

## Overviews

# World knowledge in discourse comprehension

Teruo UENO

**Abstract :** Discourse comprehension involves not only the interpretation of linguistic signs but the utilization of world knowledge, which could be summarized as the knowledge about the facts surrounding us and commonsense reasoning that relates them.

In order to access and utilize world knowledge for efficient text comprehension, world knowledge itself has to be systematically organized according to some cognitive semantic structures. In this paper, frame theories and schema theories are discussed in reference to the issue of how world knowledge is organized in order to be used in text comprehension, as well as their implications for foreign language teaching.

**Key Words :** world knowledge, discourse comprehension, frame, schema, foreign language teaching

### I. World knowledge and inference

Whether on-line or not, we are constantly making inferences in our daily life. As Ueno (1999) states, we make inferences in order to identify and/or comprehend what is perceived through our senses, to speculate on the cause of the thing perceived, to make predictions about future events, or to understand or create coherent text (written or spoken). It is not very hard to imagine that the inferences thus made reflect our knowledge of the world acquired through our past experience. In other words, our world knowledge based on our experience influences the inferences that we make, hence our perception and/or interpretation of the world we live in. Wason's (1966, 1968) four-card selection task well illustrates this point.

The world knowledge itself may vary individually from person to person, or socially/culturally from community to community. However, there must be some basic principles common to all individuals that govern the organization and utilization of such knowledge in interpreting or creating discourse text.

The theories that study the relationship between socio-cultural knowledge and language can be divided into two groups. The first group, which originates in the field of

anthropology and is called ethnography of communication, sees socio-cultural knowledge “as revealed in the performance of speech acts bounded in real time and space, and characterized by culturally specific values and norms that constrain both the form and the content of what is said” (Gumperz, 1982: 154). In this view, therefore, interpretation of what is implied in what is said is not a major concern. On the other hand, the second group focuses on how the world knowledge is applied to the interpretation of text (or, in a broader sense, to the interpretation of the world surrounding the interpreter). World knowledge has to be organized and used according to some cognitive semantic structures. These cognitive structures are called *frames*, *scripts*, *scenarios*, or *schemata*, depending on the theories.

In this paper, I would like to focus on some of the theories that address the issue of how world knowledge is organized to be used in text comprehension/creation, and their implications for foreign language teaching.

### II. Frames

The term *frame* occurs in many fields of science. Tannen (1979: 141) points out that it is used, “in the work of Bateson and Frake in anthropology, Hymes and Goffman in sociology, Minsky in artificial intelligence, and Fillmore in linguistics.” Further, in analyzing discourse, Clark and Clark (1977: 34) use the term to refer to the first phrase of a sentence to make

a functional distinction with the remainder of the sentence (*insert*).<sup>1</sup> First, I would like to examine Minsky's (1974/1996, 1977, 1986/1996) notion of frame because it directly concerns the organization and representation of world knowledge,<sup>2</sup> followed by Tannen's notion of frame, which relates to the studies in anthropology, sociology, and linguistics.

## II-1. Minsky's notion of frame

Minsky describes a frame as a "data-structure for representing a stereotyped situation" and summarizes its essence as follows:

When one encounters a new situation (or makes a substantial change in one's view of the present problem) one selects from memory a structure called a Frame. This is a remembered framework to be adapted to fit reality by changing details as necessary (Minsky, 1974/1996: 997).

A frame consists of two parts, fixed part and changeable part. The former represents the things that are always true about the supposed situation. The latter has many "slots" called *terminals*, where the specific values of a given instance of an entity notion or situation are stored according to assignment conditions that each slot specifies. For example, if we are in a kitchen in a house that we visited for the first time, we store information of its size, colors of walls, floor, and ceiling, layout, electric appliances it has, etc. in terminals. If the value for a certain terminal cannot be determined, (i.e., if the information necessary to determine the value is simply missing), then the terminal is given a default setting (or assumption) as its typical value. Thus, even if we do not see someone cooking in the kitchen, we can assume that the room is for cooking thanks to the default value of the "action" terminal of the kitchen frame.<sup>3</sup> Similarly, even if we do not see kitchen utensils, we presume that they are there in the kitchen. In this sense, as Johnson-Laird and Wason (1977)

point out, the notion of a frame approximates that of Rosch's "prototype", since a prototype could be considered as a frame whose terminals for various variables are filled in with all the default values.

An important point in this system is that different frames share the same terminals. That is why a scene of the kitchen from one particular point of place and that from another, each of which making a different frame, can represent the effects of moving from one place to another. Further, even when the frames are non-visual type, "cause-effect relations, or changes in conceptual viewpoints" can be represented by the differences between frames (Minsky, 1974/1996: 998).

