

How Primary Classes Visually Represent *While* Temporal Relations: A Preliminary Evaluation Study

Tania Di Mascio^a, Rosella Gennari^{b*}, Barbara Arfé^c

^aDIEI, University of l'Aquila, Monteluco di Roio, 67100 l'Aquila, tania.dimascio@univaq.it

^bKRDB, Free University of Bozen-Bolzano, Via della Mostra 4, 39100 Bolzano, Italy, gennari@inf.unibz.it

^cUniversity of Verona, Lungadige Porta Vittoria 17, 37129 Verona, Italy, barbara.arfe@univr.it

Abstract

We are working on a temporal reasoning web tool for novice readers. The acquisition of temporal relations and reasoning with them depends on age and experience, as well as linguistic factors. We conducted a preliminary evaluation with 6–8 olds in order to assess whether and how they would visually represent “while” temporal relations of a story. In this paper, we present and discuss our experimental evaluation, which paves the way for the visual representation of such relations in our e-tool.

Introduction

Temporal reasoning is one of the many cognitive skills that children must develop in order to integrate well in our society: “of the many cognitive skills which children must master in order to become proficient members of their cultures, the acquisition of commonsense time concepts is among the most essential” (Scott 1997). The study of (Duran et al. 2007) supports the relevance of temporal features of texts as viable cues for facilitating the coherent interpretation of the texts. Reasoning coherently with time concepts is acquired indirectly through narration, and evolves with age and experience. Language specific factors also affect the age at which temporal connectives are comprehended and mastered. In general, after the age of 5, normally developing children become able to make deductions with temporal relations, reasoning on sequences of events with “before” and “after” (McColgan and McCormack 2008). This ability seems to develop further from the age of 7 to that of 9, when children seem to be able to master the “while” temporal connective, e.g., see (Ge and Xuehong 2002).

We are working on an e-story comprehension web tool. The tool’s users include Italian primary-school children (Orsolini et al. 2008). The tool originates from LODE, a logic-based web system for the literacy of deaf readers,

e.g., see (Gennari and Mich 2007) and (LODE). The tool invites readers to reason on the temporal dimension of an e-story by using temporal relations, between pairs of events of the e-story, that can be expressed with “before”, “while” or “after”. An example is “Mammy chicken is worried about her little Gino, while Gino is telling stories to the wolf”. They correspond to Allen qualitative temporal relations (Allen 1983), namely, disjunctions of the atomic relations of Figure 1. For brevity, we will denote such relations as *novice reader temporal relations* (NRTRs).

After reading an e-story of the e-tool, a series of comprehension games ask readers to deduce coherent (possibly implicit) NRTRs between events of the e-story. The readers will need to construct a model of the flow of the e-story’s events in order to answer coherently.

Based on (Pavio 1991), numerous studies already showed significant comprehension gains when people can visualise while reading, as reported in (Johnson-Glenberg 2005). Our web tool aims to be visual as for: (1) the interface; (2) the story’s events; (3) the e-story’s NRTRs.

A non-trivial challenge comes forward in the creation of such a visual tool: How can we render NRTRs with a visual representation that is as natural as possible for primary-school children?

We took over the challenge and conducted a preliminary evaluation with fifty-six 6–8 olds in order to assess how they would visually represent “while” NRTRs between events of a story. Our evaluation aims at assessing: (1) whether 6–8 olds can all visually represent “while” NRTRs; (2) the strategies they spontaneously employ in their visual representations.

This paper reports on the results of our evaluation, and paves the way for the development of our visual tool.

