

Empirical Determination of Basic Heuristics for Narrative Content Planning

Pablo Gervás

Instituto de Tecnología del Conocimiento
Universidad Complutense de Madrid
Ciudad Universitaria, 28040 Madrid, Spain
pgervas@ucm.es

Abstract

Stories are sequential in nature but they are used to package human experience that involves many things happening at the same time, to several people or in several locations. The mechanics of this packaging process constitute an instance of content planning that has not been addressed in sufficient detail in existing NLG work. The present paper reviews a number of traditional stories in the light of the basic concepts of narratology that would be involved in the decisions involved in planning the content for tellings of these stories, proposes a number of basic principles to understand what is happening, and explores a possible way in which these principles may translate to basic heuristics for narrative content planning.

1 Introduction

Stories are a fundamental vehicle used by people to communicate and understand their environment. An interesting point is that, although stories are generally sequential in nature,¹ they are often used to package human experience that is, in its original form, anything but sequential. Complex sets of events involving many characters over many locations – which do not correspond to an ordered sequence of events but rather to a cloud of events that

¹Stories are linear in the sense that there is only a single path through them, by reading the words and sentences in the order they appear in the page. Non-linear narratives – such as branching storylines, choose-your-own-adventure books or hypertext works – exist as cultural artifacts, but they do not share the importance that linear stories have as vehicles of human experience.

may overlap in time and space – are routinely converted into narrative discourse in the form of novels, short stories, or films. The mechanics of this process constitute the basic skill that novelists and film makers exhibit. Although there has been some debate as to whether writers first come up with a world and then tell about it or directly invent the story with the world being implicitly built during reading (Dehn, 1981) a computational model of the task involved in each of these options would be a useful tool. Work on cognitive models of writing has addressed the process of text composition, but these efforts focus on the elementary composition of text structure (Flower and Hayes, 1981) or the creativity affecting the ideas to be included (Sharples, 1999), rather than the story structure. Although these are worthy research topics, and there are a number of additional issues that are indeed significant from a psychological perspective, the mechanics that underlie the transcription processes involved may provide important insights on the composition processes.

Recent efforts in the field of natural language generation have addressed the formalization of the basic representational elements that are involved in these process (Gervás, 2012; Gervás, 2014) but have stopped short of identifying specific heuristics that might be involved in the process.

The present paper reviews a number of traditional stories in the light of the basic concepts of narratology that would be involved in the decisions involved in planning the content for tellings of these stories, proposes a number of basic principles to understand what is happening, and explores a possible way in which these principles may translate to basic heuris-

tics for narrative content planning.

2 Related Work

The outlined programme requires the introduction of basic concepts of narratology and brief discussion of how they have been addressed in related studies.

2.1 Narrative

An important distinction is made in studying narrative between the content of a story as it would have taken place in the real world – or an imagined one – and the way that an author chooses to present it. The content of the story is exhaustive in detail and all of its ingredients is fixed in time and space. The way this content is presented by a given author involves selecting only particular aspects to mention, and telling those aspects in a particular linear order. There are several ways of understanding and referring to this distinction (see (Abbot, 1986) for details), but for simplicity we will refer to the exhaustive content as the *fabula* of the story and to the particular way of telling it as the *discourse* chosen for it.

This distinction forces the consideration of two different reference frames for time: the time in which events happened in the *fabula* – which we will refer to as *story time* –, and the point in the sequence of the discourse in which the corresponding events are mentioned – which we will refer to as *discourse time*. The way in which story time and discourse time differ, and the way in which they relate to one another is traditionally known as *chronology* (Abbot, 1986).

Discourse is linear and *fabula* usually is not. A *fabula* may involve a world where only one action takes place at a time, none of these actions overlaps with the next, and each action is immediately perceptible by an agent that was focusing on the preceding one. But such cases are extremely rare. More complex *fabulae* require discourse to break the telling of events that happen simultaneously into separate segments of discourse (sometimes known as *narrative threads*), where each thread follows a different character as they go through different experiences over the same period of story time. This is known as *focalization* (Genette, 1980). When discourse needs to change focalizer to go back in time

or go to a different location, these changes may need to be explicitly marked as contextualizations of the new thread with respect to the preceding one, in order to help the reader make the correct interpretation (Gervás, 2014).

