Simulation of cued-recall and recognition of expository texts by using the Construction-Integration Model (Kintsch, 1988)

Isabelle Tapiero
URA 1297 au C.N.R.S.,
Université de Paris 8,
93526 Saint-Denis (France)
ITAPIERO@FRP8V11 &,
Université de Lyon II, Institut de Psychologie,
69676 Bron (France).
Tel: 78-77-24-32 - Fax: 78-74-22-17.

&

Guy Denhière
URA 1297 au C.N.R.S.,
Université de Paris 8,
93526 Saint-Denis (France)
DENHI@FRP8V11

Abstract

The purpose of this paper is to compare the results of adults' performances in cued-recall and recognition of expository texts with the results of simulations derived from the construction-integration model proposed by Kintsch (1988, 1990). In the cued recall task, we manipulated three parameters: the weights of the connexions in the net, the size of the short-term memory buffer and the representations of the sentence used as a cue. The main results show that we need to simulate the macroprocessing and the prior knowlegde of the learners to be able to increase the simulations. In the recognition task, we simulate the representation of 4 levels (surface syntactic variation, close semantic variation, inference and distant semantic variation referring to the same situation model than the text) using different connexion weights function the decay of the memory trace. The main results show the necessity to take into account these levels to explain the subjects' cognitive processes involved in a comprehension/memorization task. For both experiments, the activation values obtained correctly fit the hierarchy of the experimental data.

Introduction

In contrast with expectation-based models of comprehension, the construction-integration model proposed by Kintsch (1988, 1990) is strictly bottom-up and doesn't use pre-stored schemata or "smart" rules for controlling the process of comprehension. According to this model, reading a text leads, in addition to the activation of the "correct" representations to the activation of non relevant, redundant, and even contradictory information, which will be, during a second phase, desactivated by a relaxation connectionnist process. We tested this model of cognitive architecture with two tasks: immediate cued recall and recognition following the reading of an expository text and we examined how the simulations fit the experimental data.

Cued recall task

In the cued recall experiment, we study the influence of prior knowledge on the learning and transfer by adults. 32 students of the University of Colorado in Boulder participated in this experiment[1]. According to their major (biology or psychology), they were assigned to two groups of 16 subjects: experts and novices. The structure of the domain knowledge consisted in 4 instanciations of the sea-animal category: two exemplars of sea-mammals "Dolphin" and "Manatee", and two categories of sea animals "Fish" and "Sea-mammal". 4 texts referring to the 4 domains presented above have been constructed: "Manatee" (Text 1), "Dolphin" (Text 2), "Sea-mammal" (Text 3), and "Fish" (Text 4), each text being structured with three levels of macrostructure. One week after reading

[1] We are indebted to Professeur Walter Kintsch who provides us the funds necessary to perform this experiment.

one of the 4 texts on the sea animals, all the subjects had to read the text "Manatee" and to recall "all that they remember about the young manatee". The instructions were given orally and the cued-recall task lasted 10 minutes.

In this paper, we will focus on only one main aspect of this research: in which extent the simulations performed fit the experimental data on the cued recall of the transfer text "Manatee" (preceded by 4 different learning texts: "Manatee", "Dolphin", "Sea-mammal"and "Fish")

Parameters. Three parameters were manipulated: the weights of the connexions in the net, the size of the short-term memory buffer and the representations of the sentence used as a cue.

Weights of the connexions. Three systems of weights have been used: p1 (1,0) ; 1 for a one step connexion, 0 for other connexions ; p2 (.9, .5, 0) ; .9 for a one step connexion, .5 for a two steps connexion, 0 for other connexions ; p3 (.9, .7, .3, 0) ; .9 for a one step connexion, .7 for a two steps connexion, .3 for a three steps connexion and 0 for other connexions. We postulated a greater connexity in the network for the experts than for the novices, so the correlations between simulations and recall will be hierarchized as follow: p3>p2>p1.

Size of the short-term memory buffer. We chose 2 values for the parameter s: s = 1 and s = 3. We assume that the capacity of short-term memory vary as a function of the prior knowledge of the learner and that, for the experts, the correlations between simulations and recall will be higher for s = 3 than for s = 1. The reverse was expected for the novices.

The representations of the cue. Four representations of the sentence "Recall all that you remember about the young manatee" have been used. In the first one (i1), the cue was represented by one text base proposition (**P24. YOUNG[P1], P1= Manatee (x)**). The activation value of this proposition was systematically set to 1 for each iteration up to stabilization of the net. In the second one (i2), the cue was represented by 2 text base propositions (**P24. YOUNG[P1], P1= Manatee (x) and P33. YOUNGSTER[P1]**). The activation values of these two propositions were systematically set to 0.50 for each iteration up to stabilization of the net. In the third one (i3), the cue was represented by 2 text base propositions (P24 and P33 plus the only necessary inference generated in the text (**REFERENCE [P24: YOUNG[P1],P33:YOUNGSTER[P1]]**). Finally, in the fourth representation (i4), the cue was represented by 3 text base propositions: P24, P33, plus P32 (**P32. MOTHER[P33, P1]**). For both the third and the fourth representations, an activation value of 0.33 was assigned to each proposition of the cue sentence.

