Can Computer Comprehend Stories like Human Reader? : Representing Situation Model as Interconnected Frames

Masahiro Kiyota† Jun Ohashi† Hideji Enokidu‡ Seiichi Komiya†
† Graduate School of Engineering, Shibaura Institute of Technology
3-7-5 Toyosu, Koutou-ku, Tokyo, 135-8548 Japan
‡ Shibaura Institute of Technology
3-7-5 Toyosu, Koutou-ku, Tokyo, 135-8548 Japan

Abstract: It is assumed that the readers construct the situation model during reading a story. The situation model is the mental microworld described in the text. According to the event-indexing model, the readers monitor four conceptual dimensions of the incoming event described in a clause: characters, time, space, causality, and intentionality. They construct the situation model of each incoming event and incorporate the model into the previous model. In the present study, we proposed the procedure to represent the situation model using four kinds of frame and the referent relationship between the frames in order to develop the computer system that comprehends the story like the human readers do.

Key-Words: situation model, frames, four kinds of frame, situation frame, reference relationship, update

1 Introduction

What is the story comprehension? Most discourse psychologists agree with van Dijk and Kintsch's idea that the readers construct three levels of mental representation while comprehending a story: the surface code, the textbase, and the situation model [1]. The surface code preserves the exact wording and syntax of clauses. The textbase contains explicit propositions derived from each sentence. It also involves inferences that are needed to establish local coherence. The situation model is the mental microworld, that is, the integrated mental representation of a state of affairs described in the text. Therefore, it refers to the people, spatial setting, actions, and events developing in the mental microworld. At this point of view, language is a set of processing instruction on how to construct a situational model [2], [3], [4].

The historically prevalent theory of knowledge representation in cognitive science has been that of the amodal (propositional) symbol system [5], [6]. Recently, Barsalou argued that perceptual symbols rather than amodal propositions are building block of cognition. Unlike amodal symbols, perceptual symbols have an analog relationship with their referents [7]. He assumes that perceptual symbols are used in perceptual simulations in human cognition. In the domain of language comprehension, many studies supported his idea that the readers activate and manipulate perceptual symbols during the story comprehension [8], [9], [10].

For research on situation models, the perceptual symbol theory gave an important suggestion that the readers have access to perceptual symbols of all characters and objects contained in the situation model [11]. Because it has already been demonstrated that the readers have access to spatial and temporal information relevant to the protagonist and other characters, this suggestion implied that the readers step into the mental microworld and observe the people, the spatial settings, actions, and events as a participant or a bystander [12], [13].

In the present study, we proposed the procedure to represent the situation model of the children's story using several kinds of frames and the reference relationships between frames. Frames are a notion that Minsky introduced as means to represent the generic knowledge [14]. Because he implicated frames mainly in visual perception, we decided to use frames as the knowledge representation in our enterprise. It is assumed that the readers construct the situation model inferentially through interaction between the linguistic cues in the explicit text and the background world knowledge. In order to make our problems easy, we wanted to avoid dealing with the background world knowledge as much as possible. Since the children's story contains so many linguistic cues that the younger children can construct the situation model with a little world knowledge, we

focused on the situation model of it. Therefore, we had only to devise some simple inference mechanisms based on the linguistic cues. Of course, as our ultimate goal is to make the computer comprehend stories, so the present study is a first step to attain it.

2 Construction of Multithreaded Situation Model

How do the readers construct the situation model while reading the stories? This is a very important question, because we tried to devise the procedure that can not only represent the situation model but also construct it like the human reader. Several discourse psychologists have proposed the cognitive models on how to construct the situation model: e.g., the structure building model, the construction-integration model, the event-indexing model [2], [15], [16], [17]. Among them, the event-indexing model most definitely specifies the language processing tasks, that is, the mental operations that are required to construct the situation model. Consequently, based on this model, we decided to devise the procedure which can construct and represent the situation model in the computer.

The event-indexing model explains the reader's construction of a multithreaded situation model while comprehending simple stories. It makes general claims about both on-line comprehension and about the representation stored in the reader's long-term memory. According to this model, the reader monitors five conceptual dimensions that compose each incoming event in the story: characters and objects, time, space, causality, intentionality (i.e., the character's goal). Each dimension is explained as following respectively;

1) Space It refers to the spatial setting where the incoming event occurs.
2) Time It refers to the temporal setting when the incoming event occurs.
3) Characters and objects They refer to the persons and the things that are involved in the incoming event.
4) Intentionality It refers to a person's goal and his actions in order to attain his goal. Of course, they are embedded in the incoming event.
5) Causality It refers to the causal relationship between the incoming event and the prior event. To take into account that the action instigates the event, the causality seems to be implied by the relationship between the actions.