2.2 Narrative Planning in Natural Language Generation

The mechanics that this paper sets out to clarify correspond to the particular instantiation of the *content planning* task (Reiter and Dale, 2000) for the case when the content to be conveyed is a *fabula* and the text to be generated is a discourse, in the sense described above. With respect to prior work in the field of NLG, the present paper focuses on the task described as *narrative planner* in Callaway’s work on narrative prose generation (Callaway, 2002), which focused on narrative realization rather than narrative planning. It also correlates reasonably well with efforts to generate discourse to describe sports events (Lareau et al., 2011; Bouayad-Agha et al., 2011; Allen et al., 2010) if the events that took place in the corresponding games are considered the *fabula*. The present work addresses the task in more detail by considering aspects such as characters and protagonism. Also related is work on the automatic generation of cinematic visual discourse (Jhala and Young, 2010), which shares the goal of identifying the best linear sequence of restricted views to convey a given content (or *fabula*). Where cinematic visual discourse focuses on restricted views as determined but the part of a scene that can be covered by a camera take, the present paper focuses on restricted views as determined by focalization on a given character – and telling only what that character might have perceived.

A set of elementary data structures to capture some of the concepts of narrative composition as described in section 2.1 has been proposed in (Gervás, 2012; Gervás, 2014). This work introduced concepts of a *fibre* – the restricted view of a given *fabula* as perceived by a given focalizer character –, and the tasks of *heckling* a *fabula* into fibres and *splicing* a selection of those fibres into a single linear discourse.

- 1 Mother pig tells boys to build
- 2 Pig1 builds house of straw
- 3 Pig2 builds house of sticks
- 4 Pig3 builds house of bricks
- 5 Wolf blows house of straw away
- 6 Pig1 runs to house of sticks
- 7 Wolf blows house of sticks away
- 8 Pigs 1 & 2 run to house of bricks
- 9 Wolf fails to blow house and leaves
- 10 Pigs and their mother rejoice

Table 1: Story of the Three Little Pigs

2.3 Narrative Planning in Cognitive Models of Writing

The classic model of the writing task from a cognitive point of view (Flower and Hayes, 1981) focuses on the production of informative documents, with little attention devoted to the particular case of narrative discourse. The work of (Sharples, 1999) addresses writing as a task of creative design, focusing on the interplay between following an initial set of constraints and revising those constraints as a result of reflection on partial results obtained during discourse production. Such a high-level abstract view of the process is clearly relevant for narrative generation, but the particular case of narrative discourse as considered here was not considered.

3 Empirical Study of How Known Stories are Planned

In order to understand how the mechanics of building a discourse for a given fabula operate, we turn to the analysis of two traditional stories. For each one, we try to infer what the fabula for the story might be, and correlate that with the discourse as we have come to know it. From his comparison, we hope to obtain insights on the decisions that need to be taken and the heuristics that may be employed to inform them.

3.1 The Three Little Pigs

The story of the Three Little Pigs (outlined in Table 1) provides an interesting example of how the process of content planning moves from a fabula to a discourse. In this case, the fabula would be a record of the activity of every character from the start of the story to the end. A sketch of this would correspond to the representation given in Figure 1a, which pro-