Results

The correlations between the simulations and the subjects' performances show four main results:

1. For the 4 groups, higher correlations were obtained for the novices than for the experts (.40 vs. .32, in average), the correlations values vary between .35 and .48, for the novices, between .10 and .44 for the experts, with the lowest correlations for the interference group (Fish-Manatee). The prior learning of the fish characteristics interfere with the learning of the Manatee characteristics, and this interfering effect is more important for the experts.
2. p1(1, 0) and p3 (.9, .7, .3, 0) led to higher correlations than p2 (.9, .5, 0): .38 and .40 vs. .33.
3. No difference between experts and novices and between the two sizes of the short-term memory buffer was observed.
4. i2, i3 and i4 led to higher correlations than i1.

Conclusions

These correlations are lower than those obtained with narrative texts. We know that prior knowledge play an important role in the comprehension and memorization of expository texts, therefore we postulate that it is necessary to simulate the macroprocessing of the text and the prior knowlegde of the learners to increase the correlations.

Recognition task

The goal of this second experiment is to study the recognition of an expository text by adults and to extend the construction-integration model proposed by Kintsch (1988, 1990) to the simulation of the recognition of six types of statements. We make the assumption that reading a text leads a subject to construct different levels of representation: the surface structure, the syntactic structure, the local and global semantic structure (the micro- and macrostructure) and the situation model. Related to these levels, we constructed six types of statements to be recognized: verbatim (verb), syntactic surface variation (ssv), close semantic variation (paraphrase: csv), inference (inf), distant semantic variation dealing with the situation model referring to the text (mdsv), distant semantic variation dealing with an other situation model than the one referring to the text (odsv). We suppose that these different statements to be recognized will inform us on the mnemonic traces associated to the representation of these different levels and on the differential forgetting of these traces related to these levels. 4 groups of 26 psychology students read a text untitled "Manatee". This text was composed by an introduction (3 sentences), a first topic on "Breathing" (10 sentences), a second topic on "Reproduction" (10 sentences) and a conclusion (3 sentences). The recognition task only dealt with the topic "Reproduction".

The reproduction of the manatee. We present below the 10 sentences dealing with the "Reproduction of the Manatee".

1. The manatee female has a single pair of mammary glands attached to the chest.
2. The reproduction period for a manatee lasts from april to august.
3. The gestation period for a manatee lasts six months.
4. the mother give birth in the water to only one young.
5. The mother pushes its youngster to the surface of the water and permits him to breathe.
6. The mother suckles its youngster during eighteen months.
7. The youngster's suckling is done at the surface of the water.
8. At birth, the manatee weighs in average 60 pounds.
9. The skin of the young manatee, pink in color, is bare.
10. The young manatee becomes an adult at the age of three years.

The statement to be recognized. For each text sentence, we constructed 2*6 types of statements to be recognized

1. Verbatim (verb)

1. The gestation period for a manatee lasts six months.
2. The young manatee becomes an adult at the age of three years.

2. Surface syntactic variation (ssv)

1. From april to august, lasts the reproduction period for a manatee.
2. Pink in color, the skin of the young manatee is bare.

3. Close semantic variation (csv)

1. The manatee female gives life to a single youngster in the water.
2. The mean weight of the new born manatee is around 60 pounds.

4. inference (inf)

1. At birth, the manatee has to be pushed at the surface of the water to breathe.
2. The youngster manatee is feeding by its mother.

5. Manatee distant semantic variation (mdsv)

1. The manatee female has two pairs of mammals attached to the pectoral limbs.
2. The manatee female weans its youngsters one month after the birth.

6. Other distant semantic variation (odsv)

1. The children enjoy the show of the sea-lion balancing a ball on its nose.
2. Hanged by their legs, upside down, the bat gets out during the night.

Results

The recognition time

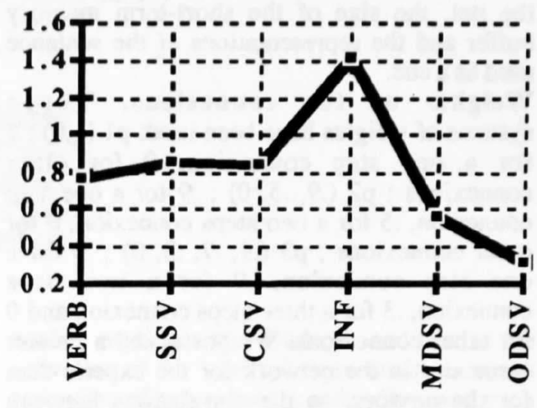

Figure 1: Recognition time (cs.) of the different statements to be recognized. [Legend: VERB = Verbatim ; SSV = Syntactic Surface Variation ; CSV = Close Semantic Variation ; INF = Inference ; MDSV = Distant Semantic Variation referring to the Manatee ; ODSV = Distant Semantic Variation referring to an Other animal].

