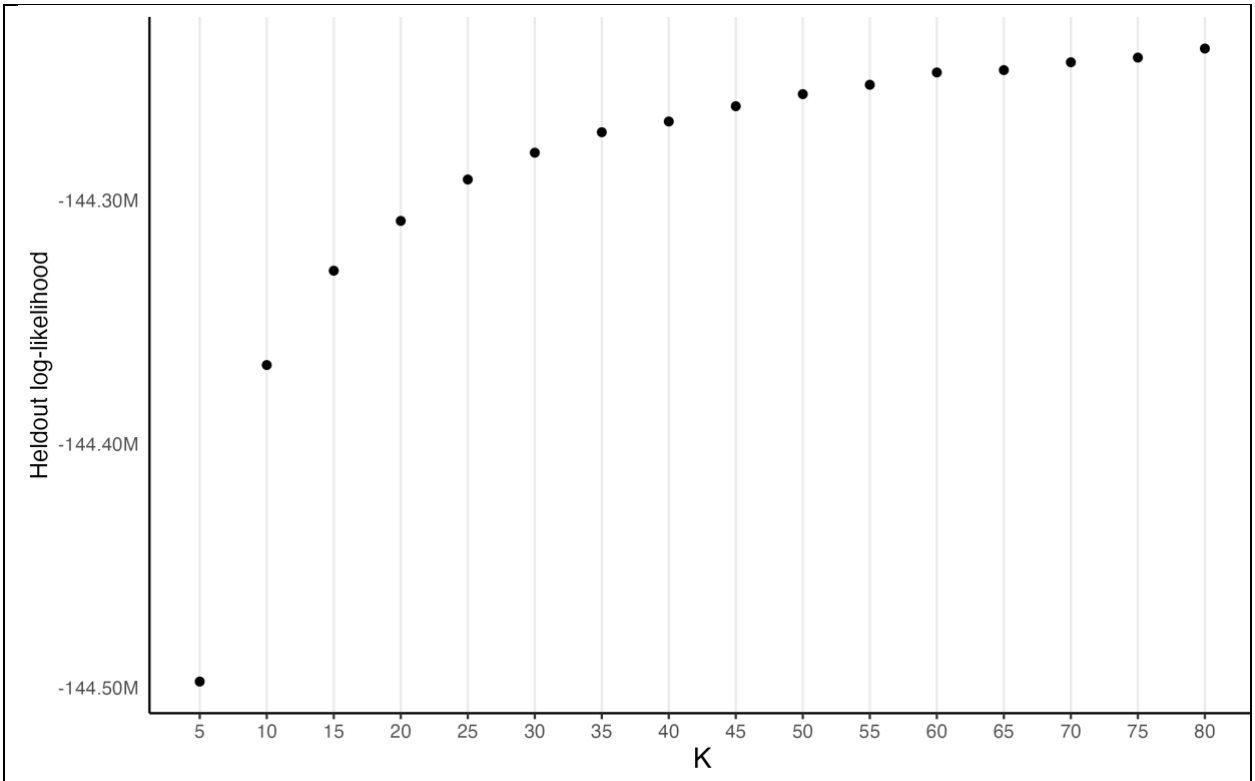

Figure A1. Held-out log likelihood for different values of K.

Relationship frame	Actions
Omitted frame I	regard, treat, invite, greet, address, approach
Omitted frame II	know, like, trust, do, understand, help, say, think, know-about, mean

Table A1. Relationship frames and associated actions for omitted frames

Relationship frame	Actions
Observe	glance-at, look-at, smile-at, stare-at, turn-to, grin-at, study, gaze-at, glare-at, eye, turn-toward, face, regard, watch, hand-to, hand, look, lean-toward, ignore, peer-at, frown-at, hold-to, turn-on
Know & think-of	see, remember, think-of, know, think-about, meet, imagine, forget, look-for, tell-about, recognize, find, hear-of, know-about, picture, watch, like, talk-about, speak-to, ask-about, run-into, mention, hear-from, lose, notice, talk-to, talk-with, feel, hear-about, see-of
Dislike & forgive	hate, forgive, believe, despise, understand, trust, listen-to, hurt, blame, accuse, pity, believe-in, like, dislike, be-like, hear, be-to, envy, betray, think-of
Converse-with I	say-to, turn-to, whisper-to, look-at, explain-to, call-to, ask, say, go-to, answer, smile-at, sit-with, speak-to, shout-at, come-to, point-to, go-with, tell, talk-to, hand-to, wake, shout-to