As long as our proposed procedure is based on the event-indexing model, it must be able to represent these dimensions of the situation model using the several kinds of the frames and the reference relationships.

Zwaan and Radvansky distinguished three states of the situation model in the process of construction and the retrieval of situational information [11]. The current model is constructed, that is, at Time t_n while a reader is reading a particular clause or sentence, called c_n. The integrated model is the global model that was constructed by integrating, one at a time, the models that were constructed at Time t_1 to t_{n-1} while the reader reads clauses c_1 to c_{n-1}. Finally, the complete model is constructed at Time t_1 through t_x and stored in long-term memory after the person has read all clauses (i.e., c_1 to c_x). They called the process of incorporating the current model into the integrated model updating.

In the event-indexing model, the events are building block of integrated situation model. So, when the reader reads a clause, he constructs a model of the situation denoted by the clause. It is reasonable to expect that each event is indexed on each of five dimensions indicated above. Many researchers have actually confirmed that the reader constructs a situation model composed of five dimensions [16], [18], [19], [20], [21], [22]. Specifically, the reading time of an explicit event in a story is found to decrease depending on how many indexes the current model shares with the integrated model. This finding is significant to us, because it suggests how to generate the instance from each frame, establish the reference relationships, and update the frames.

3 Procedure for Representing Five Dimensions

In order to represent five dimensions of the event described in a clause, we prepared several kinds of frames: the character, the space, the time, and the action frames. The character, the space, and time frames corresponds to the characters, the space, and time dimension of the situation model respectively. The action frame is used to represent the causality and intentionality dimensions. We gave the core role to the character frame, then referred to the set of frames interconnected with it based on the reference relationships as the situation frame.

3.1 Character Frame

As described later, the character frame represents the situation of each character involved in a story. Therefore, the situation cahnges in the story can be ultimatly grasped by the character frames of all characters.

At the first setout, we introduce the procedure of generating the character frame here. The character frame represents the character dimension in five conceptual dimensions of situation model. It is equipped with a slot of the name whose value is his/her name as a string. Whenever each character is described in a clause, the frame of him/her is generated as an instance of the character frame. Fig. 1 shows the character frame and the instance of it as an example.

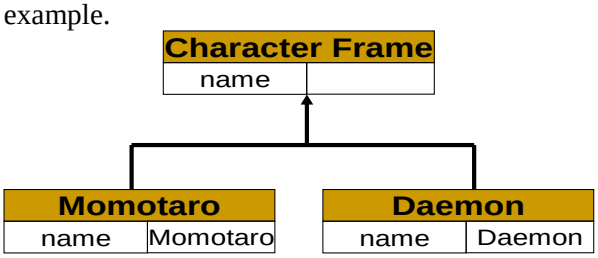

Fig. 1. Character frame and its instances.

3.2 Space Frame

The space frame represents the space dimension of situation model. The space frame has a slot of the name whose value is a string referring to the space. Whenever a space is described in a clause, the frame of the space is generated as an instance of the space frame. Fig. 2 shows space frame and instance of it as an example.

The space in the story is generally described as a place connected with the character (e.g., the place where he/she lives or carries out something). Because the character frame plays the core role in addition to this case, it is equipped with a slot of the space. The value of this slot is a string referring to the space. Accordingly the character's spatial setting can be represented by the reference relationship between the value of the space slot in the character frame and the instance of the space frame. Fig. 3 shows relationship between the character frame and the space frame.

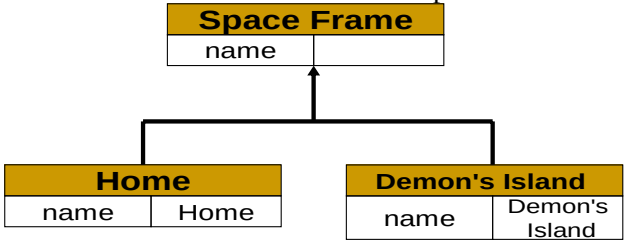

Fig. 2. Space frame and its instances.

Fig. 3. Relationship between character frame and space frame.

