Temporal Coherence in Discourse: Theory and Application for Machine Translation

Cristina Grisot

Abstract Temporal coherence in discourse is provided through several temporal cohesive ties, such as tense, aspect and discourse connectives. In the relevance theoretic framework and more specifically in the Geneva school of pragmatics, these cohesive ties are considered as encoding procedural information important for guiding the hearer towards the intended interpretation of the discourse. Jacques Moeschler and his team studied temporal cohesive ties and proposed original theoretical models that have been validated with human and automatic annotation experiments, as well as in language acquisition studies (Zufferey and Popescu-Belis, this volume). In this paper, I show that Jacques Moeschler’s model for inferring temporal discourse relations and his description of tenses expressing past time in French is cross-linguistically valid and can be modelled for improving the results of statistical machine translation systems.

Keywords Verbal tenses · Discourse coherence · Natural language processing · Machine translation · Cross-linguistic · Relevance theory · Cohesive ties

The ideas presented in this paper are based on a series of articles published with Bruno Cartoni, Thomas Meyer, Andrei Popescu-Belis, Michele Costagliola and Jacques Moeschler, with whom I collaborated on two research projects. I am very grateful to Jacques Moeschler for his guidance and resourceful discussions since the beginning of my research. A theoretical model of temporal reference in tensed languages based on empirical work (corpus-based and experiments with native speakers) is proposed in my Ph.D. dissertation (Grisot 2015), in which I make the proposal that temporal coherence in discourse is triggered by the hearer’s need to acquire temporal coherence at the cognitive level. In order to attain this purpose, he treats information coming from several sources (tense, grammatical aspect, lexical aspect, temporal adverbials and temporal connectives) and their rich interrelations in a coherent manner.

The ideas presented in this paper are based on a series of articles published with Bruno Cartoni, Thomas Meyer, Andrei Popescu-Belis, Michele Costagliola and Jacques Moeschler, with whom I collaborated on two research projects. I am very grateful to Jacques Moeschler for his guidance and resourceful discussions since the beginning of my research. A theoretical model of temporal reference in tensed languages based on empirical work (corpus-based and experiments with native speakers) is proposed in my Ph.D. dissertation (Grisot 2015), in which I make the proposal that temporal coherence in discourse is triggered by the hearer’s need to acquire temporal coherence at the cognitive level. In order to attain this purpose, he treats information coming from several sources (tense, grammatical aspect, lexical aspect, temporal adverbials and temporal connectives) and their rich interrelations in a coherent manner.

C. Grisot (✉)
Département de Linguistique, Université de Genève, Geneva, Switzerland
e-mail: Cristina.Grisot@unige.ch

© Springer International Publishing AG 2017
J. Blochowiak et al. (eds.), Formal Models in the Study of Language,
DOI 10.1007/978-3-319-48832-5_19
1 Introduction

Natural Language Processing (NLP), Machine Translation (MT), Semantics and Pragmatics are fields that have had an increasing interest in linguistic phenomena such as discourse connectives, pronoun anaphora and verb tenses due to their important role for discourse coherence. The essential feature that makes a piece of a text a discourse is the coherent succession of sentences forming a whole and referring to the same entities (nominal or eventualities). However, it seems that this condition is not sufficient, as shown in example (1) (Hobbs 1979:67), which is not coherent even if “he” can refer to “John”. Hobbs argues that there is an expectation of coherence, which is deeper than the notion of a discourse just being “about” some set of entities. Sentences in (2) and (3) are coherent because the hearer infers a causal relation in (2) and a temporal relation in (3).

(1) John took a train from Paris to Istanbul. He likes spinach.

(2) John took a train from Paris to Istanbul. He hates airplanes.

(3) John took a train from Paris to Istanbul. He went by boat from there to Cyprus.

The Geneva School of Linguistics, and Jacques Moeschler in particular, investigated linguistic expressions on which discourse coherence depends, such as French connectives and verb tenses. Zufferey and Popescu-Belis (this volume) discuss how Moeschler’s classification and description of the role of French discourse connectives and of discourse relations have been validated empirically in language acquisition studies, NLP and MT.

In this paper, I will discuss Moeschler’s proposal for inferring temporal and causal discourse relations and show that his model is cross-linguistically valid and it can be used for improving the results in terms of coherence of statistical machine translation systems (SMT). I will concentrate particularly on verb tenses in English (EN), French (FR), Italian (IT) and Romanian (RO) and their usage in multilingual translation corpora. I review theoretical accounts of temporal coherence in discourse (classical and pragmatic descriptions) in Sect. 2. Section 3 is dedicated to the link between verb tenses expressing past time and temporal coherence in discourse and to pragmatic features proposed for empirical testing. Section 4 provides

---

1The COMTIS Project (Improving the Coherence of Machine Translation Output by Modeling Intersentential Relations; project n° CRSII2_127510, March 2010–July 2013) and the MODERN Project (Modeling discourse entities and relations for coherent machine translation; project n° CRSII2_147653, August 2013–August 2016) belong to the Sinergia interdisciplinary program funded by the Swiss National Science Foundation.