Related Work

The AI and HCI literatures count several spatial representations of qualitative temporal relations between pairs of events, e.g., see (Gennari, Di Mascio and De Gasperis 2009) and (Gennari and Di Mascio 2009). Figure 1 is the standard visualisation of the so-called atomic relations of Allen. (Hibini and Rundensteiner 1997)

*Work partially supported by the CARITRO Foundation, Trento, Italy (<http://lode.fbk.eu>).

proposes an interesting alternative, capable of expressing indefinite information such as the Allen relation “before or meets”. More recently, (Chittaro and Combi 2001) proposed three alternative visual metaphors for the representation of events and their Allen relations, see Figure 2. They can scale up to the visualisation of relations in a network with more than two events, e.g., see Figure 2(c). Their metaphors are based on concrete objects and phenomena from the physical world: elastic bands, springs and paint strips.



Figure 1. The atomic Allen relations.

All such visualisations of temporal Allen relations are based on linear spatial orderings. Is this assumption consistent with children’s visual representation of time, in general, and of NRTRs, in particular? To the best of our knowledge, there are very few studies that assess such questions with novice readers.

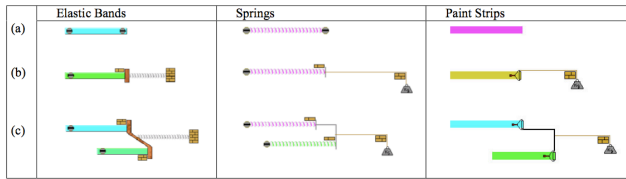


Figure 2. The visualisation by (Chittaro and Combi, 2001).

For instance, (Koerber and Sodian 2008) evaluated preschool children’s ability in mapping a sequence of three temporal events onto spatial relations on a panel. The authors tested whether children would spontaneously produce and comprehend a spatial linear ordering. According to their results, children’s “mapping appears to be influenced by cultural conventions (such as the left-to-right direction of reading and writing), but space is used to represent non-spatial relations spontaneously before these cultural conventions are learned” (*ibidem*, p. 394).

The results of (Arfé, Gennari and Mich 2009) seem to indicate that 7–8 olds can comprehend NRTRs better with a spatial linear representation, based on the one in Figure 1, than with a choice-box textual visualisation.

However, the spontaneous visual representations of “while” seems by far less explored than that of sequential NRTRs, like “before” or “after”. Our evaluation leaves away sequential NRTRs, and concentrates on “while” sentences, correlating events of a story.

Rationale and Goals of the Evaluation

We asked 6–8 olds of an Italian primary school to represent “while” NRTRs of an entire story by drawing, and by arranging the illustrations of the related events on an A4 paper.

We opted for a story in our evaluation instead of, say, brief isolated sentences like in (Koerber and Sodian 2008). Stories give children a meaningful context, and a multiple-sense connective is particularly sensitive to the context. As reported in (Winskel 2004), in general, children are more successful at interpreting texts where the context reinforces the interpretation of the temporal connective. This is the case of “while” in English as well Italian. It can correspond to different Allen temporal relations, according to the context. For instance, it can correspond to “during” as well as to “during or finishes”.

We chose the first short story, of circa 250 words, from the picture-book (Gunthorp and Cassinelli 2002) for children, older than 5. The story characters and settings are well defined and clearly depicted, the temporal flow is linear (e.g., no flashbacks), there are “while” NRTRs between significant temporal events, and the language is suitable for 6–8 olds alike.

We also decided to read the story, instead of letting children read the story on their own. The decoding skills of some of the children involved in the evaluation might still have been rather immature, and this might have had an affect their reading comprehension. Story telling, instead, is an experience familiar to all, from their preschool year.

Our evaluation aims at assessing the following:

- (1) if class, grammar comprehension, and working memory capacity affect the child’s capability in visually representing the “while” NRTRs;
- (2) the type and quality of their visualisation strategies, e.g., the dimension of events in drawing.

Method

The experimental evaluation we conducted is based on a classical HCI user based schema, e.g., see (Di Mascio *et al.* 2005). It consists of the following steps.

- *User Analysis*, that is, the description of the involved users and definition of users’ categories.
- *Experiment Design*, that is, the definition and description of study models, metrics and tasks.
- *User Teaching*, that is, exhaustive explanation for people involved in the evaluation about the *modus operandi* in the experiment sessions.
- *Experiment Execution*, that is, the description of experiment sessions, a.k.a., experiment diary.
- *Result Analysis*, that is, the collection of data and description of significant results.

