Free indirect discourse: an insurmountable challenge for literary MT systems?

Kristiina Taivalkoski-Shilov[1]
School of Languages and Translation Studies
Koskenniemenkatu 4
20014, University of Turku
Finland
`kristiina.taivalkoski-shilov@utu.fi`

Abstract

This paper argues that an essential element affecting literary translation – the structure of narrative discourse – has been overlooked in research on literary MT systems so far. After a brief survey of basic concepts of structuralist narratology (Genette 1972), which are necessary for understanding essential aspects of literary translation, a type of reported speech called *free indirect discourse* is taken as an example of the translation problems which successful literary MT systems would have to tackle.

1 Introduction

Over the last few years there has been an increasing number of studies that investigate the possibilities of using literature-specific MT systems in literary translation (see e.g. Lee 2011; Besacier 2014; Toral & Way 2015; Toral & Way 2018). As stimulating as these studies are, most of them do not discuss any narrative aspects of literary texts and therefore overlook an essential dimension of literary translation.

In this paper I argue that developing a successful literary MT system requires knowledge of the narrative structure of literary texts – as well as technological expertise, knowledge on translation workflows and readers' expectations. Being a specialist of (human) literary translation myself, my aim is to explain some basic aspects of narrative texts as well as their challenges in literary translation and that way hopefully feed into ethically responsible research on this topic.

In what follows I first define some key concepts that are necessary for understanding literary translation from a narratological point of view. Then I illustrate challenges that the developers of literary MT systems must address by discussing a particularly thorny question of literary translation: rendering *free indirect discourse* (henceforth FID, for a definition see below) in different languages.

2 Narratological Key Concepts for Literary Translation

The key concepts presented in this section come from classical, structuralist narratology that was designed to account for universal phenomena of narrative discourse regardless of cultural and historical context. In this sense structuralist narratology followed the pattern of structural linguistics that investigated the general rules and conventions of language (Steinby and Mäkikalli 2017, 9). Even though the representatives of classical narratology did not take into account changes that occur in the narrative structure when a text is translated, nor other aspects of translatedness (see e.g. Schiavi 1996; Tahir Gürçağlar, 2002) its key concepts offer a solid ground for observing narrative aspects in literary translation.

Steinby and Mäkikalli (2017, 10) point out that Gérard Genette's theory, presented in his seminal "Discours du récit" (in *Figures III*, 1972) became the essence of structuralist narratology thanks to the clarity and usability of his concepts. They write: "Although several of Genette's concepts, particularly focalization, voice, person, the status of the narrator, and the story-discourse distinction (--), have been the subject of extensive critical discussion, it is his conceptualization – with some additions, such as Wayne Booth's 'implied author' –

that forms the hard core not only of 'classical' narratology but also of more recent applications of narratology in other approaches to literary research." (Steinby and Mäkikalli 2017, 10) Owing to the centrality of Genette's notions and their usefulness in translation studies as well, the basic narratological concepts presented here are taken from his "Discours du récit" .

2.1 Story, Discourse and Narrating

Genette's (1972) theory is based on a fundamental division between three narrative levels which are interdependent, but all characterized by their own temporality (Scheffel et al. 2013, section 2). The first level is that of *story* (*histoire*), by which Genette (1972, 72) means narrative content, in other words "the events of the entire narrative in chronological and causal order prior to any verbalization thereof" (Mani 2013, section 3.1.). Naturally these events may not have had real existence, in which case they are inferred from *discourse* (*récit*) that is Genette's second level. For Genette (1972, 74) discourse is the only tangible level of narrative that can be the object of analysis. Discourse does not necessarily present the events in a chronological order and its time dimension is fixed by the text whereas the story-level time dimension is set in the narrated world (*diegesis*) (Scheffel et al. 2013, sections 3.1.1.–3.1.2.). Discourse is also the level where translation takes place and shifts on this level might have a repercussion on the two other levels. For instance, the fact that the first-person narration of *Robinson Crusoe* was shifted into third-person narration in some of the nineteenth-century German, Swedish and Finnish translations turned Crusoe from the narrator of the novel into a mere character (see Taivalkoski-Shilov 2015, 63). Genette's (1972, 72–73) third level is the narrating act itself (*narration*) and the situation where the *narrating* takes place (for instance Marcel relating his past life in *A la recherche du temps perdu*). The narrating should not be confused with the real-life composition of the fiction.