2At the beginning of the eighties, the label “Geneva School” was given to a series of publications on discourse and conversation that applied basic principles of syntactic analysis to the domain of discourse (Roulet et al. 1985; Moeschler 1985). In the beginning of the nineties, two different directions could be identified in the Geneva School: (i) a general discourse-oriented framework of language based on the modular hypothesis (Roulet 1997) and (ii) a radical pragmatic perspective on discourse sequencing and discourse interpretation (Moeschler 1993, 1996) (see detailed presentation in Moeschler 2001).
the results of corpus analysis, human and automatic annotation of data, as well as the improvement in the translation of EN Simple Past (SP) into FR by automatic systems. Finally, Sect. 5 concludes this paper.

2 Temporal Coherence in Discourse: Theoretical Accounts

2.1 Discourse Relations

In theoretical linguistics, studies have aimed at describing the factors that contribute to discourse coherence and categorizing the different types of coherence relations that connect clauses and sentences. Halliday and Hasan (1976) proposed the terms *cohesive ties* and *cohesion* for the linguistic devices used to build coherence between sentences. A number of theories made use of relations in explaining coherence (e.g. Longacre 1983; Hobbs 1985; Grosz and Sidner 1986; Mann and Thompson 1988). Relations that link clauses are known as “rhetorical predicates” (Grimes 1975), “conjunctive relations” (Halliday and Hassan, 1976) or “intentions” (Grosz and Sidner 1986). The term of “coherence relations” is due to Hobbs (1979).

From the theoretical linguistics perspective, the taxonomy of discourse relations consists of several types, the most well known are those proposed by Halliday and Hassan: additive, temporal, causal and adversative (contrast). Kehler (2004: 244) points out that “an explanatory theory of coherence requires a set of externally driven principles to motivate and ultimately constrain the relation set.” This is the direction taken by Sanders and colleagues (Sanders et al. 1992; Sanders and Noordman 2000; Sanders 1997, 2005) by proposing a theory in which psychological plausibility is the primary motivating factor. In their view, discourse relations are composed of more fine-grained and more primitive features than in the previous taxonomies, such as *basic operation* (causal and additive meanings), *order of segments* (basic or non-basic), *polarity* (positive or negative) and *source of coherence* (semantic and pragmatic). Thus, Sanders et al. proposed a principled and explanatory theory of coherence, which leaves open the possibility to add other factors that interact with the listed ones.

Following Hobbs (1979), Sanders and colleagues (1992) state that coherence relations account for coherence in the cognitive representation of a discourse and they see coherence relations as *cognitive entities* (Mann and Thompson 1986; Sanders et al. 1992). Sanders (2005) suggested that when language users process a discourse, they connect discourse segments by inferring coherence relation on the basis of a very limited set of cognitive principles, such as *causality* and *subjectivity*. At a more general level, he aimed at shedding light on human cognition by investigating the mechanisms underlying discourse coherence. According to Sanders, the causality and subjectivity cognitive principles account for the use of linguistic expressions (connectives and other lexical items), discourse coherence
through discourse relations, language acquisition and discourse processing. Sanders et al. (1992) argued that the essential property of cognitive discourse relations is that they establish coherence in the cognitive representation language users have or make of a discourse.

In a different framework, pragmatic theories (Grice 1989, neo3 and post4 Gricean trends, Geneva School) adopted the idea that human communication in general, and discourse coherence in particular, are inferential processes driven by the desire to express and recognize intentions. In this paper, I am interested in temporal coherence in discourse and its cohesive ties. A question that was asked in the literature is how are temporal relations in discourse inferred. Answers have been proposed by several trends: Discourse Representation Theory (DRT: Kamp and Reyle 1993) and Segmented Discourse Representation Theory (SDRT: Asher 1993; Lascarides and Asher 1993; Asher and Lascarides 2003), Relevance Theory (RT: Wilson and Sperber 1998) and Geneva School (Reboul and Moeschler 1998; Moeschler 2000, 2002; 2005). I will discuss these approaches in what follows.

Since SDRT, it is generally accepted that discourse relations are inferred on the basis of discourse, linguistic and world types of knowledge. Let’s consider two temporal discourse relations: Narration and Explanation (Asher 1993). The Narration discourse relation is characterized by a forward temporal inference, i.e. time advances while Explanation is characterized by a backward temporal inference, i.e. time goes backward. In what follows, I will briefly describe these types of knowledge as they are approached in discourse semantics.

