Information flow in excerpts of two translations of *Mme Bovary*

Alexandre Sévigny
McMaster University

This article explores how information is accumulated and collated in a cognitively realistic fashion in two very short excerpts of translations of Flaubert's 'Mme Bovary'. The approach taken is a formal cognitive linguistic one using Discourse Information Grammar (DIG), a theory of grammar based on the intuitive idea that texts are understood by the reader incrementally, in a left-to-right fashion. Thus, a cognitive pragmatic approach is taken to the study of the excerpts, highlighting how much information is accumulated as the reader develops an understanding of the text in question. The analysis discusses the differences in the build-up of information in the source text and in its translations. The conclusion indicates that translation studies contribute much to the development of formal linear cognitive linguistic theories.

1.0. Introduction

This article explores how the information contained in two short extracts from *Mme Bovary* (Flaubert 1857/1966) compares in translations by May (Flaubert 1953) and Hopkins (Flaubert 1949), when examined from the perspective of time-linear information flow. We adopt Discourse Information Grammar (DIG), a formal cognitive linguistic approach developed to model the dynamic left-to-right accumulation of information in a text (Sévigny 2000, 2002, forthcoming). The brevity of the excerpts studied is intentional as it is part of our purpose to highlight the extraordinary amount of information involved in the linear understanding of even short texts. It is our claim that taking this approach provides insights into the reader's incremental reconstruction and comprehension of the texts through time. What relations, if any, can be made between the information contained in the source text and the translation(s)? This paper examines three of these relations, as textual information develops incrementally: (i) sequencing, (ii) ambiguity, and (iii) situational equivalence.

1.1. Sequencing, Ambiguity, and Situational Equivalence

DIG brings out aspects of translation which human processors may not be aware of for a number of reasons. We will look at three of these aspects: (i) **sequencing of information:** as we will see, the information contained in texts, to a large extent, tends to be order-independent; (ii) **ambiguity in information:** texts contain a surprising number of ambiguous instances, many of which might not be apparent to a human language processor; (iii)

the problem of equivalence: translation introduces an additional complexity: situational equivalence. Here, there are several points to consider, among which we may list: participant equivalence, relational equivalence, circumstantial equivalence, logical equivalence and degree of equivalence.

2.0. General Description of Discourse Information Grammar

Formal grammars, in the generative tradition, attempt to construct systems which model the initial state of the language faculty, using the most economical set of rules possible (Chomsky 1995, 2000, p.c.). However, traditional generative grammars have always skipped over the incremental linear process of utterance construction, focusing, rather, on the structures which are its end result. Another current in formal grammar theory, only recently explored, is that of the time-linear construction of sentences. This approach has been the subject of few, but serious, attempts at grammar construction: Dynamic Syntax (Kempson et al. 2001), LA-Grammar (Hausser 2001), LPSG (Shin 1987) and DIG (Sévigny 2000, 2002, forthcoming). These grammatical theories approach the formal derivation of linguistic structures as a left-to-right model of the natural language understanding process. Two of these theories, Dynamic Syntax and DIG, are inspired by the relevance-theoretic model of natural language understanding (Sperber & Wilson 1986/1995), which predicts that hearers understand discourse from left to right, accumulating information, matching it against their knowledge base of accumulated information and anticipating possible relevant interpretations.

Now a very brief description of how Discourse Information Grammar works. DIG rests upon several principles: (i) a strict adherence to linear, defeasible, incremental information accumulation; (ii) a lexicalist approach to language and (iii) the use of complex-feature structures for capturing information rich enough to allow the simulation to be cognitively probable. The use of feature structures (Carpenter 1992; Shieber 1986) allows additional refinements without needing to undo what has been established, and the logics necessary to manipulate feature structures are relatively well understood. In addition, DIG also uses several sorts of information types. Of these, the following will be mentioned in this short paper: lexical, structural, functional, anaphoric, semantic and situational. The DIG method is based on the common sense notion that words are processed as they are input/read. Each newly input word receives minimally specified default lexical information. Since DIG uses default and/or partial specification whenever it is necessary and underspecification otherwise, it is frequently the case that lexical and feature values are not specified until sufficient information has been accumulated for such specifications to be possible. DIG also uses grammatical rules, which form part of its accumulated linguistic pattern database. Such knowledge is meant to simulate our ability to recognize instantly correct patterns or usage or even to reject a correct pattern as 'ungrammatical' in those cases where we have never encountered such a pat-

tern before. Gradually, structures are built up, each structural type contributing types of information as well as constraints and checks for consistency in the various types of information accumulated. These structures are integrated into the discourse stream according to strict rules.