Although originally developed as a theory of visual perception and visual memory, Minsky's notion of frame extends its application to language comprehension. To illustrate this point, let us take a simple example from Minsky (1986/1996: 736).

- (1) Mary was invited to Jack's party.
- (2) She wondered if he would like a kite.

Having read (1) and (2) in this order, we can infer that probably the party was Jack's birthday party, and that the kite was a birthday present from Mary. We can also infer that both Mary and Jack are children or (at least Jack is a child). Such interpretation would not be possible just by the identification of anaphoric referents of *she* and *he* in (2). According to Minsky's frame theory, (1) arouses a "party-invitation" frame in the reader's mind. Its terminals are attached to certain memories of various concerns such as "Who is the host?", "What present should I bring?", and "What should I wear?" Each of these concerns is represented by a subframe with its terminals filled with the most usual solutions to that particular problem. The most usual solutions are based on our previous experience, that is assigned, in advance, as default values to the corresponding terminals.

As seen in the example above, Minsky sees the activation of frames and filling of their terminals as a key to the comprehension of text. The question here is whether we

<sup>1</sup> Clark and Clark (1977) explain the function of *frame* as to orient the listeners "toward a particular area of knowledge—to give them a point of departure for the sentence." Therefore, the notion of *frame/insert* substantially coincides with that of *theme/rheme*, and more importantly with the notion of *starting point*, a linguistic unit from the viewpoint of the serial word order (i.e., linearisation).

<sup>2</sup> Schank and Abelson (1977) propose the notion *script*, which is similar to the notion of *schema* or *frame* but especially deals with event sequences.

<sup>3</sup> Since the theory allows frames of other entities to be *fillers* for the terminals, the action of "cooking" and kitchen utensils such as a colander or an eggbeater themselves form their own frames (i.e., subframes).

really have to activate a frame for each and every thing we see, hear, or read in order to comprehend it. Brown and Yule (1983) raise the same question, citing the following sentence from *The Sunday Times*.

- (3) The Cathedral congregation had watched television monitors as Pope and Archbishop met, in front of a British Caledonian helicopter, on the dewy grass of a Canterbury recreation ground (Brown and Yule, 1983: 240).

The question here is whether a Cathedral frame, a television-watching frame, a meeting frame, a helicopter frame, and a recreation-ground frame are all activated by the reader to comprehend this sentence without creating any mess in his/her short-term memory. If only one/some of them is/are activated, which ones will it/they be? Brown and Yule suggest that this is a major problem with Minsky's frame theory. Examining Minsky's explanation carefully, however, we could say that their criticism might be a little misleading.

It is true that Minsky emphasizes the importance of activation of various frames to understand discourse properly, but we have to be aware that such activation occurs in order to enable the comprehender to access stored data or knowledge called *common sense* (i.e., default assumptions that fill the relevant terminals) about a given object (concrete or abstract) or situation. (Remember that most terminals are already filled with default assumptions.) These excited (or "aroused" in Minsky's term) frames lay a ground for the comprehender to process the next sentence. In the example above, (1) activates a party-invitation frame together with other related frames and subframes, including such linguistic frames as a noun frame, a verb frame, and a sentence frame. Now, common sense related to (i.e., filled in the terminals of) these activated frames is ready to recognize possible references to presents, clothes, the protagonist, etc. in (2).

We have to remind ourselves of Minsky's definition of a frame. He states that a frame is a data-structure for storing and representing a stereotyped situation, and that it is selected from memory when one encounters a new situation or makes a substantial change in one's view of the present problem. In other words, if there is no change in the situation or no change in one's view, there will be no selection of a frame. It is selected only when it is necessary to adapt one's common sense acquired from previous experience to the reality that he

is experiencing now. Modification to one's common sense is done by filling some of the terminals with new data.

Here, no selection does not mean no activation of frames. Various frames could be activated, but the activation is done only to facilitate comprehension by making common sense ready to provide proper context for the text. Still, if there is nothing that would threaten the validity of the existing common sense, no selection occurs because it is not necessary to modify it. I argue, therefore, that criticism against Minsky's frame theory resulted from confusion between selection and activation (arousal, or excitement) of a frame is invalid. It would not be too much to say that the notion of *frame* as a data-structure for representing a stereotyped situation has provided helpful hints on how new situations are dealt with, and how new data are stored. As such, it has also benefited linguistics in the analysis of text comprehension. For example, Fillmore's (1976, 1985) frame semantics emphasizes the importance of reference to a set of culturally familiar frames in understanding meaning.