3.3 Time Frame

The time frame represents the time dimension of the situation model. The time frame is equipped with a slot of the name whose value is a string referring to the space. Whenever a time is described in a clause, the frame of the time is generated as an instance of the time frame. Fig. 4 shows the time frame and the instance of it as an example.

The character fame is equipped with a slot of the time whose value is a string referring to the time. Like as the spatial setting, the character's temporal setting is can be represented by the reference relationship between the value of the time slot in the character frame and the instance of the time frame. Fig. 5 shows relationship between the character frame and the space frame.

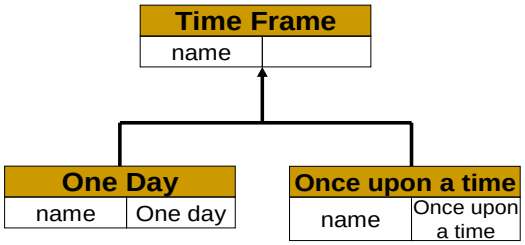

Fig. 4. Time frame and its instances

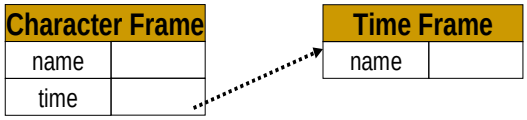

Fig. 5. Reference relationship between character frame and space frame.

3.4 Action Frame

The action frame represents the character's behavior. Whenever a character's behavior is described in a clause, the frame of the behavior is generated as an instance of the action frame. The action frame is equipped with a slot of the action. The quality of the slot value, however, is different from the value to the slots stated above. In other words, the behavior can't be represented by a static value. Therefore we decided to use the method as the value of the action slot. For example, the method of changing the character's space is used as the value of the action slot in the action frame "go". Fig. 6 shows the action frame and its instances.

The character fame is equipped with a slot of the action whose value is a string referred to the behavior. So the character's behavior is represented by the reference relationship between the value of the action slot and the instance of the action frame. Fig. 7 shows relationship between the character frame and the action frame.

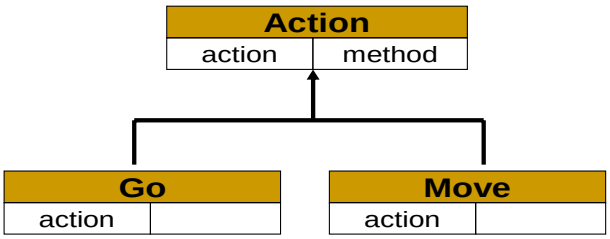

Fig. 6. Action frame and its instances.

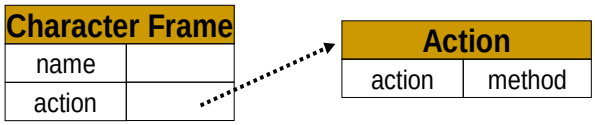

Fig. 7. Reference relationship between character frame and action frame.

The action frame is used to represent the causality dimension and the intentionality dimension of situation model. We introduce how to use the action frame to represent these dimensions in order.

When a behavior causes another behavior, we can confirm the causality of them. So we decided to equip the action frame with a slot of the cause to represent the causality. The value of the cause slot is a string referring to the behavior that causes another behavior. Accordingly the causality can be represented by the reference relationship between the value of the causal slot in the action frame and the instance of the action frame. For example, when the behavior "help" causes the behavior "say reward", this causality is represented by the reference relationship between the value "Help" of the action frame "Say Reward" and the action frame "Help". Fig. 8 shows this causality as an example.

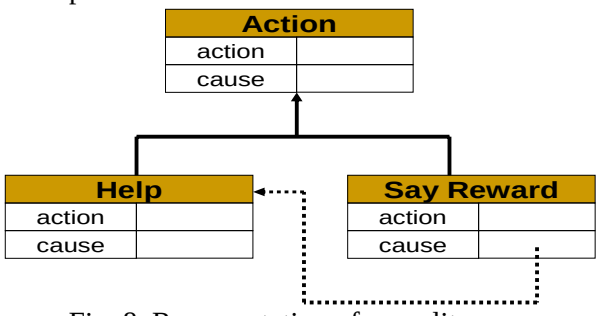

Fig. 8. Representation of causality.