Some scholars, such as Meister (2005, 2011) have developed computer-based markup tools that tag and analyze temporal expressions in literary texts (Scheffel et al. 2013, section 3.2.3.4.). Such tools could turn out useful if they were integrated in literature-specific CAT tools. However, using them in fully automatic MT systems would yield low-quality translations because temporal expressions can have several functions in a literary text:

for instance, tense variation is a marker of certain forms of reported speech in some languages (e.g. English and French).

2.2 Focalization

By creating the term focalization (*focalisation*) Genette wanted to distinguish two aspects of narrating that, according to him, had been hitherto mixed by several narratologists: narrative voice (who speaks?) and focus of narration (who sees?) (Genette 1972, 203–206). The notion of focalization is a means to answer to the question *whose point of view orients the narrative perspective*? Focalization designates the way narrative information is restricted in relation to the narrator, the characters and other possible entities in the storyworld (Niederhoff 2013, sections 1–2). Genette (1972, 206) divides focalization into three categories. In the case where the narrator knows more than the character(s) and relates this information to his audience (the so-called "omniscient narrator"), the focalization is zero. In the case where the narrator tells as much as the character knows, the focalization is internal. In the third case where the narrator shares less information than the character knows the focalization is external.

Focalization is a central concept for FID even though Genette later stressed that FID (belonging to the domain of who speaks?) and focalization (that answers to the question who sees?) should be distinguished from one another. As Kathy Mezei (1996, 70) points out, "(--) FID is frequently the mode by which a narrator focalizes through a character, appropriating that character's words to make the reader see through his/her eyes."

2.3 Reported Speech and FID

Reported speech or the way in which the discourse or the thoughts of literary characters are textually represented is an inherent part of narrative fiction. Genette (1972, 189–203) calls reported speech *récit de paroles*, which highlights the narrator's role as a mediator. The discourse and the thoughts of literary characters take place in the narrated world (the story-level) and even when characters seem to talk without the narrator's intervention, as

in direct discourse (see below),[2] the narrator only pretends to give voice to the character (Genette 1972, 192).[3]

Reported speech appears in many forms ranging from a mention of a speech act to a direct quote that seems to reproduce also stylistically the character's speech (see Taivalkoski-Shilov 2010, 6–13). Types of reported speech can be located on a scale according to different criteria (see e.g. Genette 1972, 191–194; McHale 1978; Leech and Short 1981). For the purposes of this paper it suffices to distinguish between three basic types of reported discourse:

indirect discourse (e.g. Mrs. Smith **answered that she had not seen him that morning**.)

direct discourse (e.g. Mrs. Smith answered: "**No, I have not seen him this morning.**")

free indirect discourse (e.g. After Watson's question Mrs. Smith looked startled for a moment and then composed herself. **No, she had not seen him that** [or this] **morning**.)

The last type, FID, is a hybrid one. The range of its formal possibilities is extremely large (McHale 1978, 253). It is a combination of the narrator's and character's discourse that can appear in first-person or third-person narratives. Ordinarily it combines features of both indirect discourse (back-shift of tenses in retrospective narration) and direct discourse (deictic adverbs like "here" and "now", exclamation marks etc.). (Taivalkoski-Shilov 2006, 142.)

Genette (1972, 192) points out that one of the characteristic aspects of FID is its ambiguity. This is partly caused by the fact that FID is not dominated by a "higher clause" (McHale 1978, 253) and is not preceded by a reporting verb. That is why the interpretation that readers make of it depends on contextual cues and extra-linguistic phenomena (Tammi 2003, 43; Taivalkoski-Shilov 2006, 142). As Genette (1972, 192) observes, it is not always clear whether FID represents the character's speech or thought. Another ambiguity is between the narrator's and the character's voice; who is speaking, the narrator or the character? Furthermore, is the narrator empathetic or ironic towards the character? FID is sometimes also difficult to distinguish from *non-reporting narration*, which means the narration of other events than the speech of the characters (Taivalkoski-Shilov 2006, 137; Taivalkoski-Shilov 2010, 3).