Tannen (1979, 1986, 1990) also uses the term "frame" in relation to *metamessages* which is implicitly conveyed by words or actions to refer to the relations among the people involved, and their attitudes toward what they are saying or doing and toward the people they are saying or doing it to. Although Tannen's notion of frame (together with the term *metamessages*) is from an anthropologist Gregory Bateson and is not directly related to Minsky's frame theory, both share a basic concept about a function of world knowledge in language comprehension: Discourse cannot be successfully understood without reference to culturally constrained assumptions about the world (i.e., *common sense* shared by the people within a community).

Regardless of the similarities in names and basic concepts between these two notions of frame, it would be necessary to distinguish them here. In order to avoid a possible confusion between the two notions, I would like to touch upon Tannen's notion of frame or *framing* in the next section.

## II-2. Tannen's notion of frame and framing

Tannen explains the concept of frame as "a way of showing how we mean what we say or do and figuring out how others mean what they say or do (Tannen, 1986: 74-75)." Our utterances are framed by conversational signals (pacing

and pausing, loudness, pitch and intonation), conversational devices (expressive reaction, asking questions, complaining, apologizing), and non-verbal signals (facial expression, gestures) so that they carry some specific metamesages, such as *this is a joke, I am serious, this is a warning, or I am more competent than you are.*

For example, the sentence “I am sorry” uttered loudly with angry facial expressions might be good enough to mean something like “I am saying this not because I feel sorry or I want to apologize. I am saying this to satisfy you so that you will shut up.” Further, Tannen points out that framing may reveal what position the speaker is assuming, and what position he/she assigns to the listener. By putting the listener down, the speaker is taking a superior position (or “alignment” in sociologist Erving Goffman’s term). Tannen’s (1990) example of a driver waving his hand to another driver to yield the right-of-way illustrates this point. Instead of taking this action as kindness, the second driver interprets it as “I am granting you my own right to go ahead.” Consequently, he turns down the offer and waves back to the first driver. Here, the second driver sensed the first driver’s higher position framed by the first and reframed the relative position as “No, I am higher,” or at least, “No, we are on the equal position” by rejecting the offer and waving the first driver on.

Reactions to the frames set by others (i.e., whether to resist or to accept them) may reflect gender difference as Tannen suggests. Or, they might reflect individual habits, the person’s state of mind on that particular day, and/or his/her cultural background. If interpretation of frames and reaction to them vary individually or socially, there is always a possibility of miscommunication. In other words, since conversational signals, conversational devices, and non-verbal signals frame our utterances and actions, different expectations about these communicative signals between interlocutors may make communication unsuccessful. It is on this point that Tannen’s notion of frame and Minsky’s agree. Whether the different expectations are based on gender difference as Tannen (1986, 1990) emphasizes or cultural difference, communication will fail if we are unaware or ignorant of such differences in expectations, or if our world knowledge does

not contain the relevant information on them.

Now, going back to the example of the driver who declined the offer of the right-of-way from another driver, why did he think that the first driver was assuming a higher position? What is his reaction based on? Would that be gender, culture, or just his personality? Tannen’s answer consists in two important concepts, *involvement* and *independence*.

Tannen (1986) states that everyone has two conflicting needs in terms of relationship to others; the needs for involvement (i.e., the needs to be connected to others) on one hand, and the needs for independence (i.e., the needs to be left alone) on the other.<sup>4</sup> She claims that generally men feel the needs for independence, while women feel the needs for involvement. Here, supposing that the second driver is a man, he may have taken the offer from (or a frame set by) the first driver as a threat to his independence. To him, accepting the frame was to admit that he was not independent, or that he was subordinate to the first driver. If the second driver was a woman, she might have taken the offer without reframing the frame set by the first driver.

Tannen explains that many instances of miscommunication between men and women are derived from this gender difference in attitudes toward others. Accordingly, she claims that a key factor to successful communication between men and women is to know these conflicting needs of the other gender and to try to balance them within themselves with the help of conversational signals and devices.

This suggests that if communicative signals are not available, communication might fail even between the two persons of the same gender. We can easily find many instances of misunderstanding in electronic mail exchange if we subscribe to some mail lists on the Internet, which distribute every message posted to all their members. In e-mail, as well as in any other type of writing, we cannot use conversational signals as in the face-to-face conversation. The only thing we could do is invent alternative ways of signaling our metamesages. We could use bold face or upper case letters to mean that we are emphasizing the words, (though upper case words usually mean *I am shouting!* in an e-mail

---

4 Gray (1992) makes a similar point, expressing the conflicting needs within men as the needs for “autonomy” and the needs for “intimacy”. According to him, men are balancing them by pulling away to achieve autonomy and getting close for intimacy just like a “rubber band”.

message.) A simplified illustration of a smiling face called Smiley might be useful from time to time.