Give-to	give, send, pay, show, teach, get, keep, buy, want, owe, make, offer, treat, put, do, hand, bring, order, lend, let, take, beat, do-for, use, drive, sell, promise, hire
Marriage & courtship	marry, like, care-for, love, refuse, accept, meet, hate, treat, dance-with, admire, think-of, dislike, want, write-to, be-to, forgive, forget, know, propose-to, go-with, think, despise, laugh-at
Distant communication	write-to, write, hear-from, visit, ring, send, think-of, go-to, meet, receive-from, telephone, spend-with, persuade, call-on, phone, accompany, miss, send-to, learn-from, live-with, see, discuss-with
Love & romance	love, want, lose, kiss, miss, forget, think-of, need, hurt, adore, marry, care-for, feel-for, win, write-to, see, deceive, write, think-about, hate, hold, belong-to, be-with, worship, forgive, be-to, dream-of, make-to
Tell & assure I	remind, assure, inform, warn, point-to, urge, ask, explain-to, advise, order, persuade, eye, reassure, protect, convince, agree-with, instruct, study, suggest-to, question, regard, encourage, discuss-with, join, accuse
Beat & kick	hit, grab, push, slap, beat, drag, pull, kick, shake, shove, glare-at, throw, strike, catch, punch, fuck, do-to, ignore, yell-at, shout-at, knock, force, stare-at, nudge, turn, stand-over
Tell & assure II	tell, call, talk-to, explain-to, show, lie-to, assure, trust, warn, believe, hear, promise, phone, get, come-to, ask, answer, talk-with, agree-with, know, listen-to, mention-to, keep-from
Hold authority over	send, command, teach, order, spare, summon, call, bid, call-to, instruct, bless, send-for, speak-with, grant, save, scold, dismiss, forbid, forgive, speak-to, receive, lead, permit
Lead & guide	lead, show, hand, teach, introduce, pull, hand-to, guide, help, put-around, hug, wink-at, nudge, walk, join, invite, whisper-to, explain-to, grin-at, steer, smile-at, take, drive, pat
Physical affection	kiss, hold, pull, put-around, touch, draw, lift, hug, release, want, hurt, carry, grab, make-to, embrace, take, lead, reach-for, smile-at, catch, push, turn, press, enter, lie-beside, pick, sit-beside, gather, lay, guide
Physical care	carry, take, lift, pick, put, lay, feed, hold, set, bring, place, leave, get, sit, keep, buy, take-of, drop, raise, catch, lead, nurse, look-after, cover, drag, help, drive, read-to, gather, draw
Shoot & kill	kill, shoot, murder, hate, hit, do-to, beat, strike, hurt, kick, fight, throw, threaten, grab, push, attack, knock, stare-at
Search	find, look-for, spot, ride-with, go-for, get, come-for, save, lose, play, seek, believe, follow, talk-to, rescue, miss, run-into
Long-for	call, miss, need, think-about, want, talk-to, sleep-with, phone, live-with, stay-with, spend-with, be-with, tell-about, talk-about, get, do-to, lie-to, imagine, owe, lose, care-about
Converse-with II	ask, call, talk-to, remind, invite, look-at, hug, think-about, listen-to, smile-at, thank, phone, pick, believe, sleep-with, wake, go-with, wait-for, like, miss, hear, come-with, spend-with, walk, meet, stay-with, drive, follow
Motion-with & motion-to	follow, watch, join, glance-at, call-to, catch, hear, stand-beside, catch-with, take-from, approach, reach, gaze-at, sit-beside, run-after, run-to, go-to, go-with, hand-to, stare-at, whisper-to, draw, come-with, pass
Thank, beg, & promise	thank, beg, assure, pray, cry, beg-of, entreat, wish, promise, serve, implore, bid, inform, swear-to, advise, receive, reply, accompany, do, say, understand, exclaim, obey, congratulate, request
Be-with	come-to, leave, be-with, go-to, stay-with, come-with, go-with, marry, be-to, live-with, stand-before, sit-beside, belong-to, cling-to, read-to, wait-for, care-for, stand-beside, bring, draw

Table A2. Relationship frames and associated actions.
Note: This table replicates Table 1 from the main manuscript but expands the FREX score cutoff to 50 instead of 30, leading to an average of 24.1 listed actions per frame. The actions are ordered according to their FREX score such that the first actions are more likely and exclusive to the respective frame. Note that some of the last-mentioned actions may not appear central to the frame. This has to do with the fact that the total number of actions used in the analysis is just 500, so this expansion necessarily goes beyond the very core of the frame.

Appendix B. Hierarchical clustering of relationship frames

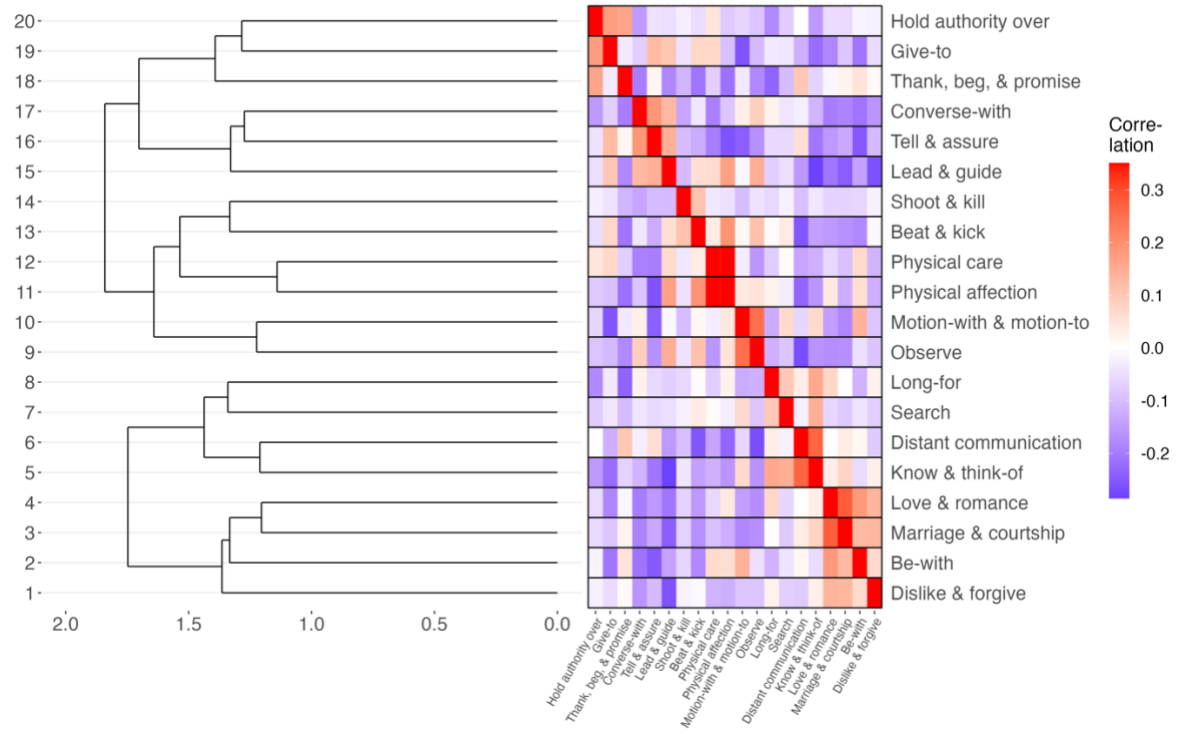

Figure A2. Association of relationship frames within action profiles (complete tree). Note: This figure replicates Figure 2 of the main manuscript but shows the complete tree derived from hierarchical clustering.