In the following, we describe the experimental evaluations, using the aforementioned steps.

User Analysis

Our experiment participants were fifty-six children at end of the school year: one first-year class (1a), with seventeen 6–7 olds; two second-year classes, namely, (2a) with nineteen 7–8 olds, and (2b) with twenty 7–8 olds.

The grammar comprehension level, and the working memory skills of (1a) and (2a) were assessed by means of the digit span test, briefly DS (WSC-R), and a receptive grammar test, briefly G. See (Rustioni and Lancaster 1994). Each child of (1a) and (2a) thus obtained a DS score (min 0, max 28) and a G score (min 0, max 16). This was not possible with (2b) due to time restraints. On (1a) and (2a), we can then compare:

- the children's DS scores to the DS median values (DSm), that are equal to 9 for (1a) and 12 for (2a);
- the children's G scores to the G median values (Gm), that are equal to 10 for (1a) and 12 for (2a).

The comparison allows us to group children of (1a) and (2a) in the DS/G categories of Table 1. Note that X-HC means that a child of the X class has both DS and G scores higher than DSm and Gm. X-MC means that a child of the X-class has either DS higher than DSm and G lower than Gm, or vice-versa. X-LC means that a child of the X class has both DS and G scores lower than DSm and Gm. These are then used in the result analysis.

DS/G categories	School class	DS/G scores
1-HC	(1a)	DS > DSm and G > Gm
1-MC		either DS > DSm or G > Gm
1-LC		DS < DSm and G < Gm
2-HC	(2a)	DS > DSm and G > Gm
2-MC		either DS > DSm Or G > Gm
2-LC		DS < DSm and G < Gm

Table 1. DS/G categories, according to school class (2nd col.), and DS or G scores (3rd col.).

Experiment Design

We divided our experiment into three phases: Phase (1), assessment of working memory and grammatical skills; Phase (2), arrangement of transparent illustrations on a white paper; Phase (3), spontaneous drawing.

Phase (1) served to classify children in the Table 1 categories. Phases (2) and (3) served to study the capability of visually representing the “while” NRTRs. Whereas the children's preferred visual patterns of representation may emerge in Phase (2), common drawing strategies of children can emerge in Phase (3).

Since we cannot assume that all the experiment's children equally comprehend the “while” NRTRs, the

metrics for the analysis are: frequency analysis of patterns for Phase (2); frequency of drawings special patterns and strategies for Phase (3). We used the classical score metric of DS and G tests for Phase (1).

The tasks for children in the three phases are summarised in Table 2. Note that all the sentences assigned to children in T3–6 narrate significant episodes of the story. According to the story, the “while” between the first and the second event introduced in T3–5 can be mapped into the “during” Allen atomic relation. The T6 “while” is more vague, and can be mapped to “during or finishes”.

Phase	Task acronym	Task description
1	T1	Repeat aloud the sequence of numbers read by the evaluator
	T2	An examiner reads a sentence. The child is asked to point to one out of four pictures, which he or she thinks to portray the sentence best
2	T3	Represent the following with transparencies: “The wolf reaches Gino while Gino is picking up strawberries in the woods”
	T4	Represent the following with transparencies: “Mammy chicken and the other animals go to the wolf's house while Gino is telling stories to the wolf”
3	T5	Draw “Mammy chicken and the other animals go to the wolf's house while Gino is telling stories to the wolf”
	T6	Draw “While mammy chicken is worried about her little Gino, Gino is telling stories to the wolf”

Table 2. Task descriptions per phase.