However, these devices that mimic conversational signals are very much limited both in number and in function. To make matters worse, e-mail messages are written quickly without giving much attention to wording, not to mention rhetoric. The result is, as Levine and Baroudi (1993) warn, that e-mail always seems ruder than it is supposed to. We may sometimes even feel that we are “flamed”. This happens because we cannot detect a frame that the writer set for the message. Even if we have invented a large number of symbols to label frames, they would not be of much use if the receiver of a message does not know the conventions that we are trying to establish. It parallels the situation where we do not have adequate information about cultural differences in the use of communicative signals. We have to say, once again, that in order to comprehend communicative signals correctly, our world knowledge reservoir has to contain relevant information to decode them. If not, communication could fail.

Likewise, lack of relevant information about different cultures could lead to misinterpretation of the speaker’s intention. Tannen gives an example of a Japanese woman talking about her husband’s recent death with a smile (not with “laughs” as Tannen states, though.) Expressing too much emotion might offend the listener because it could be taken as imposition of one’s emotion on others (, which Tannen might call a threat or a challenge to the listener’s independence). Indeed, this might even be a universal principle across different languages and cultures. However, in Japan, the principle goes to the extreme: Woman may smile while talking about her husband’s death. As Tannen explains, the woman might be misdiagnosed by a Westerner who does not know that laughter (or smile) is “the customary and expected Japanese way of masking emotions (Tannen, 1986: 41).” A smiling face when feeling intolerable sadness deep inside would be totally inappropriate in other cultures. In Japan, however, this is not an uncommon strategy for avoiding making a conversational partner feel uncomfortable. Death of a person is certainly not a pleasing topic to anybody. In order to cancel out the negative or unpleasant connotation of the

topic, the speaker tries to use whatever strategies available. To most Japanese people, a smiling face is a tool to set a frame to show their solicitude for the listener’s feelings. With such knowledge, we can correctly interpret the woman’s smile as “I do not want you to feel bad like I am feeling now.”<sup>5</sup>

To sum, if socially-culturally conventionalized norms about communicative signals and their meanings are not shared by the interlocutors, or in Gumperz’s (1982) terms, if conventionalized *co-occurrence expectations* between content and surface style (i.e., *contextualization cues*) are not shared by them, there would be a mismatch between conversational inferences that the speaker has in mind and those that the listener makes, and consequently communication would fail. In the next section, we will see some more evidence for the relations between socio-cultural knowledge and discourse comprehension.

### III. Schema theory

Similar to the notion of frame is that of *schema* originally put forth by Bartlett (1932). Having English speaking subjects recall the stories of Native Americans, Bartlett (1932) found that the subjects not only had difficulty remembering them but even reorganized them. Based on these findings, he argued that recalling a bizarre story was poor because the subjects tried to alter its details so that it fit into their existing *schemata*. Referring to this study, the notion of schema is generally explained as “a structure in semantic memory that specifies the general or expected arrangement of a body of information” (Carroll, 1986: 231). As for a story schema, it may be culture-specific as shown by Bartlett, or may be genre-specific. For example, in reading an article in a newspaper, we expect that the headline provides the information on what the article is about, that the lead summarizes the incident or accident, and that its detailed account will follow later in the article. In reading detective stories, we do not expect a suspect will be caught early in the story. Such expectations are possible because we have a semantic structure or knowledge structure called *schema* that we have made based on our previous encounters with many types of discourse.

Bartlett explains that the past experience works as a

---

5 Culturally imposed meaning on this type of strategy has to do with the concept of politeness shared by the Japanese. The concern about listeners’ feelings may have been the origin of honorifics in the Japanese language.

whole, not as an each individual instance; “the past operates as an organized mass rather than as a group of elements each of which retains its specific character (1932: 197).” In other words, schemata are generalized assumptions (about organization of a body of information) abstracted from the aggregate of each instance that we have encountered in the past.<sup>6</sup> Therefore, by definition, a schema may also undergo modification according to our new experience. Here, its similarity to Minsky’s frame is obvious.

Considering the fact that various schemata are constructed and modified not only through our personal experience but also by the norms of a society in which we live, there is no doubt that some schemata will be culture specific. Using R.N.Ross’s term “structures of expectations,” Tannen (1979) also treats schemata as structures of expectations organized according to “one’s experience of the world in a given culture (or combination of cultures)”. The organized knowledge about the world is, in turn, used to “predict interpretations and relationships regarding new information, events, and experiences (1979: 138-139)”.