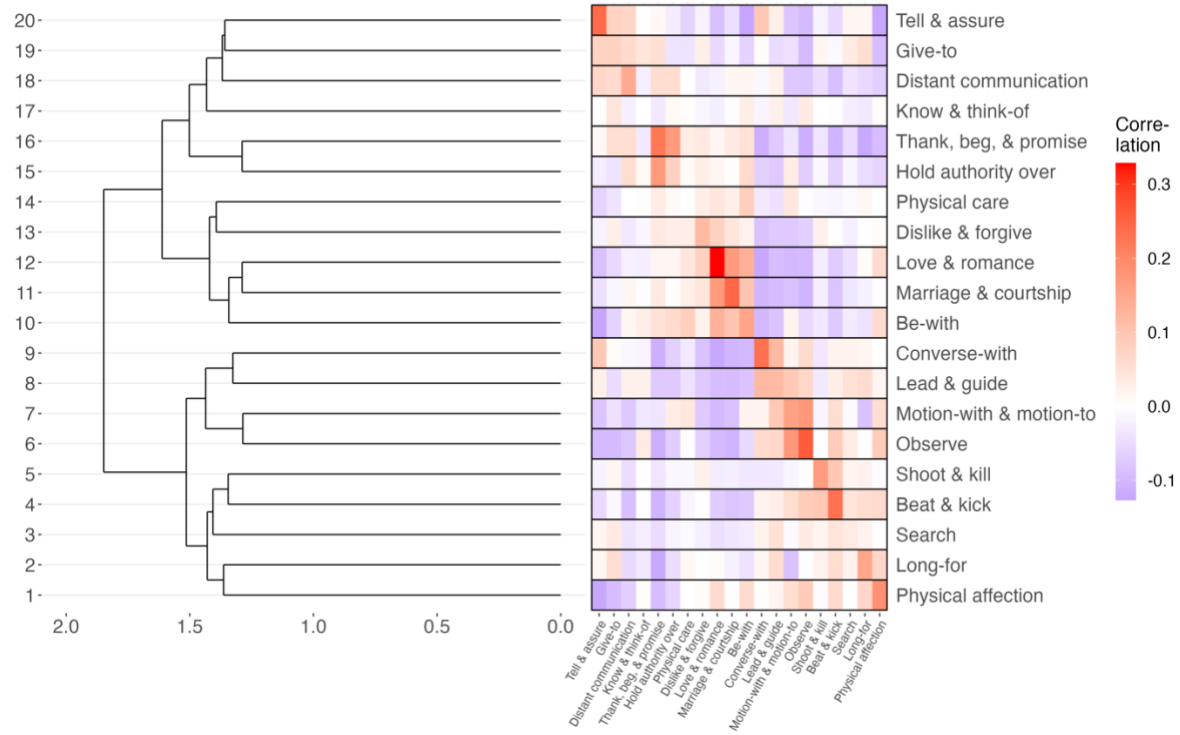

Figure A4. Association of relationship frames with alter's action profile (complete tree).

Note: This figure replicates Figure 3 of the main manuscript but shows the complete tree derived from hierarchical clustering.

Appendix C. Spatial mapping of relationship frames and frame prevalence over time

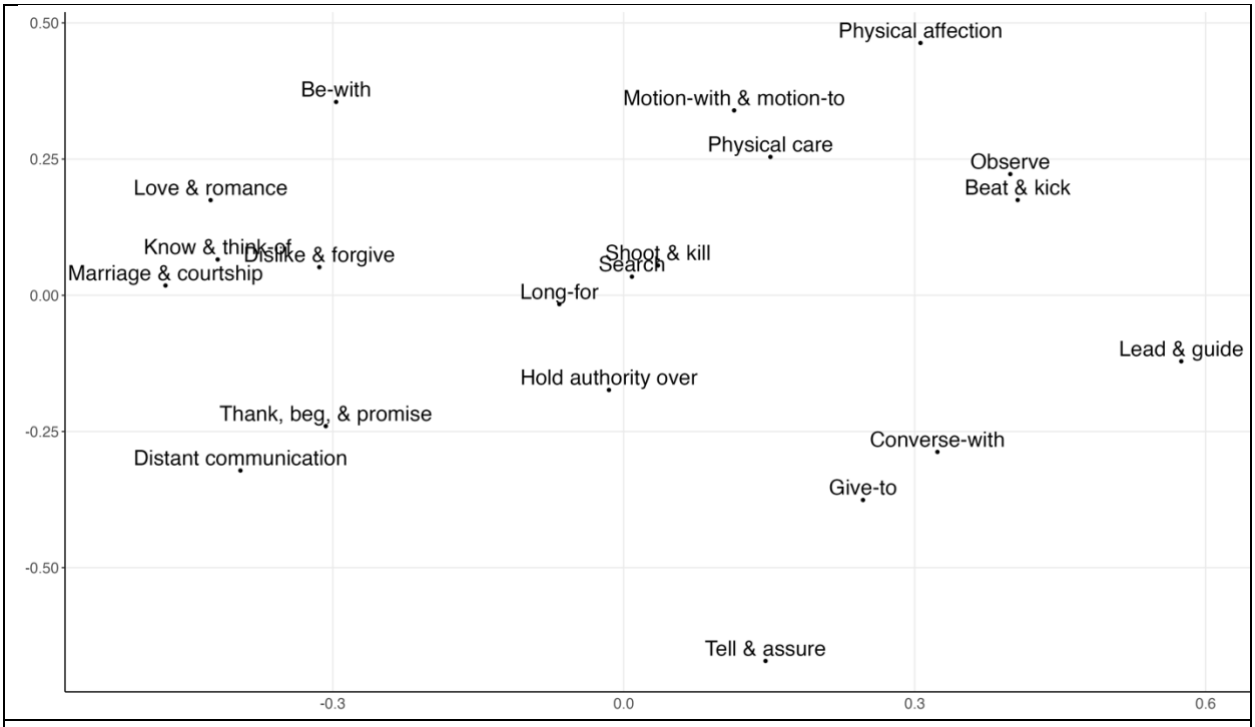

Figure A3. Metric MDS of association of relationship frames within action profiles
Note: The figure contains the results of a metric MDS fit on the association of relationship frames within action profiles. The distance between relationship frames was respectively measured as $d_{ij} = \sqrt{2(1 - c_{ij})}$, where c_{ij} is the correlation between relationship frames i and j. The goodness-of-fit GOF score is at .33, suggesting that two dimensions are not capable of capturing the distances among frames well. This is not too surprising, as the BTM algorithm's objective is to identify distinct clusters.