As DIG gathers and integrates information, situations are generated. These situational constructs are similar in some ways to the notion of situation used in situation semantics (Barwise & Perry 1983), but are more general in their application. Informally, a situation in DIG consists of the triplet: <P, R, C>, where

> P = the set of participants, usually denoted by nominal structures of various sorts;
> R = the set of relations, usually denoted by verbs and linkers with their respective arguments;
> C = the set of circumstances and modifications, usually either adverbial or adjectival.

The generic structure of a situation is quite simple. This is not a surprising fact given the limitations of human short-term memory, which must be respected if simulations of human language processing are to be cognitively probable. However, situations can become quite involved because structural embedding, serialization and links of various sorts operate to create composite Ps, Rs and Cs. The result is usually a rich and complex network of relations which act together to stabilize ongoing relation construction. Meaning and information are also constrained without compromising the integrity of the unified attribute feature values of the various components built up so far. Situations are thus meta-linguistic constructs which constitute contexts for the development of linguistic information.

Some results of applying DIG to a text are: (i) an incremental representation of the accumulation of the pieces of information sufficient for the understanding of a text; (ii) a fully tagged text, including words, structures and functional relations; (iii) a comprehensive and relational representation of the accumulation of the text's content through time, including various relations: anaphoric, structural, functional, logical, semantic. The results can be used to extract various types of information about the time-based linguistic analysis of text and to arrive at a synthesis of the text as a sequence of cognitive procedures, rather than a static synthesis of the text as a simple result. This leads to comparative interpretations of the text's content, and incremental development in the simulated mind of the receiver. Several of these results are illustrated in this paper.

3.0. Example 1: Text 1 from *Mme Bovary*

Take the following short text from Mme Bovary:

> (1) "L'enfant d'Emma dormait à terre, dans un berceau d'osier. Elle la prit avec la couverture qui l'enveloppait ... " (Flaubert 1857/1966: 126)

To illustrate, to a certain extent, how DIG incrementally accumulates information of various sorts, the information contained in this text will be described in terms of its unfolding in partial stages. In actuality, DIG works using one word at a time, as mentioned earlier. Also, many details have been left out. For instance, DIG must resolve whether 'l'' is the definite article or the object form of the personal pronoun before a word beginning with a vowel.

Input: <L'enfant d'Emma>:
'enfant': category: common noun; structure type: nominal structure; semantic features: +human, +animate, +young, +definite, belongs to 'Emma'; gender unspecified. 'Emma': proper name for a person: +adult, +feminine, +singular. 'Enfant' is at the top level (meaning it is not embedded within another structure) and +singular, +3rd person. Functional role: SUBJECT? TOPIC? DIRECT OBJECT? It is not yet possible to specify the role unambiguously.

Input: <dormait>:
'nominal-1 as DOER or SUBJECT. Situation-1 is initiated: DOER: 'L'enfant d'Emma'; relation: EVENT: dormir(l'enfant d'Emma, {complement(s)}). Logical type: ASSERTION. Situational type: DESCRIPTION or NARRATION, as indicated by tense and aspect values of 'dormir()'.

Input: <à terre>:
'à terre': fixed prepositional phrase which indicates static location. This information triggers the functional relation MODIFICATION: ADVERB: LOCATION. It applies to the earlier verb 'dormait', since it cannot apply to a nominal. The situation has now grown to 'dormir(l'enfant d'Emma, à terre)'. Its status of +DESCRIPTION remains valid.

Input: <,>:
DIG treats commas as separators. Separators trigger several possibilities: elaboration, continuation, serialization, or error.

Input: <dans un berceau d'osier>:
'dans': embedding linker which triggers MODIFICATION. It may be +time or +location. It initiates a prepositional phrase which modifies the word preceding it. Since 'dans' is preceded by a separator, error is eliminated. The prepositional phrase triggers the functional relation MODIFICATION: ADVERB: LOCATION, and the relation triggered by the separator ',' can now be resolved to serialization, which in this kind of context can be reduced to a form of elaboration. The situation can now be described as 'dormir(l'enfant d'Emma, à terre, dans un berceau d'osier)'. Note also that 'berceau' has, among its several semantic attributes +object, +container, +place. This reinforces the location attribute of +place for the linker 'dans'. It is also compatible with the static nature of the event captured by 'dormir()'. The following phrase 'd'osier' is processed in similar fashion. Since 'de

(= d')' is an embedding linker following a nominal ('berceau'), the prepositional phrase it triggers will, in turn, trigger the functional relation MODIFICATION: ADJECTIVE: ORIGIN/MATERIAL which will be linked to 'berceau'.