A schema may reveal itself as a prejudice against or a bias for something or somebody, influencing our interpretation of stories, movies (including interpretation of physical movements of something/somebody which are not connected from one scene to another, as continuous motions), or virtually anything that exists in this world as long as we have encountered them enough times to create a certain schema. If it really exists, a schema will surely help us infer the things that are not explicitly said or written in text, what is going to happen next, and the things that are not perceived by our senses.

In the next section, limiting our topic to discourse comprehension, I would like to look at some of the schema theories on discourse structure.

### III-1. Schema and discourse structure

Among several models that have been proposed to deal with how discourse is structured, Labov & Waletzky’s (1967) study is perhaps worth mentioning first here for two reasons: (1) The data analyzed were actual narratives told by the subjects (i.e., they were not created by the researchers to serve for a specific study), and (2) the data were collected from a large number of subjects who widely ranged in age.

Seeking for basic structures of complex narratives, Labov & Waletzky (1967) analyzed the tape-recorded data of oral narratives of personal experiences gathered from about 600 interviews. The English speaking subjects of 10 to 72 years old were drawn as narrators from both black and white communities of both rural and urban areas in the U.S.

Although narrative units recapitulate experience of series of events, the order of those units in a narrative does not always correspond to the temporal sequence of the original events. Hence, focusing their attention on the types of clauses and their orders in narratives, Labov & Waletzky attempted to find some organizational principles according to which narratives were constructed.

First, they found that some clauses could be placed anywhere in the narrative without changing the inferred sequence of events in the original semantic interpretation (hence, named *free clause*), and some were restricted to some degree in terms of possible positions they could take (hence, named *restricted clause*). Further, those clauses that were

---

6 Schemata in this sense corresponds to Oller’s (1995) *formal schemata*. In relation to Peirce’s three modes of inference, Oller classifies schemata into three kinds, *content schemata*, *formal schemata*, and *abstract shemata*. Content schemata, which are based on abduction, concern particular arrangements or relationships of things in the world that we perceive in a particular context of experience. Formal schemata, which are mainly based on induction, are concerned with generalized assumptions about arrangements or relationships of the things that we have abstracted to some degree from the particular facts indexed as being similar. In Abstract Schemata, whose concept parallels deduction, the assumptions obtain the complete generality of pure symbols.

For example, the number of classrooms, names of students, cultural background of each student, and the rules of a particular school are all part of content schemata. If some similar facts are observed at many elementary schools to the extent that we can generalize them as applicable to all other schools, the generalized assumptions are now part of formal schemata. On the other hand, when we hear the words ‘elementary school’, we all can infer that there must be classrooms where teachers teach children of a certain age range. The concept of ‘elementary school’ as a place for education is embedded in the meaning or definition of the words itself. The definition of the words (i.e., symbols) ‘elementary school’ is not dependent on the similarities observable among many elementary schools. Rather, it is integrated in our abstract schemata. With this classification, Oller examines the results of recent empirical studies to show how these schemata influence discourse processing. He claims that abstract schemata have the greatest impact of the three, and that formal schemata have greater impact than content schemata. I will refer to some of the studies in Section III-3.

strictly fixed in position in the sequence were named *narrative clause*.

By observing the distribution of clauses over the entire narrative, Labov & Waletzky found that four basic components were operative in narratives; that is, orientation, complication (complicating action), evaluation, and resolution. The orientation section provides background information such as person, place, time, and behavioral situation. It is mainly composed of free clauses that precede the first narrative clause. The complication section consists of series of events and is represented by the main body of narrative clauses whose order in the story corresponds to that of actual events. A long string of events may consist of several cycles of simple narrative, with many complication sections, each of which is normally terminated by a result.

In the evaluation, narrators express their attitudes, personal feelings or judgments towards the events by emphasizing the relative importance of some narrative units as compared to others. The section also serves as a marker to separate the complication from the following resolution. Groups of free or restricted clauses are often found in this section, but a single narrative clause may states both the result and its importance. A narrative may have more than one evaluation, each of which is followed by the resolution, the fourth component. When this happens, a narrative becomes more complex in its structure than the primary sequence.