The correct responses

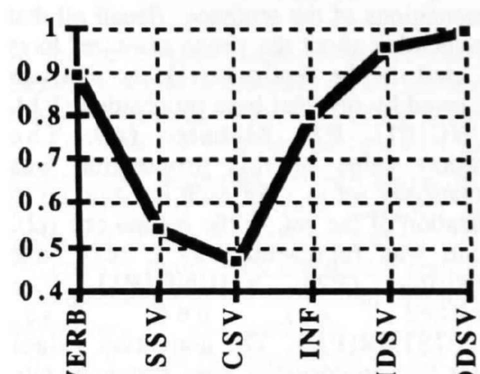

Figure 2: Correct responses of the different statements to be recognized. [Legend: VERB = Verbatim ; SSV = Syntactic Surface Variation ; CSV = Close Semantic Variation ; INF = Inference ; MDSV = Distant Semantic Variation referring to the Manatee ; ODSV = Distant Semantic Variation referring to an Other animal].

Principles of the simulation

In the simulations performed, we not only take into account the level of the local semantic structure (the propositions) but also the level referring to the surface structure introduced by words. We assigned an initial weight of 1 for the relations between the propositions and the words, and between the words. A weight of 2 relates the propositions of the text base. These weights of 1 and 2 are supposed to explain the differences in the strengths between the mnemonic traces of the surface representation and those of the semantic structure of the text. For the **Surface Syntactic Variation**, we consider that it differs from the verbatim statement only by **the order of words** presentation in the sentence and that the semantic text base is identical. A weight of -3 is assigned to the relation between the two orders, showing that the activation of one inhibits the activation of the other one. The **Close Semantic Variation** is supposed to differ from the verbatim statement only by the type of **lexical items**. We assigned a weight of -3 to the relation between the verbatim lexical items ("Birth manatee", "To weight", and "In average") and the close semantic lexical items "New-born", "Weight", and "Around"). "Order 1" corresponds to the words presentation order in the verbatim statement, and order (x) to that in the close semantic variation. A weight of -3 is assigned to the relation between these two orders. We notice that the lexical item "60 pounds" refers both to the verbatim statement and to its close semantic variation. For the **Inference**, we not only need to include the new lexical items and their presentation order but also new propositions to explain the semantic distance between the inference and the verbatim statement. As for the inference, the **Distant Semantic Variation referring to the Manatee**, differ from the verbatim statement by the type of lexical items and by the type of propositions.

Results

The figure 3 below presents the experimental data and the results of the simulations.

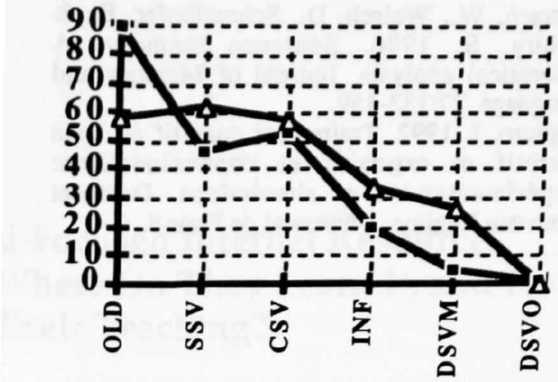

Figure 3: Probanility of responses "yes" as a function of the different types of statements to be recognized.[Legend: the white triangles represent the predictions of the model and the black squares represent the experimental data].

The hierarchies obtained for the activation values of the simulation of the different types of statements to be recognized and for the responses "Yes" are almost identical. Hierarchy of the responses "YES": **VERB>SSV<CSV>INF>MDSV>ODSV** Hierarchy of the activation values **VERB≤SSV>CSV>INF>MDSV>ODSV**

Conclusion

The construction-integration model proposed by Kintsch (1988) permits us to formulate precise predictions concerning the subjects' representation and the processes they apply in cued-recall and recognition tasks, in the study of expository texts with adults. The first consequence of the extension of this model is to proceed to more accurate texts analyses using the representation of different levels: the surface structure, the syntactic structure, the local and global semantic structure and the situation model. The second consequence consists in the psychological relevance of the adopted approach: we applied the same principles for all the simulations and no ad hoc procedure have been used.

References

Denhière, G., & Baudet, S. 1992. Lecture, compréhension de texte et science cognitive. Paris, Presses Universitaires de France.

Kintsch, W., & van Dijk, T.A. 1978. Toward a model of text comprehension and production. Psychological Review, 85:363-394.

Kintsch, W. 1988. The role of knowledge in discourse comprehension: A construction-integration model. Psychological Review,

95:163-182.
Kintsch, W., Welsch, D., Schmalhofer, F., & Zimny, S. 1990. Sentence memory: A theoretical analysis. Journal of Memory and Language, 27:133-159.
Tapiero, I. 1992. Traitement cognitif du texte narratif et expositif et connexionnisme: Expérimentations et simulations. Doctorat Nouveau Régime, Université de Paris 8.