Input: <.>:
DIG treats the period as a terminator. Terminators trigger complete closure of all previous structures and following this, a resetting of DIG for new input. Thus far, DIG has recorded the following information (Note that this is not an exhaustive synopsis!):

Situation:
PARTICIPANT: DOER: l'enfant d'Emma
RELATION: dormir(l'enfant d'Emma, {complement})
COMPLEMENT: à terre, dans un berceau d'osier

Input: <Elle la prit avec la couverture qui l'enveloppait ...>:
To avoid needless technical repetition, processing internal to this sentence fragment is assumed. 'Elle' is resolved to Emma, with functional role of SUBJECT. Hence a situation has begun with DOER identified as Emma. The word 'la' is resolved to be a personal pronoun, object form, not the definite article. Its antecedent is 'enfant'. This, in turn, allows the specification of the unspecified attribute gender for 'enfant' to be marked +feminine. The word 'prit' is processed as 'prendre(nominal-1: +animate, nominal-2, {complement})'. Through functional role assignments and bindings, 'prendre()' becomes specified to 'prendre(Emma, enfant, {complement})'. The tense is +past, ASPECT: +punctual. The event type is +process. The phrase 'avec la couverture' is somewhat problematic at first analysis: is 'couverture' +instrument or +accompaniment? There are two interpretations:

I-1: more likely: prendre(Emma, l'enfant et la couverture) +process, +action, +narration
envelopper(couverture, enfant) +static, +description.

I-2: less likely: prendre(Emma, l'enfant, avec la couverture +agent)
envelopper(couverture, l'enfant) +static, +description

Here, we have a mild example of participant-2 ambiguity. It is unclear which of the following two situations is the correct one:

Situation-2a:
PARTICIPANT-1: DOER: Elle = Emma; PARTICIPANT-2: DO: l'enfant et la couverture
RELATION: prendre (Emma, la = l'enfant, la couverture, {complement})
COMPLEMENT: 0

Situation-2b:
PARTICIPANT-1: DOER: Elle = Emma; PARTICIPANT-2: DO: l'enfant
RELATION: prendre (Emma, la = l'enfant,)
COMPLEMENT: avec la couverture: + agent

However, when 'qui l'enveloppait ....' is processed, the default or preferred interpretation is +accompaniment. The text is marked as +DESCRIPTION /NARRATION. The ambiguity just described might appear unusual to a human processor, but the linker 'avec' can signify accompaniment or instrumentality. Until it is perfectly clear what is meant here, both possibilities exist. As lexical properties become refined, a number of semantic links can be built in to such verbs as 'prendre' and 'envelopper' to enable default resolution of such ambiguities. It is also noteworthy to point out that, as a text develops, specifications act as constraints on the possible attribute values of words and structures. This usually makes the discursive context clearer as input accumulates.

3.1. Translation of Text 1, by G. Hopkins

Examination of translations reveals that there is more ambiguity than at first assumed. For instance, let us process the following text by G. Hopkins:

> (2) "Emma's child was sleeping on the ground in a wicker cradle. She took it up, wrapped in its blanket ...." (Flaubert 1949:111)

'Emma's child' is straightforward and poses no problem. The functional role is not clearly specifiable at this stage, but as soon as 'was sleeping' is processed, it becomes possible to bind 'Emma's child' to nominal-1 in 'sleep(nominal-1, {complement})'. This generates a situation where the relation is 'sleep(child, {complement})'. The DOER role is unambiguously assigned to 'child'. The tense is marked +past, +durative, a tense used to express description in the past. Continuing, the phrase 'on the ground' indicates location. Therefore, it can be bound with the optional complement to yield: sleep(child, on the ground)'. The fact that the next phrase 'in a wicker cradle' follows directly after 'ground' triggers some ambiguity: is this the ground in the wicker cradle (which ground? the ground in the wicker cradle?), or is this the wicker cradle on the ground? To a formal system, it is not completely clear which 'ground' is referred to here: the ground (= bottom) of the wicker cradle? or the ground (= floor) of the [room])? The problem lies with the fact that 'ground' has several possible meanings: dirt, floor, bottom. Additional information would be needed to reduce the number of possibilities, preferably to one. The next text fragment also generates ambiguity. First, 'She took it up' can be resolved to 'she = Emma' and 'it = child'. The choice of 'it = wicker cradle' or 'ground' is less likely because 'take up' is not exactly the same as 'pick up'. Usually, 'take up(nominal-1, nominal-2, {complement})' when it means to 'pick up', and not 'to start', will be