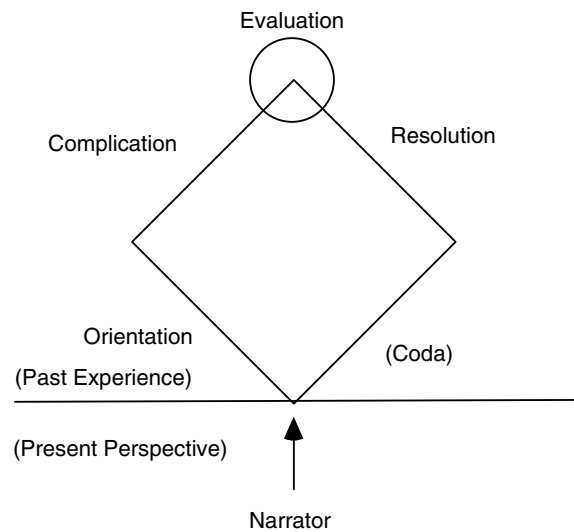
The resolution section, following the evaluation, mainly consists of narrative clauses, and shows the result or outcome of events. When the evaluation is the last element, the resolution may coincide with the evaluation. Further, an optional section called coda may appear to return the verbal perspective to the present moment.<sup>7</sup>

Identifying each narrative component as the above, Labov & Waletzky represented the standard structure of narratives in the diagram as in Figure 4 below.

The significance of this study lies in the fact that Labov

Figure 4.

Adapted from Labov & Waletzky (1967: 41)



& Waletzky showed the relationship between the sequence of clauses in a narrative and the sequence of events inferred from the narrative. If there were no mechanism relating the two, we would not be able to reconstruct the sequence of events in the original order when we try to comprehend a narrative. To see the narrative structure in terms of the arrangement of its components is compatible with the theories that view the discourse structure (or macrostructure) in terms of a schema (Mandler and Johnson, 1977; Rumelhart, 1975, 1977; Schank, 1976; Thorndyke, 1977).

For example, Thorndyke (1977) claims that narrative discourse is structured with several components which are hierarchically organized according to a set of rules like phrase-structure rules.<sup>8</sup> He further tried to show the psychological validity of this structure by claiming that in recall and summarization tests, the components at the top of the hierarchy were recalled most easily and included in the summaries.

Here, we have to note that classification of clauses into such categories as those specified in the rules (See footnote

<sup>7</sup> Labov & Waletzky list three main devices for this function. (1) Use of deixis: By using the words such as “this” or “here”, for example, we can stand at the present moment of time, and point to the end of the narrative with “that” or “there”. (2) Use of an incident in which one of the actors can be followed up to the present moment in actions. (3) Extension of the effect of the narrative on the narrator to the present moment.

<sup>8</sup> The following is an extract from Thorndyke’s rules (Thorndyke, 1977: 79).

1. Story-> Setting + Theme + Plot + Resolution, 2. Setting-> Characters + Location + Time, 3. Theme-> (Event)\* + Goal, 4. Plot-> Episode\*, 5. Episode-> Subgoal + Attempt\* + Outcome, 6. Attempt->Event\*/Episode, 7. Outcome-> Event\*/State, 8. Resolution-> Event/State, 9. Subgoal/Goal-> Desired State, 10. Characters/Location/Time-> State

8) relies on reader/listener's subjective judgment about the meaning and the function of each clause in the entire discourse. We can label them because we have comprehended the story as a whole. It should not be the other way around: We comprehend the story because we have successfully labeled the units. It is true that a tree structure derived from the rules of a story grammar is helpful to capture the overall structure of a story. However, as for the details in the structure (e.g., the structure of the components at the low level of the hierarchy), it would be difficult to prove its validity. Actually, the overall picture of a tree structure could easily get much more complicated than the surface form of the discourse. Would it not be good enough or more practical to state that a story is generally composed of Setting, Theme, Plot, and Resolution, or of Orientation, Complication, Evaluation (including its cyclic function to transform the primary sequence into a more complex normal form of a narrative), and of Resolution as Labov & Waletzky argued? It may be only these types of basic components that guide our discourse processing.

### III-2. van Dijk's macrostructure, and superstructure

In an attempt to formulate the formal means of identifying a topic for discourse, van Dijk (1977) also tried to analyze the structure of discourse. His more recent work on the analysis of the structure of news articles is an example of application of his theory. Van Dijk's (1988a, 1988b) analysis is based on his notions of Macrostructure, Microstructure, and Superstructure (van Dijk, 1980). Macrostructure refers to a semantic global structure, whose functions are (1) to organize complex information as a coherent whole, and (2) to reduce complex information to relevant, abstract, or general information (i.e., topics or themes represented by macropropositions) by way of three reduction rules called macrorules<sup>9</sup>. Microstructure is a semantic detailed structure at a more local level, such as the meanings of words, phrases, clauses, sentences, and connections between sentences. Superstructure is a conventional schematic global structure of a discourse. In discourse analysis, it may be characterized

as a syntactic skeleton of text. (Superstructure is also called *supersyntax* by van Dijk (1988b). It could be understood as what traditionally called *schema*).