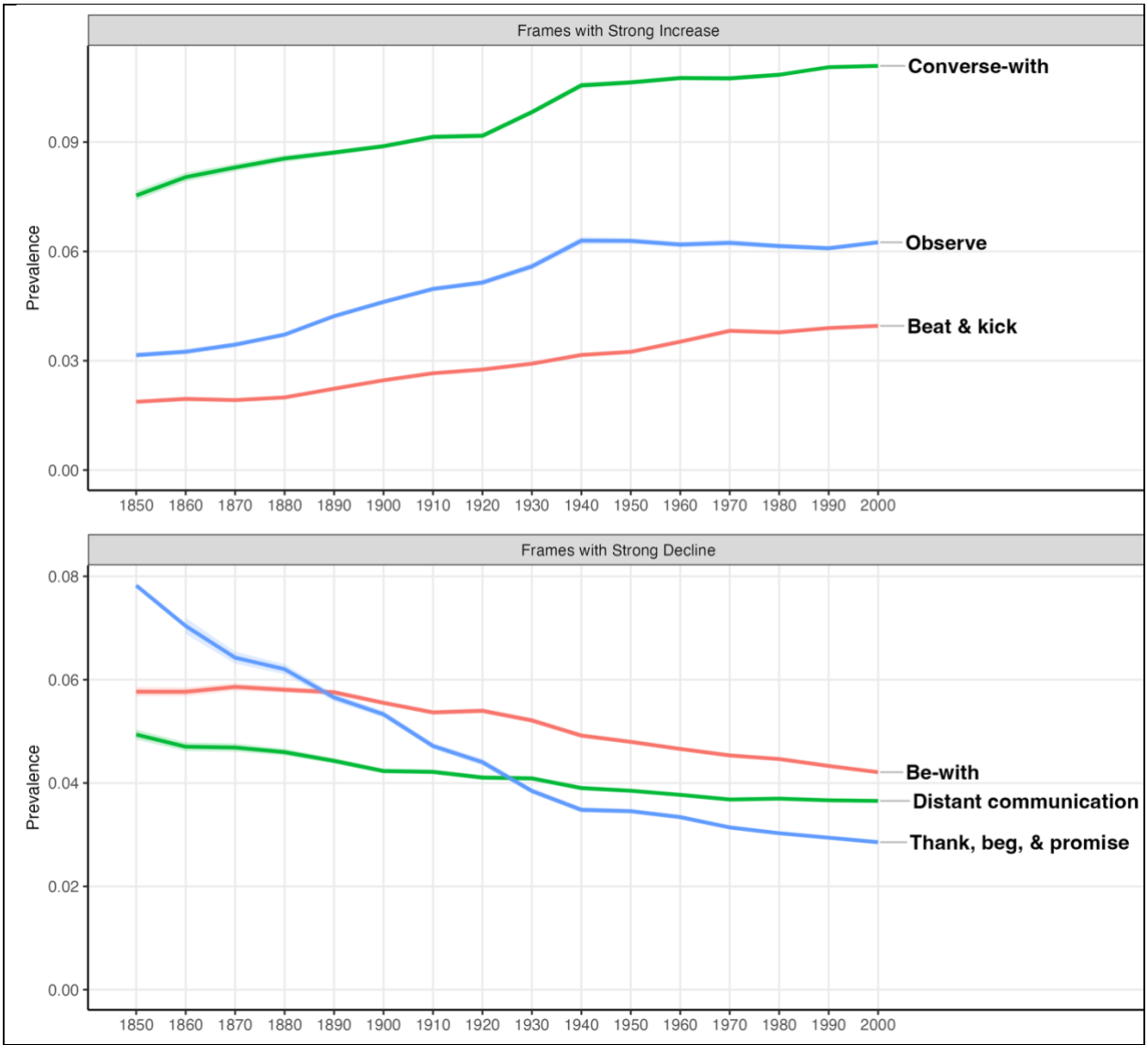

Figure A6. Changes in frame prevalence

Note: The figure shows the trends in the prevalence of the different frames. Ties were averaged within books. The figure includes parametric 95% confidence intervals, using the books as unit of analysis. Averages were computed for each decade. To group the trends, I fit linear models on the 16 averages for each frame, using the decades as independent variables. I then group the frames by effect size, identifying those with the strongest linear decline and increase. The colors are merely for distinguishing the lines, but lines of the same color in different panels have no relation to each other.