Firstly, discourse knowledge is provided by discourse type, for example, narrative or non-narrative discourses. Labov and Waletzky (1967) argued that Narration is highly preferred in narrative discourses whereas Explanation is preferred in non-narrative discourses.

Secondly, linguistic knowledge is provided by temporal connectives and verb tenses. Kamp and Rohrer (1983) argued that French verb tenses expressing past time encode time direction necessary for inferring discourse relations. For example, the Passé Simple (PS) encodes a forward temporal inference, the Plus-que-parfait encodes a backward inference and the Imparfait (IMP) encodes an inclusive temporal inference. Unfortunately, this idea has numerous counterexamples discussed by Kamp and Rohrer (1983) themselves as in (4), Moeschler (2000), Saussure (1997, 2000; Tahara 2000 for the PS; Saussure and Sthioul 1999, 2005 for the IMP) among others.

(4) Bianca chanta le récitatif et Igor l’accompagna au piano.

‘Bianca sung the recitative and Igor accompanied her on the piano.’

Thirdly, world knowledge is stronger and has priority over linguistic knowledge for determining the inferred temporal discourse relation, as in (5). According to Kamp and Rohrer, the PS encodes a forward inference, which should hold in both examples (a) and (b). But in (b), there is a backward inference based on the causal relation that comes from world knowledge.

(5) (a) Max poussaJean. Il tomba.
    ‘Max pushed John. He fell’
(b) Jean tomba. Max le poussa.
    ‘John fell. Max pushed him.’

The weakness of this approach is circularity: discourse type is defined based on the usage of the appropriate verb tenses corresponding to the intended temporal inference (forward or backward) and discourse relation is inferred based on the discourse type. Moeschler (2005) points out that the most appropriate explanation for establishing temporal inferences in discourse is a pragmatic one. Specifically, we need models that explain how linguistic and contextual information are combined. The following section is dedicated to pragmatic models of temporal relations.

### 2.2 Pragmatic Models

One of the first pragmatic propositions for explaining temporal inferences in discourse is based on one of Grice’s (1967) maxims “Be orderly”. In the gricean framework, temporal inferences are conversational implicatures triggered by the respect or lack of respect of conversational maxims. In other words, these inferences correspond to beliefs attributed to the speaker who is expected to have obeyed the Cooperative Principles and the maxim of order. Gricean conversational implicatures have no contribution to the truth conditions of utterances. Cohen (1971) was the first to question the treatment of temporal inferences as conversational implicatures by showing their link to connectives such as and, because and but. Furthermore, relevance theorists (Sperber and Wilson 1986/1995; Carston 1988, 1993; Wilson and Sperber 1998) treat temporal inferences as “pragmatically determined aspects of what is said” (Wilson and Sperber 1998: 172). In other words, temporal inferences are part of the explicature of an utterance, and they provide the same status to causal relations. Moeschler (2000) argued that causality plays an important role for temporal coherence in discourse and he proposed a theoretical model for temporality that includes this parameter. As it will be shown in Sect. 4.2, Moeschler’s theoretical model was validated experimentally in offline experiments with linguistic judgement tasks.

Moeschler proposed several arguments in favour of treating temporal relations as part of explicatures. I will briefly recall them below.
Firstly, the temporal interpretation corresponds to a pragmatic enrichment of the propositional form of the sentence and they contribute to its truth conditions. In example (6) from Wilson and Sperber (1998: 171), the disjunction is not redundant because each disjunct brings a genuine contribution to the truth-conditions of the utterance. This is based on the assumption that the events presented in each disjunct happened in a different order.

(6) It’s always the same at parties: either I get drunk and no-one will talk to me or no-one will talk to me and I get drunk.

Secondly, relevance theorists’ explanation focuses on processing efforts rather than on cognitive effects. Example (5)(a) produces two interpretations (either forward temporal inference or backward causal inference) and neither syntactic nor semantic structures indicate how the sentence should be interpreted. Moeschler argued that the interpretation is consistent with the cognitive relevance principle. This means that a temporal or a causal interpretation will be chosen depending on which manifest facts are more accessible to the hearer and based on the mutual cognitive environment.

Thirdly, forward temporal inference (called temporal sequencing) and backward causal inference (called reverse-causal interpretation) are not the only possible relations among eventualities. There are two other possible relations, namely simultaneity as in (7) and indeterminacy as in (8).

(7) Bill smiled. He smiled sadly.

‘Bill souriait. Il souriait tristement.’

(8) That night, our hero consumed half a bottle of whisky and wrote a letter to Lady Anne.

‘Cette nuit-là, notre héros but la moitié d’une bouteille de whisky et écrivit une lettre à Lady Ann.’

The relations of simultaneity and indeterminacy are defined as follows:

(A) E1 covers (partially) e2 is a part of the eventuality denoted by e1 is included in the temporal interval defining e2
(B) The relation between e1 and e2 is undetermined if the determining the relation is not necessary for understating e1 and e2 or if determining the relation is not possible.