used to refer to someone small and dear. Applying DIG to Hopkins's text, brings out these nuances of 'take up', with the result that this short harmless looking text becomes quite ambiguous. Continuing, we process the comma which, in DIG, is a separator. Normally, it is used to generate topicalization, or elaboration, usually in the form of apposition or serialization. In this case, it is not immediately possible to be clear since visually, there are no indications of how to read the word 'in'. If the expression 'wrapped in' is to be read with stress on 'in', then 'wrapped' is a finite verb, serialized with 'took it up' and the situation generated now has two sequenced events: 'take up(Emma, child, {complement})', and 'wrap in(Emma, blanket)'. If there is no stress on 'in', then we have a form of elaboration with stress on the word 'wrapped', which is now designed as a past participle describing an aspect of the condition of the child. In this second case, the situation generated has one event and additional descriptive detail. The verb(s) is/are marked +past, ASPECT: +punctual. So far, this text has generated at least three types of ambiguity: location ambiguity, association ambiguity and event ambiguity. Without further information, it is difficult to resolve these ambiguities automatically.

3.2. Translation of Text 1, by J. Lewis May

Next, we look at the translation fragment by J. Lewis May. It is evident that May was aware of some of the ambiguities in Flaubert's text.

> (3) "Emma's baby was asleep in a wicker cradle on the floor. She took it in her arms, wrapped up in its coverlet ...." (Flaubert 1953:108)

The fragment 'Emma's baby was asleep' generates the same information as Hopkins's text, except that the word 'baby' is more specific than 'child', in terms of specified feature attributes. The next fragment 'in a wicker cradle' clearly specifies that the complement of 'be asleep(baby, {complement})' is the prepositional phrase 'in a wicker cradle', which is assigned the functional role of MODIFICATION: ADVERB: LOCATION. The following segment 'on the floor' is processed as a prepositional phrase. Since it follows directly after 'cradle', a nominal, it can be assigned the functional role MODIFICATION: ADJECTIVE: LOCATION: SPECIFICATION. That is, the phrase modifies the nominal 'cradle' by specifying its location. This, in turn, limits its identity. If the system interpreted 'on the floor' in a relation of elaboration through serialization it could yield an interpretation where 'on the floor' modified 'be asleep(baby, {complement})'. This would lead to an absurd situation where there are two locations for the baby: in a wicker cradle and on the floor, presumably near or beside the cradle. This would be rejected unless there was textual indication that this was the desired interpretation. In this instance, 'in' clearly indicates that such is not the case.

Sometimes, it is difficult to imagine how these ambiguities can come about. Part of our difficulty, as human language processors, is that we bring

an enormous amount of various forms of knowledge (about the world, our experiences, our culture) to all situations. For instance, it is common sense that a child would sleep in a wicker cradle, not beside it on the floor. However, a formal system has no such knowledge, unless we have programmed this knowledge into its database. This is one of the purposes of DIG: to extract information which humans take for granted. Only in doing so, can we hope to simulate the associative richness of the words and relations which humans use to communicate on a regular basis.

The situation generated by May's text, then, is very similar to the situation generated by Hopkins's text, but with essentially no ambiguity. The next fragment is also quite clear: 'She took it in her arms' resolves to 'she = Emma', 'it = baby', 'her = Emma'. This generates a clear situation where 'take(nominal-1, nominal-2, {complement})' is specified to 'take(Emma, baby, in her arms)'. The event 'take()' is marked +past, ASPECT: +punctual. Next, the separator ',' is processed, generating possible elaboration or serialization. Following this, 'wrapped up' generates temporary ambiguity, similar to the case discussed for Hopkins, but the phrase 'in its coverlet' makes it clear that 'wrapped up' does not have a second nominal. This eliminates 'wrap(nominal-1, nominal-2, {complement})' and leaves the interpretation 'wrap up(nominal-1, {complement})' as the contextually acceptable interpretation. It is specified to 'wrap up([Emma], baby, in its coverlet)'. Additionally, the event represented by 'wrap up()' is marked +passive. This, in turn, makes it possible to interpret 'wrapped' as a past participle, modifying the nearest applicable nominal. In this situation, that nominal is 'baby'. The participial phrase 'wrapped up in its coverlet' is assigned the functional role MODIFICATION: ADJECTIVE:DESCRIPTION. The verbal element 'wrapped up' is marked +state. Finally, the suspension points followed by '.' indicate that the text has ended. The prepositional phrase 'in its coverlet' is itself assigned the functional role MODIFICATION: ADVERB: MANNER. The final result is summarized below:

Situation:
PARTICIPANT: DOER: Emma's child
RELATION: be asleep(Emma's child, {complement})
COMPLEMENT: in a wicker cradle on the floor

SEQUENCE. This is an operator which indicates that the situations are to be serialized and unified into a single situation.