Borrowing Labov's notion of narrative units, van Dijk claims that when a discourse of any type is organized according to a conventional schema, its structure is realized as the ordering of functional categories, such as setting, complication, resolution, etc., which act as sort of slots to be filled with the content of the text, namely sequences of propositions (*macropropositions*). Consequently, superstructures impose some constraints on macrostructures by specifying what kind of information is appropriate or needed in each slot.

With this theoretical framework, van Dijk (1988a) analyzes 729 articles from 138 newspapers in 99 countries. Although the expected organizational (schematic) differences supposedly rooted in cultural, economical, and political differences between the First World and the Third World press were not found in this study (probably because most of the Third World press use materials from the major news agencies in the First World such as the AP or the Reuters and literally copy the news text for their own papers), we may be able to look at it as the evidence that the news press all over the world (and probably the readers as well) are following "globally shared code of journalistic practices".

As a consumer of the news articles, we consciously utilize our knowledge of the structure of news articles. When we do not need detailed information on some news, we can just read the Headline and Lead which are placed at the very beginning of the article. From the position of a text on the page and its type face, we can also infer whether or not it is a major news of the previous day. This is certainly not the way we read other types of text, for example, empirical research papers.

Van Dijk's approach to discourse analysis seems to be parallel to the approach of traditional transformational generative grammar for the analysis of a sentence. In this context, his claim that superstructures and macrostructures serve as underlying structures of sequences of sentences in

---

<sup>9</sup> The rules concern (1) deletion of all information that is not relevant in the rest of the text, (2) generalization of a sequence of propositions by a single proposition, (e.g., "We have pets," generalized from "We have a dog, a cat, and a canary."), and (3) construction of overall event out of its constituent details, (e.g., "I took a plane to ...," constructed from "I went to the airport, checked in, walked to the gate, and got on a plane.")



texts, and his proposal of transformation-like rules for the organization of discourse categories (e.g., rules concerning ordering, deletion, or combination of categories), are understandable.

However, the more details the theory try to explain, the stronger the type of criticism I mentioned at the end of the previous section will be. In addition, it is not clear as to how many superstructure types should be considered if, as he claims, functional categories are specific for each type of discourse, and exactly what type of transformation rules are necessary in each type. But again, the clearer are these points made, the more ad hoc the analysis may seem (See Ueno, 2003, for more detailed discussion on Van Dijk's theory).

#### **IV. Schema and its implications in foreign language teaching**

As stated before, some schemata are culture specific. Accordingly, we can imagine that those culture-specific schemata will have some effects on second/foreign language learning/teaching. In this section, I would like to discuss some of the studies that provide the evidence of the effects of culture-specific schemata on the processing of discourse, and also refer to what these studies imply for successful second/foreign language learning/teaching.

Chihara, Sakurai, and Oller (1989) examined how minor changes in textual elements to conform to the cultural background of the subjects (Japanese female junior college students) would affect their performance on cloze tests. With a few modifications to the text (e.g., names and actions of the characters), the subjects scored significantly higher than with unmodified versions.

The findings of Chihara et al. (1989) were replicated by Al-Fallay (1994) with Arab students learning English in Saudi Arabia. Two stories, (English translation of one originally Arabic narrative and one originally English narrative), were used in the study. By changing the names of persons, places, and events, the translation version was modified to fit American culture, resulting in two versions (modified and unmodified) of English translation of the Arabic story. Likewise, the English story was modified to conform to Arabic context, resulting in two versions (modified and unmodified) of the originally English story. On the cloze tests administered five times throughout the quarter, the subjects

who were given the texts in Arabic context scored higher than those who received the texts in non-Arabic context. Moreover, by comparing pretest and posttest scores, Al-Fallay reports that by taking cloze tests in culturally familiar context over a period time, the subjects in the experimental group benefited more (i.e., learned more) than those in the control group.

What the findings of these studies suggest has a significant meaning to language teachers. In preparing teaching materials or assessing the language skills of their students, second/foreign language teachers cannot ignore the impact of such culture-specific schemata. Introducing new grammatical rules or words in a discourse which is built up on the schemata culturally different from those of students may not be as effective as expected. Such materials could be an extra burden on the part of the students; that is, they have to learn grammatical rules and words that are embedded in the materials that they can not easily understand and remember. At the same time, second/foreign language learners should be aware of the fact that knowledge of vocabulary and grammatical rules of a target language is not good enough to interpret a discourse constructed on culturally different schemata. Students' efforts may not be greatly rewarded unless cultural aspects of the language, such as different expectations (or, in Minsky's term, different default settings for terminals in a frame) about people's actions and reactions in interaction with others, narrative structure, non-verbal communicative tools, etc. are learned by the students themselves. What schema-based theories imply to second/foreign language teachers and students should not be made light of. After all, it concerns their ultimate objective; to learn the language successfully on the learners' part and to enhance effectively the learners' fluency in the language on the teachers' part, without giving unnecessary burdens to the learners.