When a character wants to do a behavior, we can confirm his/her intention. So the character frame is equipped with a slot of the intention to represent the intentionality. The value of the intention slot is a string referring to the behavior that the character wants to do. So the intentionality can be represented by the reference relationship between the value of the intention slot and the instance of the action frame. Fig. 9 shows the reference relationship between the characte frame and the action frame. For example, when Momotaro want to beat the demon, his intentionality is represented by the reference relationship between the value "Beat Demon" of the intention slot in the character frame "Momotaro" and the action frame "Beat Demon". Fig. 10 shows such Momotaro's intention as an example.

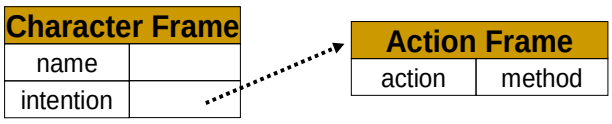

Fig. 9. Reference relationship between character frame and action frame

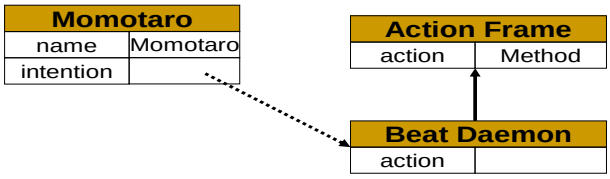

Fig. 10. Representation of intentionality useing action frame

3.5 Situation Frame

As stated earlier, the situation model is necessary to contain five dimensions: the character, the space, the time, the intentionality, and the causality dimensions. So we decided to represent the situation frame as a set of the frames that are interconnected using the reference relationship. Since the character frame is equipped with the slots of the time, the space, the intention, and the action as mentioned so far, it can play the core role in the situation frame. Fig. 11 shows the situation frame that is generated using the refrence

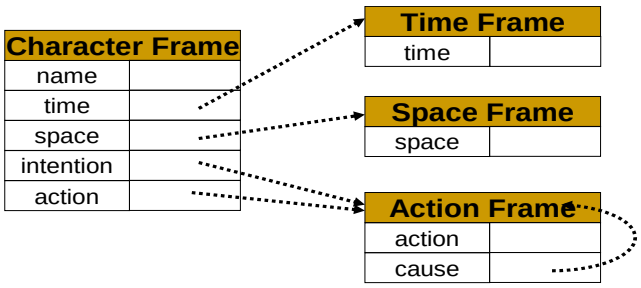

Fig. 11. Situation frame

4 Procedure for Updating Situation Frame

The story usually consists of many sentences. Therefore, whenever each sentence is processed as a input, the situation model will often be updated. The update of the situation frame, however, is virtually accompanied with the changes of the reference relationship. we introduce the procedure for updating the time, situation , the intentionality, and the causality in order.

4.1 Updating Temporal Setting

Whenever the sentence describes the time when the character behaves, his temporal setting must be upated. The temporal setting is updated by changing the new value of the time slot to switch the reference to the instance of the time frame.

it is assumed that the sentence “The day, Momotaro lifted up the washtub overhead” is processed as an input. At this time, the sentence describes the time “The day” when the character “Momotaro” behaves. So Momotaro’s temporal settings is updated by replaced with the value “The day” of the time slot in the character frame “Momotaro” to switch the reference to the time frame “The day”. Subsequently another sentence “One day, Momotaro said I must leave you for a while” is assumed to be processed. The sentence describes the time “One day” when the character “Momotaro” behaves. Then Momotaro’s temporal settings is updated by replaced with the value “One day” of the time slot in the character frame “Momotaro” to switch the reference to the time frame “One day”. Fig. 12 shows the temporal setting updated by switching the reference in this example.

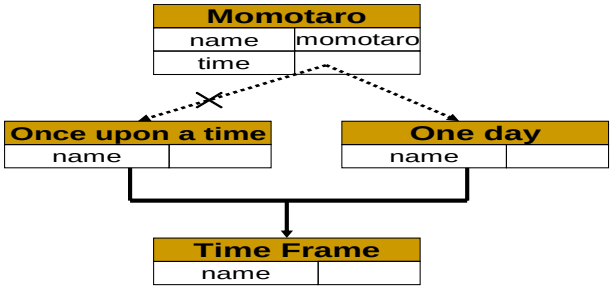

Fig. 12. Temporal setting updated by switching reference

4.2 Updating Spatial Setting

There are two kinds of the procedure for updating spatial setting. The first procedure is applied when the sentence describes the spatial setting where the character behaved. So the spatial setting is updated by same procedure as the procedure for updating the temporal setting. The second procedure is applied when the sentence involves the behavior that implies the change of the spatial setting. In this case, the value of the action slot in the character frame instigates the instance of the action frame through the reference relationship. Then the method of the action slot in the instance changes the value of the space slot and generates the new instance of the space frame simultaneously. The spacial setting is consequently upated.