Fourthly, temporal sequencing (i.e. forward temporal inferences) does not seem to be central for temporal coherence in discourse. Causality plays an important role therefore the question concerning the relation between temporality and causality should be asked. In example (9), the only possible relations are forward causal and forward temporal relations whereas in (10) several relations are possible: forward temporal and causal, forward temporal and backward causal, backward temporal and causal. Wilson and Sperber (1998) give an example where a causal relation occurs without a temporal relation as in (11).
(9) Socrate but un coup et tomba raide.  
‘Socrate drank and fell stone.’
(10) Marie cria et Pierre partit.  
‘Mary screamed and Peter left.’
(11) Susan is underage and can’t drink.

Hence, Moeschler (2000) proposed that causal and temporal relations are two sets of relations that can have a Boolean junction. This means that for two eventualities e1 and e2, there can exist an intersection of causal and temporal relations for which [e1 causes e2] implicates [e1 precedes e2]. He also suggested that two sentences can produce identical cognitive effects on the basis of different explications and implicated premises as in (12) and (13). In (12), the temporal relation [e1 precedes e2] is part of the explication while the causal relation [e1 causes e2] is an implicated premise. In (13), the causal relation [e1 causes e2] is part of the explicature while the temporal relation [e1 precedes e2] is part of the implicated premise.

(12) Max a laissé tomber le verre (e1). Il s’est cassé (e2).  
‘Max dropped the glass. It broke.’
(13) Le verre s’est cassé (e2). Max l’a laissé tomber (e1).  
‘The glass broke. Max dropped it.’

It was stated that there are several possible types of relations among eventualities, and this can be summarized in Fig. 1 from Moeschler (2000). The model considers temporal and causal relations. As far as temporal relations are concerned, they can be or not forward temporal inferences (temporal sequencing). In the case
where there is no temporal sequencing, there are two new possibilities: either there is or is not a backward temporal inference. And finally, if there is no backward temporal inference, then the cases of temporal simultaneity or indetermination can be identified. Temporal sequencing can be accompanied or not by a forward causal relation, as in (9) and (11) respectively. Backward temporal inference can be accompanied or not by reverse causality, as in (14) and (15) respectively.

(14) Max tomba. Jean l’avait poussé.
‘Max fell. John had pushed him.’
(15) Jean prepara son café. Il s’était levé sans entrain.
‘Jean prepared his coffee. He woke up without energy.’

Moeschler (2000) concludes that the natural way of presenting eventualities is not as a backward temporal inference, unless it is specified otherwise through the verb tense used and discourse connectives for example. He suggests that temporal or causal interpretations of a sentence are triggered by the speaker’s intention. The speaker may have a communicative intention of presenting the eventualities as they occurred and thus the interpretation consistent with the cognitive principle of relevance is the temporal one. The speaker may have a communicative intention of presenting eventualities focusing on the causal relation holding among them. In this case, the interpretation consistent with the principle of relevance is the causal one.

3 Temporal Coherence and Verb Tense

A contrastive analysis of verb tenses used in the examples given in Sect. 2.2 shows that FR is much more sensitive than EN to the type of relation expressed. If the EN SP can be equally used for forward and backward temporal and causal relations as well as temporal simultaneity, FR tenses seem more specialized. The FR PS seems specialized for the forward temporal relation (example (9)), the Plus-que-Parfait in backward temporal relations (examples (14) and (15)) and the IMP in the case of temporal simultaneity (example (7)).\(^5\) The Passé Composé (PC) does not impose any constraint relating temporal and causal inferences (examples (12) and (13)).

At this point of the discussion, I would like to introduce the conceptual/procedural distinction referring to types of information encoded by linguistic expressions. The conceptual/procedural distinction was proposed by Blakemore (1987) to explain differences between words with a conceptual content, such as table, cat, think or walk on the one hand, and discourse connectives, such as but, so or also. Content words

\(^5\)I assume that this specialization conducted Kamp and Rohrer (1983) to argue that the PS encodes a forward temporal inference. I argue that the PS encodes a procedure regarding directional temporal inference. In other words, it is an instruction for the hearer to ascertain the contextual value of the directional temporal inference.
encode concepts that contribute to the proposition expressed by an utterance while the meaning of a discourse connective is better described in terms of constraints on the inferential phase of interpretation than in conceptual terms. The hearer is expected to have access to the smallest and most accessible set of contextual assumptions in order to get the intended cognitive effects.

Verb tenses also have been investigated regarding the conceptual/procedural distinction and their role for discourse processing (Wilson and Sperber 1998; Moeschler 1993; Ahern and Leonetti 2004 for the Spanish subjective; Nicolle 1997, 1998; Moeschler et al. 1998; 2012; Saussure 2003; Leonetti and Escandel-Vidal 2003 on the Spanish imperfective; Améños-Pons 2010, 2011 on Spanish past tenses, to name but a few). Similarly to connectives, verb tenses are considered to encode procedural information consisting of instructions on how to manipulate mental representations of eventualities. Specifically, they give information about the temporal and causal relations holding among eventualities.