User Teaching

Before performing the experiment, on June 1 2009, the two experimenters met the classes' teachers and dean. During the meeting, they discussed the organisation of the experiment (e.g., meeting time, sequence of tasks), and their respective roles in the experiment. More precisely, teachers were asked to support the experimenters and children in all the phases of the experiment. Teachers were asked to provide considerable support in Phase (1). This phase is a sort of *vis-à-vis* interview with each child, and as such it requires a considerable amount of time.

Then the experimenters explained the experiment *modus operandi*, detailed as follows.

- There should be a relaxed and playful atmosphere during the experiment, e.g., the story is told imitating the animals' voices.
- The absolute respect of privacy must be clear to children. A personal number identifies each child, unequivocally but anonymously.

- The usage of transparent illustrations is explained.
- One side of the A4 sheet is dedicated to spontaneous drawing (for T5 and T6). The other side is used for transparencies (T3 and T4).

Finally, experimenters and teachers fix the amount of time that children can spend on each task, especially T5 and T6.

Experiment Execution

The experiment took place at school, a familiar environment for children. At 8:00, the two experimenters met the teachers at their school. Then the experiment is divided in three consecutive sessions, one session per class. *Session I* is carried on with (1a). The experiment phases are in the following order: Phase 1, Phase 2, Phase 3. Tasks are assigned in the following order: T1, T2, T3, T5. *Session II* is carried on with (2a). The experiment phases are in the following order: Phase 1, Phase 3, Phase 2. Tasks are assigned in the following order: T1, T2, T6, T4. *Session III* is carried on with (2b). There is only Phase 3 and the T5 task.

See Table 3 for a recap.

Session	Class	Phase order	Task order
I	1a	1,2,3	T1,T2,T3,T5
II	2a	1,3,2	T1,T2,T6,T4
III	2b	3	T5

Table 3. Description of sessions.

In Sessions I and II, drawing and placing transparencies are executed in different order so as to augment the independency between the DS/G scores, and the child's capability of visually representing "while" NRTRs. In Session III, time restraints compelled us to choose either transparencies or drawings. We opted for the latter as it gives more freedom to children. The remainder of this section details the three sessions.

Session I, date: 05 June 2009

(I.i) 08:30–08:45. The evaluators meet (1a) children in the classroom. They explain children their role, e.g., "... children, please, help us! We need to explain 'while Gino is telling stories, mammy chicken is worried' to younger children...". Then the evaluators give each child his or her personal number.

(I.ii) 08:45–09:15. First, one of the evaluators demonstrates how T1 and T2 must be performed with the aid of one child. Then the evaluators and the teacher collect the T1 and T2 data.

(I.iii) 09:15–09:30. First, one of the evaluators read the story while children listen carefully. Then the evaluators give one transparency set, and one white A4 sheet per child. Finally, the evaluators ask the children to write on the paper sheet their number.

(I.iv) 09:30–09:50. First, one of the evaluators reads the T3 sentence. Then the children arrange transparencies on the paper sheet. Finally, the evaluators trace the pattern created by the children on the white paper.

(I.v) 09:50–10:30. First, one of the evaluators reads the T5 sentence. Then the children draw on the blank side of the paper sheet.

(I.vi) 10:30–10:40. The experimenters and teacher collect paper sheets, leaving the transparencies as presents.

Session II, date: 05 June 2009

(II.i) 10:45–11:00. As (I.i) of Session I.

(II.ii) 11:00–11:15. As (I.ii) of Session I.

(II.iii) 11:15–11:30. As (I.iii) of Session I.

(II.iv) 11:30–12:00: As (I.v) of Session I with T6 instead of T5.

(II.v) 12:00–12:20: As (I.iv) of Session I with T4 instead of T3.

(II.vi) 12: 20–12:30: As (I.vi) of Session I.

Session III, date: 05 June 2009

(III.i) 12:45–13:00. As (I.i) of Session I.

(III.ii) 13:00–13:15. As (I.iii) of Session I.

(III.iii) 13:15–13:30. As (I.v) of Session I.

(III.iv) 13:30–14:00: As (I.vi) of Session I.