PARTICIPANT: DOER: Emma
RELATION: pick up(Emma, child, {complement})
COMPLEMENT: in her (= Emma) arms

PARTICIPANT: DOER: Emma? someone? ('unspecified' is assigned the value 'someone')
RELATION: wrap up(agent?, Emma's child, {complement}) +state
COMPLEMENT: in its (= child) coverlet

Although we assume that the agent who wrapped up the child in its coverlet was Emma, there is no way of deducing this information from the text. By default, however, the agency ambiguity is resolved by assigning this role to 'Emma' unless there are overt indications that this is not to be done, which is not the case in this example.

4.0. Example 2: Text 2 from *Mme Bovary*

The first excerpt involved two situations which were eventually unified into a single situation. In this example, we will deal with a single situation in the original text which was resolved into two or more situations in the translations of that text. This example was chosen to illutrate the problem of situational equivalence. Take the following short extract from *Mme Bovary*:

> (4) "Alors on vit descendre du carrosse un monsieur vêtu d'un habit court à broderie d'argent, chauve sur le front, portant toupet à l'occiput, ayant le teint blafard et l'apparence des plus bénignes" (Flaubert 1857/1966:170)

Processing this text segment up to its first separator, we obtain the following information:

> **Situation**:
> RELATION: voir(nominal-1, nominal-2, {complement})
> PARTICIPANT-1: DOER: nominal-1 = 'on'
> PARTICIPANT-2: DO: nominal-2 = 'descendre(nominal-1, {complement})
> nominal-1 = un monsieur
> COMPLEMENT = du carrosse

'vêtu': category: participle, from 'vêtir(nominal-1, {nominal-2}, {complement})', marked +passive. 'vêtu' initiates a participial phrase to which the functional role MODIFICATION: ADJECTIVE: DESCRIPTION is assigned. The participial phrase is connected to nominal-1: 'un monsieur'. The phrase 'd'un habit court' is assigned the functional role of MODIFICATION: ADVERB: MANNER and bound to complement-1, yielding 'vêtir$_{+PASSIVE}$(un monsieur, d'un habit court). Since 'court' has category 'adjective', it is assigned the functional role MODIFICATION: ADJECTIVE: DESCRIPTION and connected to the nominal 'habit'. By default, DIG will attempt to connect the prepositional phrase 'à broderie' to the immediate word, 'court', but relevance-driven semantic attribute checks will indicate an incompatibility between the attributes of 'broderie' and 'court'. 'Court' being rejected, DIG will try 'habit' and find compatibility. This will require that 'à broderie' be assigned the functional role MODIFICATION: ADJECTIVE: DESCRIPTION and be connected to 'habit'. Thus, 'vêtu' has only one complement: 'd'un habit court', whose

HEAD 'habit' is itself modified by 'court' and 'à broderie d'argent'. Thus, 'court' and 'à broderie d'argent' are serialized and modify the nominal 'habit'. 'D'argent' is embedded and, by default, connected to 'broderie'.

The details resulting in these functional value assignments follow from the linkers and verb forms used. For example, 'vêtu', a past participle, triggers MODIFICATION: ADJECTIVE: DESCRIPTION. The linkers 'de', 'à' trigger MODIFICATION. The specific type of MODIFICATION is constrained by the word to which the modifiaction is bound. Normally, it will be the word immediately before the linker but, as we saw above, this condition cannot always be met. These decision simulations mirror the human ability to connect words, including our occasional difficulty to decide exactly what is being described and how. There are times, moreover, when humans cannot find a relevant utterance interpretation and must make a special inquiry in order to establish the desired connection.

Schematically, we now have a series of embedded modification relationships which specify various semantic attributes pertaining to 'monsieur', 'vêtu', 'habit', 'court' and 'broderie'.

So far, we have been dealing with a straightforward mechanism in French: using serialized, increasingly nested complementizer phrases to effect increasingly constrained descriptions. There has been only one active event: 'voir()' with two implied sub-events: the first nominalized via the infinitive 'descendre()' and the second passive and reduced to a past participle 'vêtu'. Being verbal, these categories are assigned the normal arguments for verbs. When the separator ',' is processed, it triggers several possible functional roles: elaboration, serialization or topicalization.