Since the first studies on verb tenses in a relevance theoretic framework, it is believed that they have rigid procedural meanings that help the hearer reconstruct the intended representation of eventualities (Nicolle 1998; Améños-Pons 2011; Saussure 2003, 2011). Saussure (2003) proposes algorithms to follow, consisting of the instructions encoded by verb tenses, in order to grasp the intended meaning of a verb tense at the discourse level. Taking the distinction conceptual-procedural as a foundation, Blakemore (1987), Wilson and Sperber (1993), Moeschler (1994, 1998) and Nicolle (1997, 1998) claimed that tenses have a procedural meaning. Nicolle (1998: 4) argues that tense markers impose constraints on the determination of temporal reference and thus they “may be characterized as exponents of procedural encoding, constraining the inferential processing of conceptual representations of situations and events”. Concerning the status of the temporal coordinates, Saussure and Morency (2012) argue that tenses encode instructions on how the eventuality is to be represented by the hearer through the positions of temporal coordinates. They consider thus that temporal location with the help of point of speech S, reference moment R and event moment E (Reichenbach 1947) is of a procedural nature.

A different view was defended in Grisot and Moeschler (2014). We argued that that location through temporal coordinates S and E does not constrain the inferential processing but contributes to the propositional content of the utterance. Seen from this perspective, temporal coordinates are conceptual parameters saturated contextually. More specifically, the hearer is brought to build an ad hoc concept as pastness or non-pastness based on contextual information.6 We considered Wilson and Sperber’s example (1993: 157) given in (16) and the propositional form given in (17). We add to this propositional form the information that eventualities of saying and of being tired took place before the moment when the sentence was

---

6For a development of the conceptualist view of tense based on experimental findings using the cognitive features of conceptual and procedural information proposed by Wilson and Sperber (1993), see Grisot (2015).
uttered. The extended propositional form would be something like the one given in (18). This temporal information cannot be cancelled or contradicted, as shown by the incompatibility with the adverb now or tomorrow in (19) and (20), as well as the compatibility with the adverb yesterday in (21).

(16) Peter told Mary that he was tired.
(17) Peter Brown told Mary Green at 3.00 pm on June 23 1992 that Peter Brown was tired at 3.00 pm on June 23 1992.
(18) Peter Brown told Mary Green at 3.00 pm on June 23 1992 (a moment before the present moment/in the past) that Peter Brown was tired at 3.00 pm on June 23 1992 (a moment before the present moment/in the past).
(19) *Peter Brown told Mary Green at 3.00 pm on June 23 1992 which is now (a moment contemporary with the moment of speech)/tomorrow (a moment which is after the moment of speech) that Peter Brown was tired at 3.00 pm on June 23 1992 which is now/tomorrow.
(20) *Now/tomorrow Peter told Mary that he was tired.
(21) Yesterday, Peter told Mary that he was tired.

In what follows, I will focus on the procedural content of verb tenses and their link to temporal coherence in discourse. Procedural information consists of instructions and constraints for contextual usages of a tense leading to the interpretation intended by the speaker. As noted in Sect. 2.2, French verb tenses expressing past time seem specialized in expressing forward and backward temporal and causal relations. Moeschler (2002) makes the hypothesis that this “specialization” is due to the procedural content encoded by French verb tenses. Specifically, one type of procedural information concerns temporal and causal relations among eventualities. In Grisot and Moeschler (2014), this procedure is called the [± narrative] feature. Narrativity is a binary pragmatic feature: in narrative usages, a verb tense expresses eventualities (events/states) that are temporally ordered accompanied or not by a causal relation, while non-narrative usages express temporal simultaneity or temporally un-related states of affairs.