Results Analysis

For space limitations, this section only gives the most significant results, and leaves out the analysis of T3.

Phase 1 (T1, 1a and 2b; T2, 1a and 2b)

In our analysis of Phase 1, children are grouped using the DS/G categories of Table 1, which are described in the User Analysis section. The analysis results are shown in Table 4. Note that there are no children in 1-LC nor in 2-LC. Thus all are at a good level, with high DS or G scores.

DS/G category	N° of children
1-HC	7
1-MC	10
1-LC	0
2-HC	9
2-MC	9
2-LC	0

Table 4. Results of Phase 1.

Phase 2 (T4, 2a)

In order to represent Event 1 ("Mammy chicken and the other animals go to the wolf's house") of T4, children had to arrange two transparencies on the paper sheet, one for Gino and the other for the wolf. The same happens with Event 2 ("Gino is telling stories to the wolf") of T4. The horizontal linear distribution of the two events is dominant (50%). Generally, children represented the two events on different horizontal lines. Only 17% of the children represented only one event horizontally. 33% of them gave no clear order between events, or used only the transparency of one character without representing any event. Children with higher G scores represented both the events in T4 horizontally more frequently than children with lower G scores: $\chi^2(1)=5,84$, $p<.01$. Children with

higher G score represented both events more frequently than children with lower G score: $\chi^2(1)=5,84, p<.01$.

Phase 3 (T5, 1a and 2b; T6, 2a)

Children's drawings were coded according to: (a) the spatial arrangement of the two events in their drawings; (b) the visual strategies for representing the “while” NRTRs.

As for (a), three major spatial arrangements emerged: horizontal (two separate events along an horizontal line), vertical (two separate events on a vertical line), diagonal (two separate events on a diagonal). Representations consisting of one single event or no clear distribution of events are coded as “other”.

As for (b), the analysis also revealed three main visual strategies for “while”, explained as follows.

S1. Children draw a background common to both events. That is, the child represents the two events within the same background frame (e.g. sky, ground).

S2. Children blend both events in a unique scene. The child represents the two events as part of a single scene (e.g., mummy chicken and the other animals are close to Gino, while Gino is telling stories).

S3. Children represent two separate events and join them with a link (e.g., a path), albeit they happen in different locations in the story.

Let us see (a) and (b) in the context of the T5 and T6 tasks.

T5. The predominant representation is horizontal for (2b). See Table 5.

Class	Horizontal	Vertical	Diagonal	Other
(1a)	77,00%	12,00%	0,00%	11,00%
(2b)	67,00%	9,50%	0,00%	23,50%

Table 5. Direction in T5.

A linear horizontal representation with the core event (“Gino was telling stories to the wolf”) on the right and the other on the left was predominant for all the experiment participants, independently of their class. A representation with the core event on the left and the secondary on the right was rare, but significantly more frequent among the younger children, that is, (1a) children: $\chi^2(1)=3,75, p<.05$. Figure 3.

The drawing strategies are not mutually exclusive. (1a) children mostly employed the S1 strategy, $\chi^2(1)=6,38, p<.01$, the least employed by (2b) children. See Table 6.

Class	Common background	No background	Other
(1a)	59,00%	12,00%	29,00%
(2b)	19,00%	62,00%	19,00%

Table 6. Background strategy in T5.

The S2 strategy was rather frequent for both (1a) children (65%) and (2b) children (48%). See Table 7.

Class	Single scene	Two scenes	Other
(1a)	65,00%	12,00%	23,00%
(2b)	48,00%	29,00%	23,00%

Table 7. Use of scenes in T5.

No participant used S3 for the T5 “while”.

T6. (2a) children tended to mostly represent the two T6 events diagonally (33%) and with two separate scenes (61%). They opted for representing the two events depicted as physically separate, and distant on the paper sheet. These participants also tended to include a common background (50%) and to use the S3 strategy (44%). Such results seem to contrast with those of their (2b) peers on T5, and are more similar to those on T5 of the younger children, namely, (1a) children. Moreover, (2a) children, with low G scores, tended to adopt S2 more than their peers with high G scores: $\chi^2(1)=3,53, p=.06$ in the drawings by the former kind of children.