Appendix D. Frame ambiguity distribution and regression model

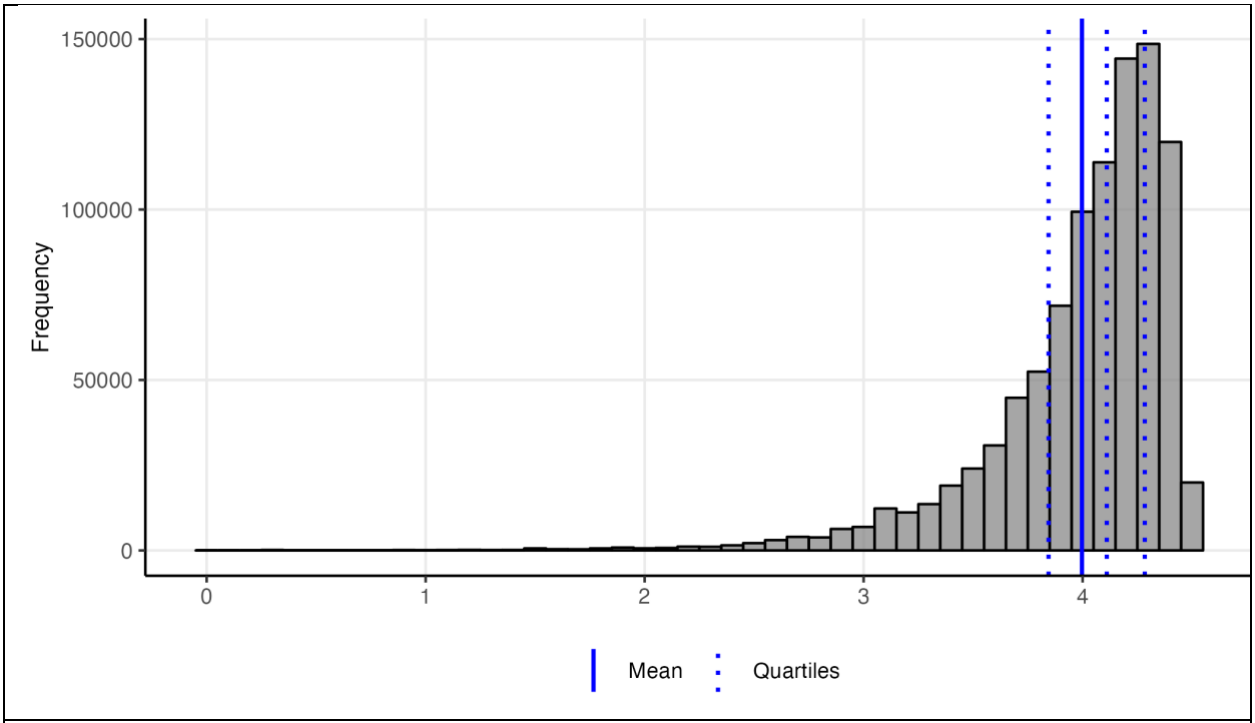

Figure A5. Distribution of action frame ambiguity (entropy)
Note: This figure shows the distribution of the measure for frame ambiguity used in the final section of the paper. The lines indicate the mean (4.0; solid) and quartiles (3.8, 4.1, and 4.3; dotted), respectively.

	Ego's action profile ambiguity		
	(1)	(2)	(3)
Actions in ego's action profile	0.025***	0.025***	0.024***
	-0.0003	-0.0003	-0.0003
Actions in alter's action profile	-0.001***	-0.001***	-0.001***
	-0.0002	-0.0002	-0.0002
Ego's betweenness centrality	0.084***	0.089***	0.088***
	-0.002	-0.002	-0.002
Ego's betweenness centrality	0.063***	0.069***	0.067***
	-0.002	-0.002	-0.002
Female ego, male alter		-0.036***	-0.026***
		-0.004	-0.004
Male ego, female alter		0.029***	0.029***
		-0.004	-0.004
Male ego, male alter		-0.106***	-0.099***
		-0.004	-0.004
Observe			0.030***
			-0.003
Know & think-of			-0.029***
			-0.003
Dislike & forgive			-0.003
			-0.005
Give-to			-0.041***
			-0.004
Marriage & courtship			0.005
			-0.005
Distant communication			0.026***
			-0.005
Love & romance			-0.036***
			-0.004
Beat & kick			-0.080***
			-0.006
Hold authority over			-0.109***
			-0.009
Lead guide			0.027***
			-0.004
Physical affection			-0.021***
			-0.004
Physical care			-0.047***
			-0.005
Shoot & kill			-0.016*
			-0.007
Search			-0.009
			-0.007
Long-for			-0.003
			-0.004
Motion-with & motion-to			-0.050***
			-0.004
Thank, beg, & promise			0.013*
			-0.005
Be-with			0.003

			-0.004
Constant	-0.579***	-0.538***	-0.443***
	-0.018	-0.018	-0.022
Decade fixed effects	Yes	Yes	Yes
Pseudo R2 (marginal)	0.11	0.12	0.12
Pseudo R2 (conditional)	0.27	0.27	0.27
Number of authors	34166	34166	34166
Number of books	72741	72741	72741
Number of male characters	0	0	0
Number of female characters	0	0	0
N	712,056	712,056	712,056

Table A1. Regression on action profile ambiguity

Note: *p < .05; **p < .01; ***p < .001

Appendix E. Qualitative analysis of the relationship between Gerda and John

Dir.	Action	Pos.	Sentence	Context	E.	Comment
G → J	anger	7.4 (3)	What irritated him in Henrietta (no, that was the wrong word — it was anger, not irritation, that she inspired)—what angered him there was Henrietta's unswerving rectitude where he was concerned.	After that summer at San Miguel ...) Curious, when you came to think of it, that the very qualities that irritated him in Gerda were the qualities he wanted so badly to find in Henrietta. What irritated him in Henrietta (no, that was the wrong word — it was anger, not irritation, that she inspired)—what angered him there was Henrietta's unswerving rectitude where he was concerned. It was so at variance to her attitude to the world in general.	✗	This is a coreference error, the angering is attributed to Henrietta, not Gerda. Henrietta is John's lover. The context shows that both identities (Gerde and Henrietta) were discussed in close proximity in the previous sentence, which is likely the cause of this error.
J → G	say-to	7.4 (3)	He had said to her once: "I think you are the greatest liar I know."	It was so at variance to her attitude to the world in general. He had said to her once: "I think you are the greatest liar I know." "Perhaps."	✗	This error carried forward to the next sentence. Again, a coreference error.
J → G	ave-with	10.7 (4)	You can' ave yer fun with me.	But there — I enjoyed the fun. You can' ave yer fun with me. I can stand it."	✗	This is another coreference error. The sentence is uttered by Mrs Crabtree, an elderly patient of John that he is treating for Ridgeway's Disease.
G → J	take	11.9 (4)	Gerda took him in deadly earnest.	He wasn't used to it. Gerda took him in deadly earnest. And Veronica had never thought about anything but herself.	✓	This is John reflecting on his relationships to various women in his life. About his relationship with Gerda, we learn that she "takes him in earnest," a complex expression the semantics of which are not well captured by the mere "take." Meanwhile, we learn that John "objects-in" aspects of her personality.
J → G	object-in	12 (4)	("Just what you object to in Gerda, in fact," said his private imp, bobbing up again.)	He wanted Henrietta to think only of him, never to let her mind stray away from him. ("Just what you object to in Gerda, in fact," said his private imp, bobbing up again.) The truth of it was that he was completely illogical.	✓	
G → J	say-to	12.3 (4)	Gerda had said to him one day: "Henrietta has asked me to sit for her."	The most furious quarrel he had had with her had arisen over that. Gerda had said to him one day: "Henrietta has asked me to sit for her." "What?"	✓	Here, John is reflecting on a quarrel he had with his lover Henrietta. Henrietta is a sculptor. John recalls how Gerda told him about her interactions with his lover Henrietta. This small statue was flattering to Gerda.
G → J	show	12.5 (4)	Then, about ten days later, Gerda had shown him triumphantly a small plaster statuette.	Something of that kind. Then, about ten days later, Gerda had shown him triumphantly a small plaster statuette. It was a pretty thing — technically skillful like all Henrietta's work.	✓	
J → G	want	12.9 (4)	"So that's what you wanted Gerda for?	He stared and then — suddenly, his neck swelled and he turned on her furiously. "So that's what you wanted Gerda for? How dare you?"	✗	In this moment, John realizes that the actual reason Henrietta used Gerda as a model was for a different sculpture ("The Worshipper"). This sculpture is much less flattering. This upsets John. It is Henrietta who "wanted"