---

7The narrativity feature is a coarse grained semantic and pragmatic feature proposed in a particular framework, that of Natural Language Processing and Machine Translation. I admit that from a theoretical point of view, only finer coarse features can explain all possible (both very frequent and less frequent) usages of a verb tense.
8Verb tense has frequently been associated with narrative contexts in various frameworks, such as in DRT and SDRT. Smith (2003) discussed discourse modes based in textual structure and aspect. If these theories focused on linguistic information and made use of non-monotonic inferences, for us narrativity is procedural information representing a cognitive (as opposed to logic in (S)DRT) discourse relation (Hobbs 1979; Mann and Thompson 1988; Sanders et al. 1992). Cognitive discourse relations are expressed lexically through verb tense and connectives (that are language-specific) and can occur in any type of stylistic register.
9At this point of the research, I consider lexical aspect as one class and do not distinguish between Vendler (1957). Moeschler (2000) discusses Dowty’s principle of interpretation of temporal discourses (1986) based on lexical aspect. Moeschler (2000) argues that this approach to temporal relations adopts a radical position and does not explain a certain number of exceptions.
If we go back to Fig. 1 presenting the possible relations between two eventu-
 alities e1 and e2, we can observe that a narrative usage of a tense assembles
temporal sequencing (with forward or no causal relation) and backward temporal
sequencing (with reverse or no reverse causality). A non-narrative usage of a tense
 corresponds to temporal simultaneity. As far as the last possible relation is con-
cerned, Moeschler (2000) suggests that two cases are possible. Firstly, one of
the three previous relations can be determined on the basis on contextual assumptions
and mutual cognitive environment. Secondly, no relation can be determined. The
case of un-related eventualities was included under non-narrative usages of a tense.

Our research in the COMTIS and MODERN projects aimed at identifying
semantic and pragmatic features that would improve the translation of verb tenses
by statistical machine translation systems. In order to achieve this purpose,
empirical research was carried out consisting of corpus work (Sect. 4.1), followed
by human and automatic annotation experiments with the [± narrative] feature
(Sect. 4.2).

4 Empirical Work

4.1 Corpus Work and Verb Tense Translation Divergences

Grisot and Cartoni (2012) studied the discrepancies between theoretical descrip-
tions of verb tenses and their use in parallel corpora. The corpus consists of texts in
EN and their translations into FR belonging to four different genres with the fol-
lowing distribution: literature (18 %), journalistic (18 %), parliamentary discussions
(31 %) and legislation (33 %). A total of 1275 predicative verb tenses have been
considered, which represents 77 % of the verb tenses occurring in the corpus.

Corpus analysis was done in two steps: (i) a first monolingual step in order to
calculate the frequency of verb tenses in a source language (SL), and (ii) a second
bilingual step in order to identify the tenses used as translation possibilities into a
target language (TL) for a certain tense in a SL. Calculating the frequency of tenses
in the corpus allowed us to verify if verb tenses that are considered to be
ambiguous, are also frequent in corpora. Quantitative analyses of tokens of SP in
our corpora enabled us formulate statistically significant observations.

The monolingual analysis of the corpus containing texts in EN as SL showed
that the most frequent tenses are: the Simple Present (32 %), the SP (25 %) and the
Present Perfect (14 %). The contrastive bilingual analysis of the parallel corpus
revealed that the SP is one of the most ambiguous EN verb tenses, as far as its
translation into FR is considered. The FR tenses used to render the semantic and
pragmatic meaning of the SP are: the IMP used most often in literature (44 %), the
PC used most frequently in the journalistic register (58 %), the PS used most
frequently in the literature register (40 %). The Present tense is used in the legislation register (10 %) in order to create a deontic effect.\textsuperscript{10} This distribution shows that genre is not a good predictor, as it could have been expected, i.e. in the literature genre the SP is translated with an IMP in 44 % of the cases and with a PS in 40 %.

In another corpus-based study (Grisot and Costagliola, 2014), we studied the translation of the EN SP into Italian (IT) and Romanian (RO). The parallel corpus consists of texts from four different genres with the following distribution: literature (37 %), journalistic (18 %), parliamentary discussions (19 %) and legislation (26 %). There are three most frequent verbal tenses used in these Romance languages for the translation of the SP, as shown in Table 1.

A question that could be asked is how can the translation divergence of the SP be explained? Two answers are possible:

(A) The SP is polysemantic – which means it has several meanings and each meaning is translated into FR, IT or RO through a different tense.

(B) The SP is underdetermined – which means that the meaning of the SP must be contextually worked out by assigning contextual values to both conceptual and procedural information. The consequence of the underdetermined meaning of the SP is that it has several contextual usages and each contextual usage is translated into FR, IT or RO through a different tense.

As suggested by Smith (1990) and based on the results of the empirical work, we adopted the second possibility. To account for the translation divergence of the SP, Grisot and Moeschler (2014) made the hypothesis that the procedural feature [± narrative] can be used for disambiguating among different possible usages of the SP. As mentioned in Sect. 2, the link between verb tenses and temporal coherence in discourse has been discussed for FR verb tenses. One of the advantages of translation corpora is to permit cross-linguistic transfer\textsuperscript{11} of semantic and/or

\begin{table}
\centering
\caption{Translation possibilities of the SP into FR, IT and RO in the corpus}
\begin{tabular}{|l|c|c|c|}
\hline
   & FR (%) & IT (%) & RO (%) \\
\hline
PC   & 37     & 33     & 49   \\
IMP  & 24     & 18     & 15   \\
PS   & 16     & 22     & 18   \\
Others & 16    & 21     & 13   \\
\hline
\end{tabular}
\end{table}

\textsuperscript{10}The translation of the SP through a present tense form can be explained by the contextual values taken by temporal coordinates S, R and E in order to lead to the speaker’s intended interpretation. Specifically, the translation with present time signals that the eventuality is viewed from the moment of speech (R = S) (see Grisot and Moeschler, submitted for publication, where we argue based on experimental results that temporal coordinates are conceptual information).