Continuing, we process until we encounter a non-structural signal, again a separator. The segment 'chauve sur le front' closes the previous structure, which recursively closes structures until the top level is reached. This serializes 'chauve + {complement}' with 'vêtu + {complement}'. Thus, 'chauve' triggers MODIFICATION: ADJECTIVE: DESCRIPTION and connects to 'monsieur'. The linker 'sur' initiates a prepositional phrase modifying 'chauve'. Hence, it is assigned the functional role MODIFICATION: ADVERB: LOCATION.

The situation initiated from 'Alors ...' is still not closed completely since separators effect partial closure, never complete closure. Resuming processing, we note that the next two phrases continue the serialization as well as the use of participial phrases, each modifying 'monsieur'. The final phrase 'ayant le teint blafard et l'apparence des plus bénignes' is simply the participialized relation 'avoir(nominal-1, nominal-2, {complement})', where nominal-2 is a compound of two nominal structures connected, on an equal basis, by the conjunction 'et'.

The net result of the entire text, which is finally completely closed by the terminator '.', is a simple situation, with an accumulation of attribute feature specifications, both embedded and serialized. For instance, if complementation were removed, the text would be simplified to:

"Alors on vit descendre un monsieur vêtu, chauve, portant toupet, ayant le teint blafard et l'apparence bénigne." Note that the missing details

have not altered the situation. Finally, DIG resolves that the situation is predominantly description with some narration involved.

4.1. Translation of Text 2, by G. Hopkins

This rendering of Flaubert's text is interesting because it uses a different sequencing of details and three situations.

> (5) "A gentleman with a pale face and a benign expression was seen to alight from the carriage. He had a bald forehead, a shock of hair at the back of his skull, and was wearing a short coat trimmed with silver braid." (Flaubert 1949:169)

Processing this text, we note that information is accumulated in a manner quite dissimilar from the French text. First, the participant is introduced and described: 'A gentleman with a pale face and a benign expression'. The functional role is not yet clear: SUBJECT? TOPIC? DO? When the verb is processed, it is marked as +passive, +past, ASPECT: +punctual. The 'gentleman' is now assigned the functional role of GRAMMATICAL SUBJECT, which is equivalent to DO when the verb is marked +process. The agent is not identified. Therefore, it is marked +indefinite and given the name 'someone' by default. This mirrors our assumption that if an event 'was witnessed or done', then it must have happened through the agency of someone or something. The infinitive phrase 'to alight from the carriage' is assigned the functional role of MODIFICATION: ADVERB: LOCATION: MANNER. Within it, the prepositional phrase 'from the carriage' is assigned the functional role MODIFICATION: ADVERB: LOCATION: ORIGIN. Finally, the terminator '.' triggers full closure and integration of the information accumulated so far.

The situation developed so far can be schematized as follows:

Situation-1:
PARTICIPANT: DOER:[1] 'a gentleman with a pale face and a benign expression'
RELATION:
1. see(nominal-1, nominal-2, {complement}) → see(someone, a gentleman, to alight(nominal-1, {complement}))
2. alight(gentleman, from the carriage)

Resuming input, we have the relation 'have(nominal-1, nominal-2, {complement})', where nominal-1 binds with 'He' which, in turn, is resolved to 'he = gentleman'. Thus, a second situation is begun and starts as:

Situation-2:
PARTICIPANT: DOER: 'he = gentleman';
RELATION: have(nominal-1, nominal-2, {complement}) → have(gentleman, nominal-2, {complement}).

The anaphoric link acts as a situational bridge between situation-1 and situation-2. This means that any information accumulated in situation-2 will apply seamlessly to situation-1. In reality, there is really only one hyper-situation. Once processing reaches the terminator '.', we have accumulated the following information:

> Situation constructed from the unification of Situation-1 and Situation-2, symbolized as Situation-1 U Situation-2, where 'U' symbolizes unification.

> PARTICIPANT: DOER: gentleman
> RELATIONS:
> 1. have(nominal-1, nominal-2, nominal-3, {complement}) → have (gentleman, a bald forehead, a shock of hair, at the back of his skull)

Finally, a third situation is created when 'wear(nominal-1, nominal-2, {complement})' is processed:

> **Situation-3**
> PARTICIPANT: DOER: gentleman
> RELATIONS: wear(nominal-1, nominal-2, {complement}) → wear (gentleman, a short coat, trimmed with silver braid).