						Gerda, so this is a false coreference.
G → J	look-at	13.6 (4)	But I think — she might be looking at you, John."	She said, and her voice had a queer note in it: "I don't know. But I think — she might be looking at you, John."	✓	The sculpture is described as a "figure offering up worship to an unseen deity". John asked Henrietta where the sculpture of Gerda is looking. Henrietta answers that the sculpture is looking at him and thus worshipping John. This is a key moment for describing the relationship between John and Gerda (she worships him). The full meaning of which is, of course, not fully captured by the phrase "looking at." This marks the end of chapter 4
J → G	choose	14.6 (5)	And he had chosen her, when he might have married somebody far more brilliant.	She had always known, from the very first, that John was brilliant and was going to get to the top of the tree. And he had chosen her, when he might have married somebody far more brilliant. He had not minded her being slow and rather stupid and not very pretty.	✓	In the previous chapter, we learned indirectly about the uneven relationship between Gerda and John — that is, we learned how it is seen by Henrietta. Now, in chapter 5, we are presented with an interior monologue that contains Gerda's thoughts about her marriage with John. These largely confirm the uneven relationship. She is glad about him having chosen" her. He looks after and takes care of her. Essentially, she frames herself as the lucky inferior partner. At the same time, we see verbs indicative of what could both be described as protectiveness or paternalism.
J → G	look-after	14.6 (5)	"I'll look after you," he had said.	He had not minded her being slow and rather stupid and not very pretty. "I'll look after you," he had said. Nicely, rather masterfully.	✓	
J → G	take-of	14.6 (5)	"Don't worry about things, Gerda, I'll take care of you ...."	Nicely, rather masterfully. "Don't worry about things, Gerda, I'll take care of you ...." Just what a man ought to be.	✓	
J → G	choose	14.7 (5)	Wonderful to think John should have chosen her.	Just what a man ought to be. Wonderful to think John should have chosen her. He had said with that sudden, very attractive, half - pleading smile of his: "I like my own way, you know, Gerda."	✓	
G → J	give-to	14.7 (5)	She had always tried to give in to him in everything.	Well, that was all right. She had always tried to give in to him in everything. Even lately when he had been so difficult and nervy — when nothing seemed to please him.	✓	
J → G	give	14.9 (5)	He never would give her anything for them, when surely it would be so easy, being a doctor.	And it did so annoy John when she had headaches. He never would give her anything for them, when surely it would be so easy, being a doctor. Instead he always said: "Don't think about it.	✓	
J → G	interrupt	15.8 (5)	John interrupted her impatiently. (68)	Gerda went on rather incoherently: "I'm so sorry, dear, it's all my fault, but first, you see, I thought you were coming, and then I thought, well, if I did send it back ...." John interrupted her impatiently. "Oh, what does it matter?	✓	The setting here is a lunch at the couple's house. Gerda apologizes for serving him a cold piece of meat, though it was John who came significantly later. Again this shows the uneven power

						relationship between the two. Gerda would likely not dare interrupt John.
J → G	raise-over	16.9 (5)	Christow raised his eyebrows over it and Gerda hurried into apologies.	The pudding was burnt. Christow raised his eyebrows over it and Gerda hurried into apologies. “I’m sorry, dear.	✗	The object here is the pudding, so we have a false coreference. The passage is very similar to the one before and bring home a similar point.
J → G	tease	22.9 (7)	John, she knew, was teasing her.	Gerda gave a little quiet sweet laugh. John, she knew, was teasing her. She stuck to her point.	✓	The context here is that John (being a doctor) earlier made a joke about hating sick people. Now, driving in the car to Lady Angkatell’s country house (The Hollow) for the weekend, Gerda very politely says that he shouldn’t make these kinds of jokes in front of the children. John then pushes back. John is “teasing” Gerda, which underscores the tilt of power in their marriage. It is hardly imaginable that Gerda would ever tease John.
J → G	interrupt	23.2 (7)	It’s such a noble life — and I’m so proud of the way you give your time and energy and never spare yourself —” John Christow interrupted her.	“You see,” Gerda stuck to her point, “I’ve always impressed on the children just what a doctor’s life is — the self - sacrifice, the dedication of oneself to helping pain and suffering — the desire to serve others. It’s such a noble life — and I’m so proud of the way you give your time and energy and never spare yourself —” John Christow interrupted her. “Hasn’t it ever occurred to you that I like doctoring — that it’s a pleasure, not a sacrifice!—Don’t you realize that the damned thing’s interesting!”	✓	These passages still take place in the car ride to The Hollow. We get another set of verbial signifiers for the uneven relationship between John and Gerda. Most interestingly, John again interrupts Gerda, stressing his eagerness to seize control of the conversation, but also his frustration with his wife. The repetition (“If he told her,” “If he were to tell her”) dramatizes this frustration. Gerda sees John as some kind of saintly artisan, and he cannot communicate his professional passion to her. We also learn that he loses his temper with her (because Gerda initially failed to start the cars and then shifted gears inappropriately during the drive). The meaning of the phrase is not entirely captured by the extract, however. I note here that this is a deliberate design choice of the software, for if we were to retain this level of information, the data would be considerably more sparse.
J → G	tell	23.3 (7)	If he told her about Mrs. Crabtree and the Margaret Russell Ward she would only see him as a kind of angelic helper of the Poor with a capital P. “Drowning in treacle,” he said under his breath.	But no, he thought, Gerda would never realize a thing like that! If he told her about Mrs. Crabtree and the Margaret Russell Ward she would only see him as a kind of angelic helper of the Poor with a capital P. “Drowning in treacle,” he said under his breath. “What?”	✓	
G → J	see	23.3 (7)	If he told her about Mrs. Crabtree and the Margaret Russell Ward she would only see him as a	But no, he thought, Gerda would never realize a thing like that! If he told her about Mrs. Crabtree and the Margaret Russell Ward she would only see him as a kind of angelic helper of the Poor with a capital P.	✓	