\textsuperscript{11}Samardzic (2013) uses this novel methodology for investigating the translation equivalents of a range of English light verb constructions into several languages. Slavic languages encode verb aspect lexically, unlike other European languages. She applies the aspectual representation obtained in the English-Serbian cross-linguistic setting to classify English verbs into event duration classes.
pragmatic information. This is due to the fact that studying instances and usages of verb tenses in a parallel corpus make it possible to control for context and register variables.

In the two experiments we designed, we made use of the methodology of cross-linguistic transfer of properties. Specifically, the [±narrativity] feature was tested and validated in two experiments with human judges. The working hypotheses of the first experiment were formulated based on theoretical descriptions of FR tenses. The second experiment was designed based on the cross-linguistic transfer of semantic and pragmatic information from FR past tenses to the EN SP. These experiments resulted in the creation of human annotated data that served as training data for an automatic classifier. The automatic classifier at its turn was integrated into a statistical machine translation system.

This interdisciplinary research from the COMTIS and MODERN projects lead to a series of publications, as follows. The human experiments are presented in detail in Grisot and Moeschler (2014). The NLP work described in Grisot and Meyer (2014) consists of the automatic labelling of the corpus with the narrativity feature. And finally, the MT work and its results are described in Meyer et al. (2013). In the following section, I will give a brief summary of the experiments, both for human and automatic annotation, as well as the implementation of the model for improving the results of statistical machine translation systems. I argue that the fine-grained theoretical model proposed by Moeschler (2000, 2002) for temporal coherence is empirically and cross-linguistically valid.

4.2 Human Judgement, Application to NLP and Improving the Results of MT

Kamp and Rohrer (1983) argued that temporal sequencing is semantic information encoded by verb tenses: the PS makes time to advance whereas with the IMP time does not advance (i.e. it is a verbal tense of background information). In other words, the PS has a narrative usage and the IMP has a non-narrative usage. Experiment 1 was designed to test this theoretical assumption. Seventy-six FR native speakers participated in this experiment and were instructed to judge sentences randomly chosen (15 sentences per participant) from the FR part of the corpus regarding the [±narrative] feature. The two possible values of feature were explained and illustrated with several examples. The sentences contained one of the three FR tenses of interest: PC, PS and IMP. The results\textsuperscript{12} of the experiment indicate that:

\textsuperscript{12}These results include only cases where inter-annotator agreement is high. Four participants judged each sentence and the result is based on the majority of answers.
(A) The PS was judged as having a narrative usage in 92% of the cases;
(B) The PC was judged as having a narrative usage in 77%;
(C) The IMP as having a non-narrative usage in 77.5%.

This leads to about 23% of the cases where the PC has a non-narrative usage and the IMP has a narrative usage. These results are consistent with the theoretical descriptions (Moeschler 2000, 2002; Saussure and Stihoul 1999, 2005 for the IMP; Saussure 1997 for the PS; Lusher and Stihoul for the PC). Moeschler (2002) suggests that the PS is specialized for forward temporal and causal and it imposes this interpretation even in cases when lexical information (order of eventualities) is not compatible. The PC is not specialized for directional temporal and causal relations allowing all possible relations presented in Fig. 1. The IMP is expected to occur mainly in non-narrative contexts (without a temporal sequencing role). The results show that in 23% of the cases, the IMP has narrative usages boosted by a subjective reading and accompanied by a temporal adverbial that increments the reference point R. The consequence is the forward temporal interpretation, as in example (22). Sentences in (23), (24) and (25) are the translation into FR, IT and RO of (20) taken from the multilingual translation corpus described in Grisot and Costagliola (2014).

(22) Alice was beginning to get very tired of sitting by her sister on the bank, and of having nothing to do. Once or twice she had peeped into the book her sister was reading, but it had no pictures or conversations in it, "and what is the use of a book," thought Alice, "without pictures or conversations?" (Literature Corpus, Alice’s adventures in Wonderland, L. Carroll)

(23) Alice commençait à se sentir très lasse de rester assise à côté de sa sœur, sur le talus, et de n’avoir rien à faire: une fois ou deux, elle avait jeté un coup d’œil sur le livre que lisait sa sœur; mais il ne contenait ni images ni dialogues: Et, pensait Alice, “à quoi peut bien servir un livre où il n’y a ni images ni dialogues?”