2.4 FID as a translation problem

FID is a *translation problem* (Nord 1991, 151) that all translators irrespective of their level of competence and of the technical conditions of their work have to solve (Taivalkoski-Shilov 2006, 138). Research on the translation of FID shows that FID tends to shift into non-reporting narration, indirect and direct discourse or into other discourse types (Taivalkoski-Shilov 2006, 138–139). There are several possible explanations for this phenomenon. From the perspective of literary MT systems, the linguistic one is the most relevant. The challenge of translating FID is that its linguistic markers vary in different languages (see e.g. Kuusi 2003). Owing to differences in tense, pronoun, adverb and punctuation systems it tends to diminish or even disappear in translation. In some cases, this is because the indices of FID (for instance, the combination of a past tense verb with a present adverb) that are acceptable in one language are unacceptable or even ungrammatical in another. For example, the temporal systems of English and French are asymmetric (Poncharal 1998, 81–82, 241, 266): English uses the preterite tense (simple past) both for the narrator's discourse and FID, whereas modern French opposes the past used in narration (*le passé simple*) and the imperfect which is the typical tense for FID. Poncharal (1998, 180) concludes that in French there is a larger gap between the levels of story and discourse than in English. (Taivalkoski-Shilov 2006, 139.)

4. Concluding remarks

FID often leads to translation shifts in human translation. However, these shifts are probably more logical and less harmful for the narrative structure of the text than those caused by a MT system that is incapable of taking narrative aspects into account. Professional literary translators are capable of making shifts that cause least loss in translation and they can also compensate for the modifications they have to make to the narrative structure of the text. All this is so far lacking in MT systems.

---

[2] For the "reproductive fallacy" of direct discourse, see e.g. Sternberg 1981 and 1982, Rosier 1999, 237–244, and Taivalkoski-Shilov 2010, 7–11.
[3] For Genette the narrator's control over the character's discourse has its limits. According to him the character's voice substitutes for the narrator's voice in the case of free direct discourse, which he calls *discours immédiat* (Genette 1972, 194).

The problem with AI so far is that machine learning of narrative information requires considerable effort and has not been very successful. As Mani (2013, section 4) writes: "[In computational narratology] *Story understanding* systems (e.g. Wilensky 1978) never got very far, since (i) inferring characters' goals involves a large search space and the inferences may need to be revised during processing and (ii) humans use a great deal of knowledge to interpret even simple stories. Given Forster's exemplifying sentence " The king died and the queen died of grief," a child has no difficulty figuring out why the queen was upset, but imparting a body of such commonsense knowledge to a computer is difficult; (iii) aspects of language that are hard to formalize but that are important for story interpretation, such as humor, irony, and subtle lexical associations, have by and large eluded computational approaches."

References

Besacier, Laurent. 2014. Traduction automatisée d'une œuvre littéraire: Une étude pilote [Automatic translation of a literary work: A pilot study]. *21ème Traitement Automatique du Langage* Naturel, Marseille, 2014. P–T 1: 389–394. URL = http://www.aclweb.org/anthology/F14-2001

Genette, Gérard. 1972. Discours du récit : essai de méthode. *Figures III*. Éditions du Seuil, Paris, 65–282.

Kuusi, Päivi. 2003. Free indirect discourse in the translations of Jane Austen's novels into Finnish and Russian. In: Pekka Tammi and Hannu Tommola (eds.) *Linguistic and Literary Aspects of Free Indirect Discourse from a Typological Perspective*. Publications of the Department of Literature and Arts 6, University of Tampere, Tampere, 25–40.

Lee, Tong King. 2011. The death of the translator in machine translation: A bilingual poetry project. *Target: International Journal of Translation Studies*, 23(1): 92–112.

Leech, Geoffrey N. and Michael H. Short. 1981. *Style in Fiction*: *A Linguistic Introduction to English Fictional Prose*. Longman, London.

Mani, Inderjeet. 2013. Computational Narratology. Paragraph 9. In: Hühn, Peter et al. (eds.): *the living handbook of narratology*. Hamburg: Hamburg University. URL = http://www.lhn.uni-hamburg.de/article/computational-narratology

McHale, Brian. 1978. Free Indirect Discourse: a Survey of Recent Accounts. *Poetics and Theory of Literature* (3): 249–287.

Meister, Jan Christoph. 2005. Tagging Time in Prolog. The Temporality Effect Project. *Literary and Linguistic Computing* 20: 107–24.

Meister, Jan Christoph. 2011. The Temporality Effect: Towards a Process Model of Narrative Time Construction. In: Meister J. Ch. and W. Schernus (eds.). *Time. From Concept to Narrative Construct: A Reader*. de Gruyter, Berlin, 171–216.