As for (1a) and (2b), children with high G scores preferred the horizontal representation, $\chi^2(1)=4,22, p<.05$. No relations between their working memory level (high or low) and their drawing strategies emerged.

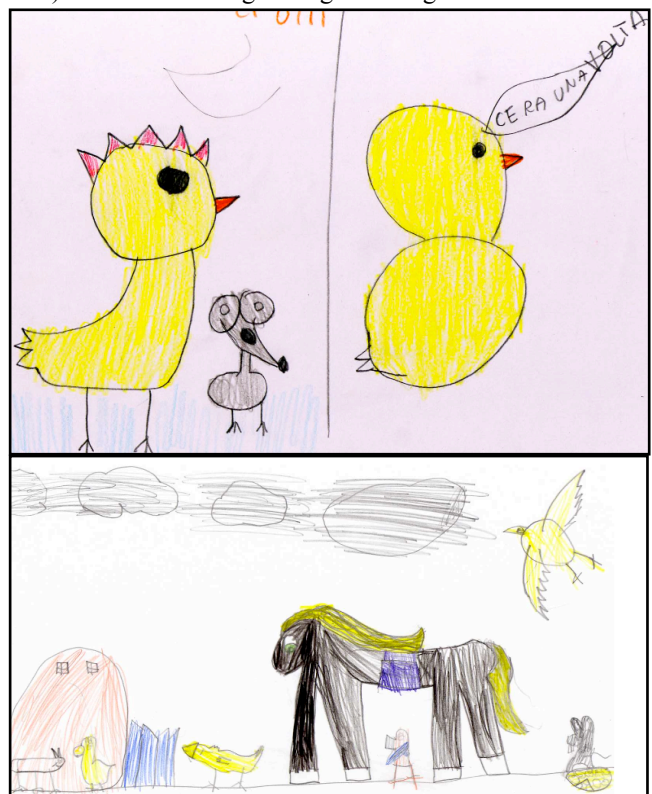


Figure 3. The top drawing for T5 is by a (2b) child, and the bottom one for T5 is by a (1a) child.

Discussion

Let us revisit the two main goals of our evaluation—see section “Goals and Rationale of Our Evaluation”.

Our results show that class and grammatical comprehension skills (DS and G scores) can have an impact on the child's visualisation strategies. Older children (in second-year classes) tended to represent "while" NRTRs between two events more abstractly (e.g. more frequent omission of the background). The visual pattern of their drawings seemed more conventional, with a dominant horizontal orientation. When the "while" NRTR is more complex as in T6 (it is vague, in that it corresponds to the Allen relation "during or finishes", and it happens in two different locations), second graders opted for more explicit drawing strategies, e.g., the introduction of a common background and a link between the two events. Grammatical skills seem to be important in discriminating those children that can be in trouble in representing the NRTRs of a story. For instance, children with low G scores in our study tended to represent a single event, when asked to depict NRTRs between two events. This result is confirmed across the different types of tasks. Contrary to our expectations, verbal working memory did not affect the children's performances at any level.

Let us now turn to the second goal of our exploratory study. This shows that linear representations are frequently used by first and second graders. However, the visual representation of the "while" between events also seems to depend on the type of events, and the type of Allen relation that the "while" corresponds to, as T6 indicates. However, other studies are needed to assess such hypothesis.

According to our preliminary findings the use of drawings offers a rich set of information, not only about the patterns of visual representations that are most common among children, but also on the strategies that children spontaneously adopt to render complex temporal relations.

Future studies will try to extend our study to a greater number of children, to a richer variety of "while" NRTRs, and consider how adults, that are expert text comprehenders, would represent the same NRTRs.