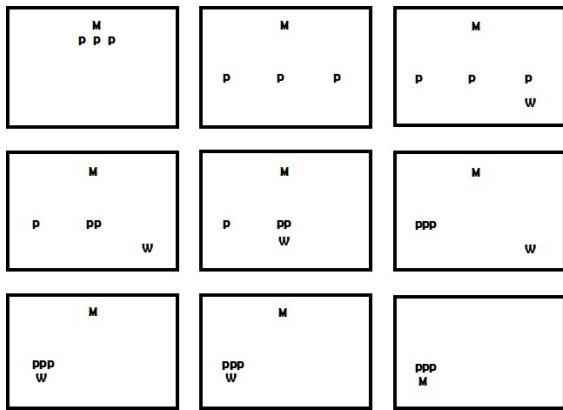
vides a graphical depiction of the fabula. Here you can follow all characters as they move around the story world and interact with one another. This representation is accurate but does not correspond to the version of the story that everybody knows. Because that corresponds to one possible discourse that “tells” this fabula. This discourse is captured in the representation given in Figure 1b. Rather than tell what happens to all characters at each point in time, the discourse focuses on a small subset of the action – that perceived by the characters that are more relevant to the story at that point. Where the actions relevant to the story at a given time point occur too far apart to be perceived by the same characters, the discourse focuses first on one possible location, and then moves back in time to focus on a different location. An example of this occurs in the second time point of the story, where the discourse tells in sequence how each little pig builds a different house, even though all the houses are built over the same time period.

This corresponds to the traditional concept of focalization: at time point 2, the discourse focalises respectively on each of the three little pigs as they build their houses.

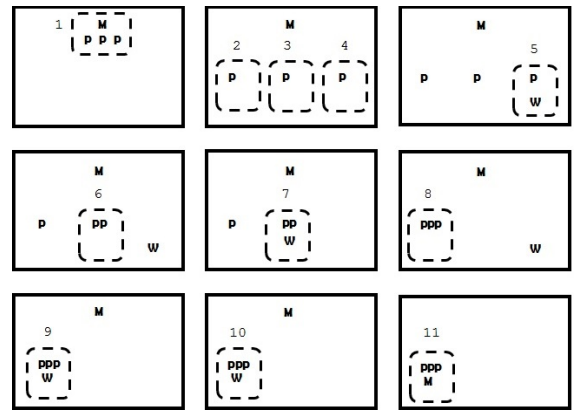
There are a number of additional interesting features to be observed in this story. First, there is a significant portion of the fabula that is not present in the discourse. The activity of the pig’s mother from time point 1 – when she send off her sons into the world – to time point 9 – when she joins them to celebrate their survival –, the activity of little pigs number two and three from they build their own house to the time they are asked to harbour their harassed brothers fleeing from the wolf, or the activity of the wolf whenever he is not threatening the little pigs, are not covered by the discourse.

We have no way of knowing – from the story as it is traditionally told – whether nothing happens to these character over those periods of time. We can only infer that nothing happens to them that is relevant to the story. This is an important point that can later be translated into useful heuristics for content planning.

Two important issues can be pointed out. First, the story is about the three little pigs: other characters such as the mother or the wolf only come into the story as they interact with the main charac-



(a) Fabula. Story time flows left to right and top to bottom.



(b) Discourse (marked out over the fabula). Discourse time marked in Arabic numerals

Figure 1: Fabula and discourse for the story of the Three Little Pigs

1	Mother send LRRH to Granny
2	LRRH sets off through forest
3	Wolf send LRRH on ``short cut``
4	LRRH lingers to pick flowers
5	Wolf reaches Granny first
6	Wolf devours Granny
7	LRRH arrives at Granny's
8	Wolf devours LRRH
9	Hunter arrives at Granny's
10	Hunter kills Wolf and victims emerge

Table 2: Story of the Little Red Riding Hood

ters. The discourse follows one or more of the pigs throughout the story. Second, even the main characters of the story may be ignored for a period of time if nothing relevant to the overall outcome is happening to them over that period.