Note that 'trimmed' is a reduction of the implied event 'trim(nominal-1, nominal-2, {complement}) + passive' where nominal-1 will be assigned the value 'someone'. Since this third situation is preceded by the separator ',', and since its INDEX values +sg, +3rd person are compatible with 'gentleman', situation-3 can be unified with situation-1 U situation-2 to yield:

> PARTICIPANT: DOER: gentleman
> RELATIONS:
> 1. see(nominal-1, nominal-2, {complement}) → see(someone, a gentleman,
> to alight(nominal-1, {complement}))
> 2. alight(gentleman, from the carriage)
> 3. have(nominal-1, nominal-2, nominal-3, {complement}) → have (gentleman, a bald forehead, a shock of hair, at the back of his skull)
> 4. wear(nominal-1, nominal-2, {complement}) + past, +durative → wear(gentleman, a short coat trimmed with silver braid).

It is interesting to note that the two texts differ in the manner of accumulation of information but that the final results are almost the same, except for the unobtrusive contrast of 'teint' versus 'face' and 'l'occiput' versus 'the back of his skull'. On the surface, it seems that one of the properties of situational information is that it is order-independent. That is, it does not matter how you assemble it, providing that the resulting constructions generate the

same or equivalent situations, be they embedded, serialized or unified. However, there may be differences if we extend our concept of information to include such cognitive psychological factors as tension, effect, interest, and logical sequencing. This situational equivalence mirrors a frequent scenario which occurs when someone is interrupted during a lengthy explanation and he or she quickly explains: 'Let me finish and you'll see that I'm saying the same thing!'.

4.2. Translation of Text 2, by J. Lewis May

May's text has been included primarily because it brings out more clearly the need for situational bridging than was the case with Hopkins.

> (6) "At this point a gentleman attired in a short coat with silver braid was observed to step out of the carriage. Bald in front, he had a tuft of hair at the back of his head. His complexion was sallow and his expression exceedingly benign." (Flaubert 1953:152)

May's first segment generates the following chunks of information: a temporal reference marking sequential order of events, a description of the first situational participant, an event marked +passive. Following the information integration triggered by the terminator '.', an initial situation is generated:

> **Situation-1**:
> PARTICIPANT: DOER (see endnote 1 below): a gentleman attired in a short coat with silver braid
> RELATION: observe(nominal-1, nominal-2, {complement}) +passive and nominal-1 = 'someone', nominal-2 = 'a gentleman attired in a short coat with silver braid'.
> COMPLEMENT: 'to step out of the carriage'.

A second situation is generated when the next segment is processed. The separator ',' following 'Bald in front' can indicate the start of a topicalized or possibly serialized functional role of MODIFICATION: ADJECTIVE: DESCRIPTION. At this point, it is not possible to know. This inability to assign a clear functional role simulates the hesitancy we would feel, if we heard someone saying to us: 'Bald in front, ...' As time elapsed, we would begin to wonder what the speaker's intention was. Thus a mild sort of tension would build up within us because normal communication happens at a surprisingly fast rate and brevity of expectation. The processing of 'he' indicates that the 'series' is over. This makes it possible to assign the functional role of MODIFICATION: ADJECTIVE: DESCRIPTION to 'he', which has already been resolved to 'gentleman' unless further information causes this anaphoric link to be blocked or changed. The remainder of the segment eventually yields 'have(nominal-1, nominal-2, {complement})', which becomes 'have(gentleman, a tuft of hair

at the back of his head)'. It is interesting to note that the final prepositional phrase 'at the back of his head' generates some ambiguity: a) is he holding/hiding a tuft of hair (not his own hair) behind his head or, b) does he have a tuft of hair (his own this time) growing at the back of his head? To a human processor, the first interpretation is highly unlikely, but to a formal system, either interpretation is possible since either is computable. Prepositional phrases often generate this type of potential ambiguity. The default solution is to resort to built-in common-sense knowledge. In this case a) hair is part of a head; b) a tuft of hair can be located at the back of the head. Therefore, the default interpretation is that the hair is his, unless there is some indication to the contrary.

This second situation, as with Hopkins, is linked with the first situation through the anaphoric link binding 'he' and 'gentleman'. Thus, all the attributes describing 'he' in this second segment unify with 'gentleman' in situation-1. The final segment contributes two more specified descriptive attributes to the 'gentleman': the 'sallow complexion' and the 'benign expression', as well as a third situation. At this point, the text ends. Since 'His' is anaphorically linked to 'he' and indirectly to 'gentleman', this third situation is also unified with the first and the second situations. Schematically, we may write situation-n = situation-1 U situation-2 U situation-3, which yields a single situation, as was the case with Hopkins's text.