			kind of angelic helper of the Poor with a capital P. “Drowning in treacle,” he said under his breath.	“Drowning in treacle,” he said under his breath. “What?”		Finally, we also learn that Gerda shows (or wants?) affection by leaning towards him. More broadly, then, the passage again underscores the asymmetry of their marriage.
G → J	lean-towards	23.3 (7)	Gerda leaned towards him.	“What?” Gerda leaned towards him. He shook his head.	✓	
J → G	tell	23.3 (7)	If he were to tell Gerda that he was trying to “find a cure for cancer,” she would respond — she could understand a plain sentimental statement.	He shook his head. If he were to tell Gerda that he was trying to “find a cure for cancer,” she would respond — she could understand a plain sentimental statement. But she would never understand the peculiar fascination of the intricacies of Ridgeway’s Disease — he doubted if he could even make her understand what Ridgeway’s Disease actually was.	✓	
J → G	lose-with	23.8 (7)	No, indeed, she’d much rather drive on for hours and hours, even if John did lose his temper with her!	Not that she wanted to get there. No, indeed, she’d much rather drive on for hours and hours, even if John did lose his temper with her! But now they were driving along Shovel Down — flaming autumn woods all round them.	✓	
J → G	enjoy-without	24.9 (7)	He wouldn’t enjoy it without me.	“He wouldn’t like that. He wouldn’t enjoy it without me. John is so unselfish.	✓	Here, Gerda is speaking to Lady Lucy Angkatell. Note that “enjoy without” is negated in the sentence, something we do not pick up. Nontheless, we learn that John enjoying something without Gerade is a topic of discussion, which tells us about
J → G	explain-to	29.8 (9)	He could never explain to Gerda.)	He could, he felt, at a pinch explain to Henrietta. He could never explain to Gerda.) And he didn’t, definitely he didn’t want to lose anything.	✓	At the country house, John unexpectedly encountered Veronica, an old love. The two took off together to her cottage and eventually had sex (or at least this is strongly implied at the end of chapter 8). This ultimately leads to jealousy and provides the motive for Gerda’s eventual murder. These sentences now stem from John’s inner monologue upon making his way back to the country house where he expects having to face Gerda (“He’d been with Veronica for three hours. She had [...] cut him out of the circle and carried him off as her prize,
J → G	know-about	29.9 (9)	How much did he really know about Gerda?	But would she have? How much did he really know about Gerda? Normally, Gerda would believe white was black if he told her so.	✓	
J → G	deal-with	30.7 (9)	Gerda would have to be dealt with, and he’d better go in and deal with Gerda as soon as possible.	But Gerda, unfortunately, was not an Angkatell. Gerda would have to be dealt with, and he’d better go in and deal with Gerda as soon as possible. Supposing it had been Gerda who had followed him tonight?	✓	
J → G	wake	31 (9)	Sorry I woke you up.	He said easily: “I’ve no idea. Sorry I woke you up. I had to go in with the woman and have a drink.”	✓	

						and he wondered now what on earth everybody had thought about it."). John will have to deal with and explain himself to Gerda.
J → G	feel-towards	31.4 (10)	He felt very kindly towards Gerda this morning.	People like the Angkatells who still managed to have butlers and servants might just as well give them something to do. He felt very kindly towards Gerda this morning. All that nervous irritation that had so fretted him of late seemed to have died down and disappeared.	✓	This is the morning after John (presumably) cheated on his wife. John now "feels very kindly" toward Gerda. One could also say remorseful. Later in the chapter, Veronica comes back to the house to make plans "for [their] future." ("'We've wasted fifteen years'") John rejects her. He wants to stay in his marriage with Gerda. This sets up Veronica as a red herring suspect for the reader. In the later parts of the chapter, there are more passages indicating that John does feel a genuine bond with Gerda and makes resolutions to be better, probably somewhat fueled by guilt ("Poor Gerda, he thought, with her unselfishness and her continual anxiety to please him. He would be kinder in future.").
G → J	stand-over	36.6 (11)	And then I came up the path to the pool and there was John lying there and Gerda standing over him with the revolver.	After all," she appealed to them all, "one doesn't! And then I came up the path to the pool and there was John lying there and Gerda standing over him with the revolver. Henrietta and Edward arrived almost at the same moment — from over there."	✓	This extraction is the first after John has been murdered. Lady Lucy Angkatell describes how she witnessed the crime scene.
G → J	kill	39 (12)	She said: "Why are you all so sure that Gerda killed John?"	Henrietta's voice, icy cold, cut into the placid atmosphere. She said: "Why are you all so sure that Gerda killed John?" There was a moment's pause — and Midge felt a curious shifting in the atmosphere.	✓	This utterance takes place in a conversation between the remaining guests at The Hollow after the murder of John. It is the first extracted verb indicating that there may be an element of violence directed from Gerda towards John. Poirot is not part of the conversation.
G → J	see	39.8 (12)	I — I just saw John —" "Yes, Mrs. Christow?"	I knew Sir Henry and Mr. Angkatell were out shooting. I — I just saw John —" "Yes, Mrs. Christow?" "John — and blood — and a revolver.	✓	This sentence is uttered in a conversation shortly thereafter between Gerda and Inspector Grange. He represents the official police, though the one to solve the case ultimately will be Poirot.