## REFERENCES

- Al-Fallay, I. (1994). *Limiting bias in the assessment of English as a foreign language: The impact of background knowledge on the proficiency of Saudi Arabian students learning English as a foreign language*. Unpublished doctoral dissertation, University of New Mexico, Albuquerque, New Mexico.
- Brown, G. & Yule, G. (1983). *Discourse analysis*. Cambridge: Cambridge University Press.
- Carroll, D.W. (1986). *Psychology of language*. Monterey, CA: Brooks/Cole.
- Chihara, T., Sakurai, T. & Oller, J.W.Jr. (1989). Background and culture as factors in EFL reading comprehension. *Language Testing*, 6(2), 143-151.
- Clark, H.H., & Clark, E.V. (1977). *Psychology and language*. New York: Harcourt Brace Jovanovich.
- Fillmore, C.J. (1976). The need of a frame semantics within linguistics. *Statistical methods in linguistics*, 5-29.
- Gray, J. (1992). Men are from mars, women are from venus: A practical guide for improving communication and getting what you want in your relationships. New York, NY: Harper Collins.
- Gumperz, J.J. (1982). *Discourse strategies*. Cambridge: Cambridge University Press.
- Johnson-Laird, P.N., and Wason, P.C. (1977). *Thinking: Readings in Cognitive Science*. Cambridge: Cambridge
- Labov, W., & Waletzky, J. (1967). Narrative analysis: Oral versions of personal experience. In J.Helms (Ed.), *Essays on the verbal and visual arts* (pp. 12-44). Seattle, WA: University of Washington Press.
- Levine, John R., & Baroudi, C. (1993). *Internet for dummies*. San Mateo, California: IDG BooksWorldwide, Inc.
- Mandler, J.M., & Johnson, N.S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Science*, 9, 111-151.
- Minsky, M. (1974). Artificial Intelligence: A framework for representing knowledge. *Artificial Intelligence Memo*, 306. (Reprinted in CD-Rom format, Shore & Marion (Ed.) *First person: Marvin Minsky* (pp. 994-1184). S.Burlington, VT: Voyager, 1996.)
- Minsky, M. (1977). Frame-system theory. In P. N. Johnson-Laird & P.C. Wason (Eds.), *Thinking: Readings in cognitive science* (pp. 355-376). Cambridge: Cambridge University Press.
- Minsky, M.L. (1986). *The society of mind*. New York: Simon & Schuster. (Reprinted in CD-Rom format, Shore & Marion (Ed.) *First person: Marvin Minsky*. S.Burlington, VT: Voyager, 1996.)
- Oller, J.W., Jr. (1995). Adding abstract to formal and content schemata: Results of recent work in Peircean semiotics. *Applied Linguistics*, vol.16, No.3, 273-306.
- Rumelhart, D.E. (1975). Notes on a schema for stories. In D.Bobrow & A.Collins (Eds.), *Representation and understanding: Studies in cognitive science* (pp. 211-235). New York: Academic Press.
- Schank, R. C. (1976). The role of memory in language processing. In C. N. Cofer (eEd.), *The structure of human memory* (pp. 162-189). Freeman: San Francisco.
- Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*. Hillsdale, NJ: Lawrence Erlbaum.
- Tannen, D. (1979). What's in a frame? Surface evidence for underlying expectations. In R. O. Freedle (Ed.), *New Directions in Discourse Processing* (pp. 137-181). Norwood, New Jersey: Ablex.
- Tannen, D. (1986). *That's not what I meant: How conversational style makes or breaks relationships*. New York: William Morrow.
- Tannen, D. (1990). *You just don't understand: Women and men in conversation*. New York: William Morrow.
- Thorndyke, P.W. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, 9, 99-114.
- Ueno, T. (1999). The roles of inference in discourse interpretation: inference in formal logic and its relation to daily inference. *The Journal of Kansai University of Social Welfare*, vol. 1, 237-293.
- Ueno, T (2003). A discourse analysis of a news article in a daily newspaper. *The Journal of Kansai University of Social Welfare*, vol. 6, 121-132.
- van Dijk, T.A. (1977). *Text and context*. London: Longman.
- van Dijk, T.A. (1980). *Macrostructures*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- van Dijk, T. A. (1988a). *News Analysis: Case Studies of International and National News in the Press*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- van Dijk, T. A. (1988b). *News as Discourse*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Wason, P.C. (1966). Reasoning. In B. Foss (ed.), *New horizons in psychology*; London: Penguin.
- Wason, P.C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, 20, 273-281.