3.2 Little Red Riding Hood

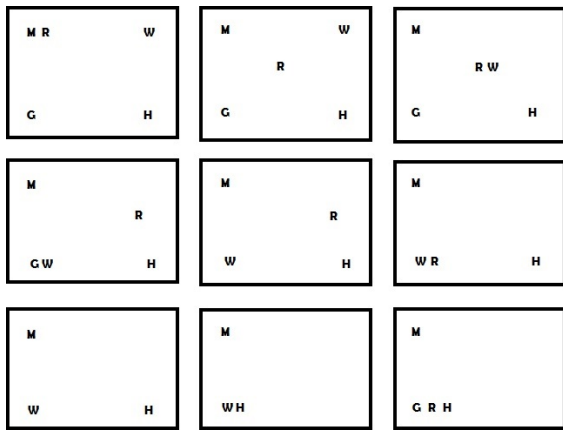
A similar analysis may be carried out for the story of Little Red Riding Hood (outlined in Table 2). Fabula and discourse for this story are shown in Figure 2. In this case, a larger set of different characters is involved, and the protagonist of the story is a single character, Little Red Riding Hood herself. The discourse indeed follows her most of the time – discourse segments 1 to 4, then briefly in 7, and finally in 10. But discourse segments 5 to 6 and 8 to 9 follow the wolf instead. This is a refinement to our previous analysis in that it introduces secondary characters that need to be followed over part of the time for the story of the protagonist to make sense. Here,

what happens to the wolf over the periods when the girl is not present is relevant to the story. This is because what happens to the wolf is partly shared with the girl and partly shared with characters that are closely related to the girl – the way in which wolf replaces the grandmother to then impersonate her in front of the girl. If this part of the fabula is not told, the story as seen from the point of view of the girl would not make sense. So the discourse needs to focalise on him over the periods where these events take place.

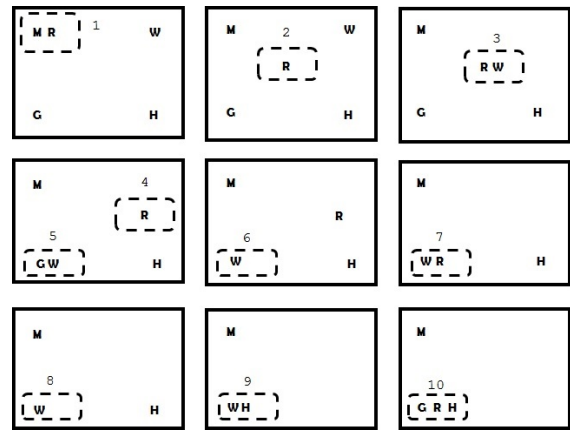
4 Identifying Abstract Principles from the Case Studies

The stories analysed above may be taken as evidence suggesting that a number of basic principles may be at play in the structuring of discourses from fabulas. Although the data set under consideration is clearly insufficient to draw any significant conclusions, a preliminary analysis of it may yield formative insights that can be used to construct baseline implementations of this task of narrative content planning or narrative composition.

An important point to consider is that any such principles would ideally be relevant not only for planning but also for interpreting narrative discourse. The interplay between interpretation and composition – with estimates of what a reader might interpret being used by writers to inform composition – has already been addressed as a plausible model in (Gervás and León, 2016), and tentative im-



(a) Fabula. Story time flows left to right and top to bottom.



(b) Discourse (marked out over the fabula). Discourse time marked in Arabic numerals

Figure 2: Fabula and discourse for the story of Little Red Riding Hood

plementations were presented in (Gervás, 2014).

4.1 Basic Principles of Narrative Economy

The telling of stories is an instance of generic situation of communication involving a speaker, who tells the story, and an audience, who listens to the story. The audience may be made up of one or several listeners, and these may be specific people or a generic public. In all cases, the process as a whole is governed by a number of implicit assumptions that help speakers and listeners to optimise the processing required of them to take part. It would be helpful if we could identify some of these assumptions and formulate them as principles to govern our attempts to model human storytelling abilities. Such principles could be considered as particular instantiations of Gricean principles (Grice, 1975) governing exchanges between two agents.