(24) Alice cominciava a sentirsi mortalmente stanca di sedere sul poggio, accanto a sua sorella, senza far nulla: una o due volte aveva gittato lo sguardo sul libro che leggeva sua sorella, ma non c’erano immagini nè dialoghi, “e a che serve un libro,” pensò Alice, “senza immagini e dialoghi?”

(25) Alice începuse să se simtă foarte obosită stând pe bancă lângă sora ei și neavând nimic de făcut: o dată sau de două ori trase cu ochiul la cartea pe care sora ei o citea dar nu avea poze sau dialoguri „Și care e rostul unei cărți” se gândi Alice, “fără poze sau dialoguri?”

These examples show that multilingual translation corpora reveal information that is hidden in a monolingual or even bilingual analysis. The EN SP from (20) has a narrative usage (forward temporal inference) and is translated into IT and RO through tenses with a preterit form: passato remoto and perfectul simplu respectively. The situation is different for FR, where the SP is translated through an IMP. A close look at the FR sentence allows us to identify that Alice’s subjective
perspective on the event of thinking and the discourse connective et (“and”) accompany the IMP. In this sentence we are dealing with a narrative IMP (“imparfait de rupture”) that becomes uncovered in a multilingual contrastive analysis.

Experiment 2 was designed to test the narrativity feature for the SP based on cross-linguistic transfer of semantic and pragmatic features. Two English native speakers participated in this experiment and were instructed to judge 458 sentences randomly chosen from the EN part of the corpus regarding the [+narrative] feature. The two possible values of feature were explained and illustrated with several examples. The results showed that annotators agreed on the label for 325 sentences (71 %) and disagreed for 133 sentences (29 %). Disagreements were resolved in a second round of the annotation experiment, where annotators were asked to insert the connectives corresponding to the temporal and/or causal relation existent between the two eventualities considered in order to test if the narrative or non-narrative label was appropriate. Annotators disagreed only on 4 sentences.

These two experiments show that the [+narrative] feature, consisting of the possible temporal and causal relations described in Fig. 1 (Moeschler 2000) is cross linguistically valid (i.e., it has been validated for FR, EN, IT and RO as discussed in Grisot 2015). In order to be use this feature for MT, a much larger amount of annotated data is needed. Because human annotation is costly, the labels should be given automatically. This is the task of an automatic classifier (described in Grisot and Meyer 2014). The automatic classifier was trained on the human annotated corpus. When labelling an unknown corpus, the classifier had similar results to human annotators.

A first run of an SMT system, which uses the classifier trained on the annotated data with the [+narrativity] feature, had slightly better results than without this pragmatic feature. When trained and tested on automatically annotated data, the [+narrativity] feature improves translation by about 0.2 BLEU points. Manual evaluation shows that verb tense translation and verb choice are improved by respectively 9.7 % and 3.4 % (absolute), leading to an overall improvement of verb translation of 17 % (relative) (for more detailed results see Meyer et al. 2013).

5 Discussion and Conclusion

Jacques Moeschler’s predictive model for temporal coherence in discourse is a complex model consisting of several factors: lexical information (e.g. push–fall), procedural contents of verb tenses and of connectives and contextual assumptions. He suggests that these factors do not have the same status, but they are in a hierarchy. Specifically, he proposed the following hierarchy:

---

13BLEU (Bilingual Evaluation Understudy) is an evaluation measure for machine-translated texts. It calculates the degree of resemblance to a human-translated text and it is a number between 0 and 1, where values closer to 1 represent more similar texts.
The Model of Directional Inferences (MDI) proposed in Moeschler (2000, 2002) uses this hierarchy. He suggested that verbs (lexical information) and tenses convey weak directional features while discourse connectives and contextual assumptions convey strong directional features. This hierarchy is useful in case of conflicting contextual and linguistic information.

I have shown in this paper the role played by one type of procedural information encoded by verb tenses, the [+narrativity] feature, for temporal interpretation of a discourse. This feature was empirically validated by human and automatic systems in a multilingual translation corpus. Moeschler’s MDI model predicts that temporal coherence in discourse depends not only on temporality (discourse temporal relations) but also on causality (discourse causal relations). I have shown that this information is useful for improving the results of SMT systems in terms of the choice of the verb and of the tense.

Finally, Grisot (2015) provides a more complex model for temporal coherence in discourse that includes grammatical and lexical aspect, tense, temporal adverbials and temporal connectives, as well as other linguistic and non-linguistic cues. Further research is needed to test each factor empirically, as well as their interaction with the help of complex statistical models.

References


14Some of these suggestions are investigated in (Grisot 2015).


Grice HP (1967) Logic and conversation: The william james lectures. Harvard University, MS


Hobbs JR (1985) On the coherence and structure of discourse. Center for the study of language and information, Stanford University, pp 85–37


Reichenbach H (1947) Symbolic logic. University of California, Berkeley
Temporal Coherence in Discourse: Theory …