5.0. Conclusions

What has been demonstrated is the fact that a formal, time-based left-to-right analysis of a piece of discourse results in a high degree of precision in representing the information contained in the discourse. Access to such meta-information makes it possible to verify and check textual apprehension, interpretation, comprehension and translation. In a fully-elaborated DIG analysis, the incremental and time-linear development of many pieces of information (semantic, structural, functional, pragmatic-contextual) can be explored thoroughly and used to delve deeply into the underlying forces that make two translations of a similar text so different in the information underlying the effects that they have upon the reader.

In spite of various small discrepancies, the two sets of texts have generated situational information which, respectively, agrees on situational structure and general content. It is interesting, however, to note some of the discrepancies which have cropped up. For instance, Flaubert has the gentleman coming down the steps, Hopkins concentrates on his stepping lightly down on the ground, while May either has the gentleman partly in and partly out of the carriage or making his way down a pathway of several steps connecting the carriage and the ground. Though the exact actions and location of our fictitious gentleman are an insignificant matter, this short analysis indicates a number of potential areas which are worthy of further exploration in the interface between the field of time-linear grammatical analysis

and the art of translation. As we have seen, different aspects of natural language understanding, translation and simulated information gathering provide many avenues for shedding light on the cognitive processes underlying the translation process and the results of those processes.

Finally, it is interesting to note that neither translation captures the smoothness or cohesion of the original text by Flaubert. Perhaps this 'feeling' results because the situational bridging entailed by both translations was not triggered by the monosituational original. If this observation proves to be correct, we will have achieved one more step toward concretizing our ability to understand a small part of the black-box phenomenon that is natural language understanding. There seems little doubt that the time-linear analysis of the comparative understanding of translations will prove valuable in the development and testing of cognitive pragmatic linguistic theories.

Bibliography

Barwise, Jon & John Perry (1983). *Situation Semantics*. Stanford, Ca.: CSLI Publications.

Carpenter, Bob (1992). *The logic of typed feature structures*. Cambridge: Cambridge University Press.

Chomsky, Noam (1995). *The Minimalist Program.* Cambridge, Ma.: MIT Press.

Chomsky, Noam (2000). *New Horizons in the Study of Language and Mind.* Cambridge: Cambridge University Press.

Flaubert, Gustave (1949). *Madame Bovary*. (tr. G. Hopkins). Oxford: Oxford University Press.

Flaubert, Gustave (1953). *Madame Bovary*. (tr. J. Lewis May). London: Collins.

Flaubert, Gustave (1857/1966). *Madame Bovary*. Paris: Garnier Flammarion.

Hausser, Roland (2001). *Introduction to Computational Linguistics*. Berlin: Springer Verlag.

Kempson, Ruth, Meyer-Viol, Wilfried & Dov Gabbay (2001). *Dynamic Syntax*. Oxford: Blackwell.

Sévigny, Alexandre (2000). *Lexically-Driven Incremental Discourse Assembly*. PhD. thesis. University of Toronto.

Sévigny, Alexandre (2002). "Discourse Information Grammar." To appear in *Linguistic Association of Korea Journal* 10(4).

Sévigny, Alexandre (forthcoming). "Fundamentals for a procedural theory of linear discourse grammar." To appear in *Canadian Journal of Applied Linguistics* 5(1).

Shieber, Stuart (1986). *An introduction to unification-based approaches to grammar.* Stanford: CSLI Publications.

Shin, Gyonggu (1987). *Linearized Phrase Structure Grammar.* Ph.D. thesis. Chonnam National University, South Korea.

Sperber, Dan & Deirdre Wilson (1986/1995). *Relevance: communication and cognition.* London: Blackwell.

---

[1] Because the situation is marked +PASSIVE, the DOER in this case is equivalent to DO in an equivalent situation marked +ACTIVE. Also to be noted, though not relevant to this particular example, is the need to change the infinitive to a participial

form in the active situation as well as the need to furnish an active DOER such as 'someone' or 'somebody' or 'they' or 'people' in a default indefinite situation where the DOER has not been explicitly indicated. These rules concerning changes triggered by voice changes are part of DIG's structural database knowledge, meant to simulate language-specific patterns which humans learn as they acquire complete mastery of a given language. Additionally, the stylistic level could be captured using a flag marker, say, +common for the utterance 'A gentleman was seen alighting from a carriage ...' and +literary for 'A gentleman was seen to alight from a carriage ...' since the latter pattern is less commonly heard in ordinary speech.