When composing a story, a synthetic speaker may apply these principles to guide any decisions it needs to make. It would do so under the assumption that a listener (whether synthetic or authentic) would apply similar principles to the process of interpreting the resulting story. For the present purposes, we will focus on the case where there is a single listener. Extension to situations where there are multiple listeners may be considered as future work. Also, we consider principles applicable to a situation where a given fabula is available to the speaker and not available to the listener, and in which the goal for the listener to become aware of this fabula by inter-

preting a discourse based on it composed by speaker. The implicit task for the speaker is to construct the discourse that most economically satisfies this goal. The precise definition of economy may require further investigation, but some baseline definitions can be provided below.

The following basic principles may be postulated.

Principle of Focalized Perception: The listener will best understand a discourse describing a part of the fabula if it is phrased in terms of what a particular character might perceive.

Principle of Faithful Reporting: If the fabula contains an event involving a character that has already been mentioned in the discourse, that event should be mentioned in the discourse.

Principle of Temporal Congruity: The ordering of events as they appear in the discourse should follow as closely as possible the ordering of events in the fabula.

These principles have been stated at a high level of abstraction, but more pragmatic considerations may be derived from them. For instance, the Principle of Focalized Perception leads directly to the presentation of narrative discourse as a sequential combination of narrative threads – or fibres, in the terminology proposed by (Gervás, 2014) – focalized on different characters. For the more traditional concepts of story, the focalizer for the most relevant thread in a given story can be considered the protagonist. This underlies the conventions applied by

both speakers and listeners in a basic storytelling situation. The Principle of Faithful Reporting is applied by speakers to optimise the length of the discourse by making no mention of certain characters when nothing is happening to them over long periods of time. The same principle is applied by listeners, allowing them to assume that if no action has been mentioned for a given character over a given period of time, that character has remained in the same conditions as it was when last mentioned. The Principle of Temporal Congruity is applied when encoding a temporal sequence, by allowing the speaker to simply enumerate the events on the assumption that, unless explicitly mentioned, the order of presentation matches with the order of occurrence. This reduces the need for temporal discourse markers or connectors to situations where divergences from this baseline occur. This generally happens when the discourse switches to a different narrative thread – which may require going back in time to where a different character was abandoned in favour of the focalizer of the thread that has just been reported – or sometimes when fragments from the same thread are presented in non-chronological order in the discourse (as in flashbacks or flashforwards on a given character).

4.2 Basic Heuristics for Content Planning

This reduced set of principles can now be used to produce a corresponding set of heuristics for narrative content planning.

A synthetic speaker faced with the task of constructing a discourse for a given fabula should:

1. identify the character in the fabula most likely to work as a protagonist
2. establish the narrative thread that focalises on this protagonist
3. for any additional characters – other than the protagonist – that appear in the resulting set of threads
 - (a) identify points in the discourse where these characters suffer changes of state due to events that are not covered by the set of thread already included in the discourse

- (b) find the minimal span of narrative thread that would ensure coverage of those events if added to the discourse
- (c) splice this minimal span into the discourse

This heuristic-driven procedure is designed to operationalise the application of the stated principles as a reference baseline. They should lead to discourses that satisfy the principles in an elementary fashion, while requiring for their application no complex sources of knowledge. The relation between the procedure and the described principles is discussed below.

Points 1, 2 and 3.b arise from application of the Principle of Focalised Perception. Point 3.a is driven by the need to satisfy the Principle of Faithful Reporting. Point 3.c would need to take into consideration the Principle of Temporal Congruity either by inserting the additional span of narrative threads at a time of the discourse where the principle is satisfied, or by inserting additional temporal markers in the discourse to indicate where deviations from the expected chronology occur.

4.3 Testing the Application of the Heuristics on the Studied Cases

The acceptability of the heuristic-driven procedure can be tested by checking that applying them to the known fabulae for the stories we have considered in detail does indeed lead to discourses that satisfy the basic principles. If the resulting discourses also match the known discourses for these stories this might be considered as reinforcement for the validity of these heuristics as a baseline.