It is assumed that the sentence “Momotaro went to Demon's Island” is processed as an input. First the value “Go” of the action slot in the character frame “Momotaro” instiagates the instance “Go” of the action frame through the reference relationship. Then , method of the action slot in the instance “Go” replaced with the value “Demon's Island” and generates the instance “Demon's Island” of the space frame. Fig. 13 shows the spacial setting updated by the second procedure.

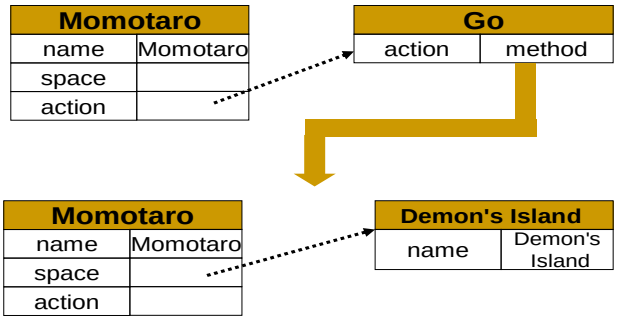

Fig. 13. Spatial setting updated by the second pattern

4.3 Updating Intentionality

The character’s intention can be often identified by linguistic cues such as “determine”, “decide”, and “want”. We introduces the procedure for updating the intentionality as an example about “determine”. At the first setout, the value “determine” of the action slot in the character frame instigates the instance “Determine” of the action frame through the reference relationship. Then, the method in the action slot operates to update the value of the intention slot in the character frame and generate the new instance of the action frame simultaneously.

It is assumed that the sentence “Momotaro determined to beat the demon” is processed as an input. The value of the action slot is replaced with “Determine” to instigate the instance “Determine” of the action frame through the reference relationship.

Then, the method of the action frame in the instance "Determine" operate to replace with the value "Beat Demon" of the intention slot in the instance "Momotaro" of the character frame and generate the instance "Beat Demon" of the action frame. Fig. 14 shows the the intentionily as an example.

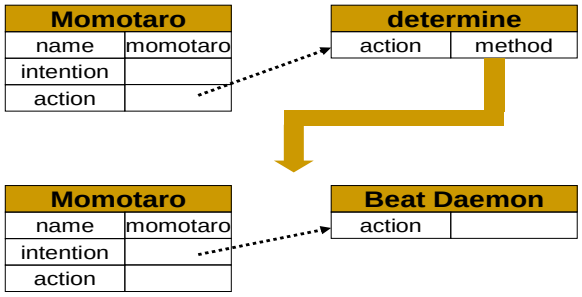

Fig. 14. Representation of updated intentionality

4.4 Updating Causality

The causality between the behaviors can be often identified by the causal connectives such as "because" and "therefore". The causality is updated by setting the value to the cause slot in the action frame.

it is assumed that the sentence "Dog became Momotaro's companion because Momotaro give the dog the millet cake" is processed as an input. At the first setout, the string "Give Millet Cake" is written in the action slot of the instance "Momotaro" of the character frame. The string to "Become Companion" is also written in the action slot of the instance "Dog" of the character frame. At the same time, the instances "Give Millet Cake" and "Become Companion" of the action frame are generated. Then, based on the causal connectives "because", the string "Give Millet Cake" is written in the cause slot in the instance "Become Companion". By such a procedure, the causality is updated by connecting one instance "Become Companion" with anothe instance "Give Millet Cake" using the reference relationship.

4.5 Problem of automatic updating

Whenever each sentence is processed as an input, the situation frame must be automatically updated by the computer. It may indicate tha the computer must be able to generate the appropriate kind of the frame accorging to the meaning of each word contained in the sentence. This indication gives us a fundamental problem. In order to solve the problem, we are examining to generate the four kinds of the frame using outcome of the case analysis.