Mezei, Kathy. 1996. Who Is Speaking Here? Free Indirect Discourse, Gender, and Authority in *Emma*, *Howards End* and *Mrs. Dalloway*. In: Kathy Mezei (ed.) *Ambiguous Discourse: Feminist Narratology and British Women Writers*. The University of North Carolina Press, 66 –92.

Niederhoff, Burkhard. 2013. Focalization, Paragraph 1. In: Hühn, Peter et al. (eds.): *the living handbook of narratology*. Hamburg: Hamburg University. URL = http://www.lhn.uni-hamburg.de/article/focalization

Nord, Christiane. 1991. *Text Analysis in Translation. Theory, Methodology, and Didactic Application of a Model for Translation-Oriented Text Analysis*. Rodopi, Amsterdam.

Poncharal, Bruno. 1998. *La représentation de paroles au discours indirect libre en anglais et en français*. Doctoral dissertation, University of Paris VII.

Rosier, Laurence. 1999. *Le discours rapporté: Histoire, théories, pratiques*. Duculot, Paris.

Scheffel, Michael, Antonius Weixler and Lukas Werner. 2013. Time. Paragraph 4. In: Hühn, Peter et al. (eds.): *the living handbook of narratology*. Hamburg: Hamburg University. URL = http://www.lhn.uni-hamburg.de/article/time

Schiavi, Giuliana. 1996. There is always a Teller in a Tale. *Target: International Journal of Translation Studies,* 8(1): 1–21.

Steinby, Liisa, and Aino Mäkikalli. 2017. Introduction: The Place of Narratology in the Historical Study of Eighteenth-Century Literature. In: Steinby Liisa and Aino Mäkikalli (eds.) *Narrative Concepts in the Study of Eighteenth-Century Literature*. Amsterdam University Press, Amsterdam, 7–37.

Sternberg, Meir. 1981. Polylingualism as Reality and Translation as Mimesis. *Poetics Today* 2(4): 221–239.

Sternberg, Meir. 1982. Proteus in Quotation-Land. Mimesis and the Forms of Reported Discourse. *Poetics Today* 3(2): 107–156.

Tahir Gürçağlar, Şehnaz. 2002. What Texts Don't Tell: The Uses of Paratexts in Translation Research. In: Hermans, Theo (ed.) *Crosscultural Transgressions*.

Research Models in Translation Studies II: Historical and Ideological Issues. St. Jerome Publishing, Manchester, 44–60.

Taivalkoski-Shilov, Kristiina. 2006. FID and Translational Progress: Comparing 18th-century and Recent Versions of Henry Fielding's Novels in French. In: Pekka Tammi and Hannu Tommola (eds.) *FREE language INDIRECT translation DISCOURSE narratology*: Linguistic, Translatological and Literary-Theoretical Encounters. Tampere university press, Tampere, 135−156.

Taivalkoski-Shilov, Kristiina. 2010 [1999] *When two become one*: Reported Discourse Viewed through a Translatological Perspective. In: ed. by Omid Azadibougar (ed.) *Translation Effects. Selected Papers of the CETRA Research Seminar in Translation Studies* 2009. URL = https://www.arts.ku-leuven.be/cetra/papers/files/kristiina-taivalkoski-shilov-when-two-become-one.pdf

Taivalkoski-Shilov, Kristiina. 2015. Friday in Finnish: A Character's and (Re)translators' Voices in Six Finnish Retranslations of Daniel Defoe's *Robinson Crusoe. Target: International Journal of Translation Studies* 27(1): 58–74.

Tammi, Pekka. 2003. Risky Business: Probing the Borderlines of FID. Nabokov's *An affair of honor (Podlec)* as a test case. In: Pekka Tammi and Hannu Tommola (eds.) *Linguistic and Literary Aspects of Free Indirect Discourse from a Typological Perspective*. Publications of the Department of Literature and Arts 6, University of Tampere, Tampere, 41–54.

Toral, Antonio, and Andy Way. 2015. Machine-assisted translation of literary text: A case study. *Translation Spaces*, 4(2): 240–267.

Toral, Antonio, and Andy Way. 2018. What level of quality can neural machine translation attain on literary text? Chapter for the forthcoming book *Translation Quality Assessment: From Principles to Practice*. New York: Springer. Retrieved from: https://arxiv.org/abs/1801.04962v1

Wilensky, Robert W. 1978. Understanding Goal-based Stories. *Yale University Computer Science Research Report*.