The case of the Three Little Pigs is very basic if one allows for the three pigs to act collectively as protagonist. Then the story arises naturally from following them through the story, separating into different narrative threads when they split apart, rejoining when they meet again, and ordering the resulting thread spans according to the Principle of Temporal Congruity.

The case of Little Red Riding Hood can be discussed over the fabula and discourse presented in Figure 2. The valid choice for protagonist is the girl. Adding the thread for the girl to the draft discourse would result in a discourse covering discourse points 1, 2, 3, 4, 7 and 10. Application of point 3 of the

heuristics would identify relevant events involving the grandmother, the wolf and the hunter that are not covered yet. Adding the thread of the wolf from fabula time point 4 to fabula time point 5 would cover all the events relevant to the grandmother and some of those by the wolf. Adding the thread for the grandmother would miss out fabula time point 5 – she has already been eaten by the wolf then – so this seems more economical. The events corresponding to discourse points 8 and 9 may be covered by adding the thread for fabula time points 7 and 8 for either the wolf or the hunter. In each case, the choice results in a different result for discourse point 9, one focusing on the wolf and one on the hunter.

In both cases, the principles outlined are satisfied by the resulting discourses. The discourses also, within minimal variation generally match the known presentation of the stories.

5 Discussion

It is clear from the discussion above that there is more to narrative content planning than the elementary principles outlined so far. In the original story of Little Red Riding Hood, the span of thread to explain the presence of the wolf is not inserted into the discourse until it becomes apparent to the girl that she is not facing her grandmother. This seems to violate the principles as outlined. However, this type of operation may be considered as an exploitation of the principles by the speaker, where the explanation is withheld until the last possible moment, to enhance the surprise that it creates when it arrives.

Another important point to note is that the stories considered so far are very simple in nature, in as much as the additional spans that have to be inserted are very short. Further work is required to explore how principles such as those stated here might extend to more complex stories, where threads covering several time points need to be combined together. Specific principles would be required to govern how such longer threads are broken down into smaller fragments to allow focalization to switch back and forth between them over long time periods. This practice is well established, for instance, in TV series following multiple characters.

Regarding the relation with previous work, the principles and the heuristic-driven procedure out-

lined would be applicable in an implemented system for narrative composition such as the one described in (Gervás, 2014). They would rely on similar operations of heckling to identify the set of narrative threads corresponding to the fabula under consideration, and they would provide content specific guidance during the splicing process which is only generically described in the original description of the system.

With respect to efforts to generate discourse to describe sports events (Allen et al., 2010), the present paper addresses the task at a slightly lower level of detail in terms of how the fabula is described. Some of the techniques presented in (Lareau et al., 2011; Bouayad-Agha et al., 2011) may be employed to identify the equivalent of a fabula for a given sports event. In that case, the principles and the procedure outlined here could be applicable to the task of constructing a discourse for such a fabula.

The issue of identifying the correct order in which to present the events of a story has been addressed in detail in (Shimorina, 2016). This work differs from the issues discussed in the present paper in that it considers the input to have already the form of a sequence. In this way, it addresses the problem of ordering events in the discourse for a story more in terms of selecting possible relative orderings for a given already linearised discourse. The present paper tackles the problem from one step further back than this, and considers temporal ordering in the context of the problem of linearization and the selection of appropriate focalizers. With respect to the introduction of markers for temporal ordering, there should be some interaction between the strategic decisions of content planning and the tactical decisions of surface realization. This will be addressed in further work.

6 Conclusions and Further Work

The study of well-known stories in terms of the basic concepts of fabula and discourse highlights the importance of focalization and chronology as tools for content planning. Based on elementary insights extracted from this study, a number of basic principles have been postulated that would apply to the task of constructing an acceptable discourse from a given fabula. These principles have been operationalised

into a heuristic-driven procedure for addressing the task planning the content for a discourse that adequately describes the given fabula. The given procedure results in discourses that satisfy the basic principles and roughly match with the discourses known for the stories that have been considered.