Conclusions

This paper presents and discusses our experimental study with 6–8 old novice readers, first and second graders of an Italian primary school. Our analyses reveal interesting common strategies that children employ for visually representing "while" temporal relations of a story. In particular, literacy maturity seems to play a relevant role in discriminating the sophistication and abstraction of the children's visual representations. Future work will continue and deepen this kind of experiment, and its implications on the design of our temporal reasoning web tool for novice readers.

References

- Allen, J.F. Maintaining Knowledge about Temporal Intervals. *ACM Comm.*, 26:832–843, 1983.
- Arfé, B., Gennari, R., and Mich, O. Evaluations of the LODE Temporal Reasoning Tool with Hearing and Deaf Children. *MCES2009*. Forthcoming.
- Chittaro, L. and Combi, C. Representation of Temporal Intervals and Relations: Information Visualization Aspects and their Evaluation. In *Proc. of TIME 2001*, pp. 13-20, IEEE, 2001.
- Di Mascio T., Catarci T., Santucci G., Dongilli P., Franconi E., and Tessaris S. Usability Evaluation in the SEWASIE project. In *Proc. of HCI2005*, L.E.A 2005.
- Di Mascio, T., & Gennari, R. (2009). Spatial Representations of Temporal Relations for Primary-School Teachers. In *Proc. of MBR 2009*. Forthcoming in 2009.
- Duran, N.D., McCarthy, P.M., Graesser, A.C. and McNamara, D.S. Using Temporal Cohesion to Predict Temporal Coherence in Narrative and Expository Texts. *Behavior Research Methods*, 2007.
- Ge, F. and Xuehong, T.. Temporal Reasoning on Daily Events in Primary School Pupils. *Acta Psychological Sinica*, 34:604–610, 2002.
- Gennari, R. and Mich, O. Constraint-based Temporal Reasoning for E-learning with LODE. In *Proc. of CP 2007*, Springer, 2007.
- Gennari, R., Di Mascio, T. and De Gasperis, G. Time Representation and Reasoning for a Story-telling Web Tool—the State of the Art. In *Proc. of ICCE 2009*. Forthcoming.
- Gunthorp, K., and Cassinelli, A. *Gino il Pulcino e Altre Storie*. Giunti, 2002.
- Johnson-Glenberg, M.C. Web-based Training of Metacognitive Strategies for Text Comprehension: Focus on Poor Comprehenders. *Reading and Writing*, 18:755–786, Springer, 2005.
- Hibino, S. and Rundensteiner, E.A. User Interface Evaluation of a Direct Manipulation Temporal Visual Query Language. In *Proc. of the ACM Multimedia Conference*, pp. 99-107, ACM, 1997.
- Koerber, S. and Sodian, B. Preschool Children's Ability to Visually Represent Relations. *Developmental Science*, 11(3):390–395, 2008.
- McColgan, K. and McCormack, T. Searching and Planning: Young Children's Reasoning about Past and Future Event Sequences. *Child Development*, 5:1477–1497, 2008.
- LODE. Retrieved August 27, 2009 from <http://lode.fbk.eu>
- Orsolini, M., Fanari, R., Cerracchio, S. and Famiglietti, L. Phonological and Lexical Reading in Italian Children with Dyslexia. *Reading and Writing*. In press.
- Paivio, A. Dual-coding Theory: Retrospect and Current Status. *Canadian Journal of Psychology*, 45:255–287, 1991.
- Rustioni, D., and Lancaster, M. Prove di Valutazione della Comprensione Linguistica. *Organizzazioni Speciali*, 1994.
- Scott, C. L. The Development of Some Working Time Concepts in Pre-school Children. Unpublished honours thesis, Foundation of Child and Youth Studies, Sydney, 1989.
- Winkel, H. The Acquisition of Temporal Reference Cross-linguistically using Two Acting-out Comprehension Tasks. *Journal of Psycholinguistic Research*, 33(4):333–355, 2004.