5 Comprehending Stories by Computer

We explained method for representing the situation model with the frame-based representation and updating procedure of it. Now, we want you to remember that our ultimate goal is to make the computer comprehend stories. Even if the situation model has been represented as the frames, it cannot be said that the computer comprehended the story. So we define the story comprehending by computer as follows;

1. The computer can construct automatically the situation model from sentences of the natural language
2. The computer can answer the question by using constructed situation model.

Probably, these definitions may not be enough to realize the computer that can comprehend the story like human being. They will give us, however, a fundamental target to be aimed at first. We will examine it in the future because we can automate these inferences.

6 Conclusion

In the present paper, we proposed a method for representing the situation model with frame-based representation. In the future, we will work on two tasks. One task is refining the procedure of constructing the situation model. Another task is devising the some mechanisms that can make necessary inferences to answer the question and mounting them into the system.

References:

[1] van Dijk, T. A., & Kintsch, W., *Strategies in discourse comprehension*, New York: Academic Press, 1983.

[2] Gernsbacher, M. A., *Language comprehension as structure building*, Hiiisdale, NJ: Erlbaum, 1990.

[3] Givón, T., *The grammar of referential coherence as mental processing instructions, Linguistics, 30, 1992, pp. 5-55.*

[4] Kintsch, W., How readers construct situation models for stories: The role of syntactic cues and causal inferences. In A. F. Healy, S. M. Kosslyn, & R.M. Shiffrin (Eds.), *From learning processes to cognitive processes. Essay in honor of William K. Estes* (vol.2, pp. 261 - 278), Hiiisdale, NJ: Erlbaum, 1992.

[5] Foder, J., *The language of thought*, Cambridge, MA: Harvard University Press, 1975.
[6] Newell, A., & Simon, H., *Human problem solving*, Englewood Cliffs, NJ: Prentice Hall, 1972.
[7] Barsalou, L. W., Perceptual symbol systems, *Behavioral and Brain Sciences*, 22, 1999, pp. 577-660.
[8] Stanfield, R. A., & Zwaan, R. A., The effect of implied orientation derived from verbal context on picture recognition, *Psychological Science*, 12, 2001, pp. 153-156.
[9] Fincher-Kiefer, R., Perceptual components of situation models. *Memory & Cognition*, 29, 2001, pp. 336-343.
[10] Zwaan, R.A., Stanfield R. A., & Yaxley, R. H., Language comprehenders mentally represent the shapes of objects, *Psychological Science*, 13, 2002, pp. 168-171.
[11] Zwaan, R. A., & Radvansky, G. A., Situation models in language comprehension and memory, *Psychological Bulletin*, 123, 1998, pp. 162-185.
[12] Morrow, D.G., Greenspan, S.L., & Bower, G.H., Accessibility and situation models in narrative comprehension, *Journal of Memory and Language*, 26, 1987, pp. 165-187.
[13] Zwaan, R. A., Madden, C. J., & Whitten, S. N., The presence of an event in the narrated situation effects its activation, *Memory & Cognition*, 28, 2000, pp. 1022-1028.
[14] Minsky, M., A framework for representing knowledge. In P.H. Winston (Ed.), *The psychology of computer vision* (pp. 211-277), New York: McGraw-Hill, 1975.
[15] Kintsch, W., The role of knowledge in discourse comprehension: A construction-integration model, *Psychological Review*, 95, 1988, pp. 163-182.
[16] Zwaan, R. A., Langston, M. C., & Graesser, A. C., The construction of situation models in narrative comprehension: An event-indexing model, *Psychological Science*, 6, 1995a, pp. 292-297.
[17] Zwaan, R. A., Magliano, J. P., & Graesser, A. C., Dimensions of situation model construction in narrative comprehension, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 1995b, pp. 386-397.
[18] Rinck, M., Williams, P., Bower, G. H., & Becker, E. S., Spatial situation models and narrative understanding: some generalizations and extensions, *Discourse Process*, 21, 1996, pp. 23–56.
[19] Zwaan, R. A., Processing narrative time shifts, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 1996, pp. 1196-1207.
[20] Trabasso, T., & Magliano, J. P., Conscious understanding during comprehension, *Discourse Process*, 21, 1996, pp. 298-327.
[21] Dopkins, S., Representation of superordinate goal inferences in memory, *Discourse Process*, 21, 1996, pp. 85–104.
[22] Hakala, C. M., & O'Brien, E. J., Strategies for resolving coherence breaks in reading, *Discourse Process*, 20, 1995, pp. 167–86.