The principles and procedures described in this paper are a tentative initial approach to bring together the terminological and conceptual frameworks of narratology and natural language generation. The paper combines the narratological view of a story – as a narrative with a fabula, a discourse, and focalization and chronology establishing the relation between one and the other – and the NLG view of content that needs to be planned into a discourse. These two views describe the same operation but are very rarely combined into a single view of the problem. By bringing them together, this paper attempts to exploit the synergies that arise. A major contribution is the formulation of the problem in terms that are reasonably specific – events, state changes, characters, protagonism – that allow for broad interpretation in the context of possibly different empirical settings or natural language generation systems.

Three different further efforts beyond the work described here can be foreseen. First, larger number of existing stories should be studied for further insights into how their corresponding fabulae relate to the discourses used to convey them. It may also consider cases where more than one discourse is available for what is nominally the same story. This would allow consideration of explicit choices made by writers in creating the different versions of the discourse. Second, more complex stories, as outlined in section 5, should be considered, with a view to extending or refining the set of principles proposed here. Third, the procedure should be implemented in existing story generation systems to test its applicability over different representations of stories and specific domains.

References

S. Abbot. 1986. *Narrative*. Cambridge University Press.
 N. D. Allen, J. R. Templon, P.S. McNally, L. Birnbaum, and K. Hammond. 2010. Statsmonkey: A data-driven sports narrative writer. In *Computational Models of Narrative: AAAI Fall Symposium 2010*.

N. Bouayad-Agha, G. Casamayor, and L. Wanner. 2011. Content selection from an ontology-based knowledge base for the generation of football summaries. In *Proc. ENLG 2011*, pages 72–81.
 C. B. Callaway. 2002. Narrative prose generation. *Artificial Intelligence*, 139(2):213–252.
 Natalie Dehn. 1981. Story generation after tale-spin. In Ann Drinan, editor, *Proceedings of the 7th international joint conference on artificial intelligence (IJCAI 1981), 24-28 August 1981*, pages 16–18, Menlo Park, CA, USA. AAAI.
 L. Flower and J.R. Hayes. 1981. A cognitive process theory of writing. *College Composition and Communication*, 32(4):365–387.
 G. Genette. 1980. *Narrative discourse : an essay in method*. Cornell University Press.
 P. Gervás and C. León. 2016. Integrating purpose and revision into a computational model of literary generation. In Mirko Degli Espositi, Eduardo G. Altmann, and François Pachet, editors, *Creativity and Universality in Language*. Springer.
 P. Gervás. 2012. From the fleece of fact to narrative yarns: a computational model of narrative composition. In *Proc. Workshop on Computational Models of Narrative 2012*.
 P. Gervás. 2014. Composing narrative discourse for stories of many characters: a case study over a chess game. *Literary and Linguistic Computing*, 29(4), 08/14.
 H. P. Grice. 1975. Logic and conversation. In P. Cole and J. L. Morgan, editors, *Syntax and Semantics: Vol. 3: Speech Acts*, pages 41–58. Academic Press, San Diego, CA.
 A. Jhala and R. M. Young. 2010. Cinematic visual discourse: Representation, generation, and evaluation. *IEEE Trans. on Comp. Int. and AI in Games*, 2(2):69–81.
 F. Lareau, M. Dras, and R. Dale. 2011. Detecting interesting event sequences for sports reporting. In *Proc. ENLG 2011*, pages 200–205.
 E. Reiter and R. Dale. 2000. *Building Natural Language Generation Systems*. Cambridge University Press.
 M. Sharples. 1999. *How We Write: Writing As Creative Design*. Routledge, June.
 A. Shimorina. 2016. Generating Stories from Different Event Orders: A Statistical Approach. Master’s thesis, University of Malta, Malta.

Acknowledgments

This work was partially supported by FP7 projects WHIM (Grant Agreement 611560) and PROSECCO (Grant Agreement 600653).