G → J	adore	60.2 (18)	She adored John and he's dead.	"Why should things be made worse than they are for poor Gerda? She adored John and he's dead. She's lost him.	✓	These sentences are uttered in a conversation between Henrietta and Poirot. Henrietta offers her interpretation of the marriage, which aligns with that implied earlier (see "The Worshipper"). Gerda certainly "adored" John. Meanwhile, he "married" her. Note that the latter might seem trivial but it can actually be read as indicative of a power relationship. It was he who "chose" and "married" her, not the other way around, and this is repeated throughout the novel.
J → G	marry	61.7 (18)	He married Gerda, whom you might describe inelegantly as a first - class chump.	His one idea was to marry someone as unlike Veronica as possible. He married Gerda, whom you might describe inelegantly as a first - class chump. That was all very nice and safe, but as anyone could have told him the day came when being married to a chump irritated him.	✓	
G → J	shoot	71.2 (20)	If she did shoot John, she's probably dreadfully sorry about it now.	And if you go and put her in prison and hang her, what on earth is going to happen to the children? If she did shoot John, she's probably dreadfully sorry about it now. It's bad enough for children to have a father who's been murdered — but it will make it infinitely worse for them to have their mother hanged for it.	✓	This is uttered by Lady Angkatell in a conversation with Inspector Grange. Lady Angkatell appears to suspect that Gerda was the murderess, but she also tries to shield her from prosecution by describing her as a sympathetic person (though we don't know whether this is genuine).
J → G	take	94.2 (29)	It was all the same as it had been — long ago, before John came and took her away.	"Oh, Gerda, don't be so slow!" It was all the same as it had been — long ago, before John came and took her away. They all thought her slow and stupid.	✓	This is Gerda having a flash-back, reflecting both on her new life as a widow and her past relationship with John.
J → G	look-after	94.3 (29)	There was nobody to say, as John had said: "I'll look after you."	They all thought her slow and stupid. There was nobody to say, as John had said: "I'll look after you." Her head ached and Gerda thought: "I'll make myself some tea."	✓	
G → J	kill	95.6 (29)	I had to kill him!	Gerda said: "I couldn't bear it! I had to kill him! I had to — you do see that, Henrietta?"	✓	Here Gerda confesses to Henrietta.
G → J	do-without	96.7 (29)	"I don't know what to do without John.	Her gaze, piteous, bewildered, went from one to the other. "I don't know what to do without John. John looked after me.	✓	These are statements by Gerda, uttered in conversation with Henrietta while Poirot is present. They are spoken moments before Gerda drinks the poisoned tea. Gerda is not aware that Poirot has identified her as the murderess. The verbs, once more, play on the themes of protectiveness and paternalism.
J → G	look-after	96.7 (29)	John looked after me.	"I don't know what to do without John. John looked after me. He took care of me.	✓	
J → G	take-of	96.7 (29)	He took care of me.	John looked after me. He took care of me. Now he is gone, everything is gone.	✓	

J → G	protect	97.5 (29)	He was asking me to protect Gerda.	It was all there in that one word. He was asking me to protect Gerda. You see, he loved Gerda.	✗	These passages are all from Chapter 29 toward the very end of the book, after Gerda has accidentally poisoned herself by drinking a cup of tea she had prepared for Henrietta. These sentences are uttered by Henrietta in conversation with Poirot, the detective. She offers an interpretation of the marriage between John and Gerda. She insists that love ran in both directions, but we are also told that John had a possessive attitude towards Gerda ("belong-to").
J → G	love	97.5 (29)	You see, he loved Gerda.	He was asking me to protect Gerda. You see, he loved Gerda. I think he loved Gerda much better than he ever knew he did.	✓	
J → G	love	97.5 (29)	I think he loved Gerda much better than he ever knew he did.	You see, he loved Gerda. I think he loved Gerda much better than he ever knew he did. Better than Veronica Cray.	✓	
G → J	belong-to	97.5 (29)	Gerda belonged to him, and John liked things that belonged to him.	Better than me. Gerda belonged to him, and John liked things that belonged to him. He knew that if anyone could protect Gerda from the consequences of what she'd done, I could.	✓	
G → J	love	98.1 (29)	Gerda loved John terribly, but she didn't want to love him for what he was.	"I think so. Gerda loved John terribly, but she didn't want to love him for what he was. She built up a pedestal for him and attributed every splendid and noble and unselfish characteristic to him.	✓	
G → J	love	98.1 (29)	Gerda loved John terribly, but she didn't want to love him for what he was.	"I think so. Gerda loved John terribly, but she didn't want to love him for what he was. She built up a pedestal for him and attributed every splendid and noble and unselfish characteristic to him.	✓	

Table A3. Verbs and their context for the relationship between Gerda and John in The Hollow.

Note: The table shows all 49 extracted actions relating John and Gerda in the book *The Hollow* by Agatha Christie. The column *Dir.* contains the direction of the action. The column *Pos.* shows the relative position token position of the action in the book in percent (e.g., 50 would mean that the action occurred in the exact middle of the book). The number in brackets indicates the chapter. The column *Sentence* contains the sentence the action was extracted from. *Context* contains the sentence immediately before and after the sentence with the action. *E.* contains a mark for whether the verb was extracted correctly or not. Erroneous extractions were left in for transparency. Finally, *Comment* contains a qualitative interpretation and contextualization of the passage and the extracted verbs.