

First- to second-language reading comprehension: not transfer, but access¹

Catherine Walter *Institute of Education, University of London*

The 'transfer' of reading comprehension skills from a first language (L1) to a second language (L2) has long been discussed in the literature. This study challenges the transfer metaphor, proposing instead a notion of access. Studies based on Gernsbacher's Structure Building Framework (SBF) show that reading comprehension draws on general, amodal cognitive processes. It follows that L1-literate learners of an L2 already have comprehension skills: their need is to access these skills from the L2. To examine whether the SBF predicts L2 readers' performance, two groups of French learners of English performed an anomaly detection task. Results corresponded to the predictions of the SBF. The difficulty of lower-intermediate readers in accessing their comprehension skills was also linked to capacity problems in L2-based working memory.

Keywords: reading comprehension, language learning, working memory

Le 'transfert' de la compétence de compréhension de l'écrit de la première langue (L1) à la seconde (L2) est étudié depuis longtemps. Cet article conteste la métaphore de transfert, proposant à sa place une notion d'accès. Plusieurs études basées sur le Système de Construction de Structures (SCS) de Gernsbacher démontrent que la compréhension de l'écrit se base sur des processus non liés à une mode d'appréhension spécifique. Il s'ensuit que les apprenants sachant lire en comprenant en L1 doivent déjà posséder la faculté de compréhension; il s'agit d'y accéder à partir de la L2. Pour examiner si le SCS prédit la performance des apprenants de L2, deux groupes d'apprenants d'anglais de langue maternelle française exécutèrent une tâche de détection d'anomalies. Les résultats correspondent aux prévisions du SCS. De plus, la difficulté des apprenants du niveau pré-intermédiaire à accéder à leur compétence de compréhension a été reliée à des problèmes de capacité de la mémoire de travail basée en L2.

Mots-clefs: compréhension de l'écrit, apprentissage des langues, mémoire de travail

A reading problem or a language problem?

In 1984, Charles Alderson asked whether L2 reading comprehension difficulties were a reading problem or a language problem; researchers have been addressing this question ever since.

Part of the response to Alderson's question has centred round learners' progress. At an L2 proficiency level commonly designated as intermediate, learners have difficulty comprehending L2 texts graded at their level. As proficiency grows, comprehension ability catches up. In Walter (2004), I demonstrated that this phenomenon corresponds to a discontinuity between sentence-by-sentence processing of L2 text and whole-text processing: intermediate L2 learners who were proficient L1 comprehenders understood sentences in L2 texts but did poorly on overall comprehension of the same texts. This discontinuity has been corroborated in L1 by a functional magnetic resonance imaging (fMRI) study where text comprehension is associated with increased neural activity in the right frontal lobe, and sentence-level reading processes with increased activity in the left frontal lobe (Robertson et al. 2000). Therefore, reading comprehension is not the additive result of understanding one sentence after another; readers can understand each sentence without understanding the text.

When L2 learners do understand L2 texts, it has been said that they *transfer* reading comprehension skills from their L1 to their L2 (see e.g. Clarke 1988; Koda 1988, 2005; Bossers 1991; Carrell 1991; Bernhardt and Kamil 1995; Lee and Schallert 1997; Walter 2004). Here I propose that *transfer* is a misleading metaphor, and that it is better to speak of *access* to an already existing, non-linguistic skill. This argument is based on Gernsbacher's Structure Building Framework (1990, 1991, 1995, 1997). Here I will report the results of a study which tests the applicability of the Structure Building Framework to L2 reading comprehension.

Following the convention of writing about comprehension skill 'transfer', advice for language teachers assumes that L2 learners with comprehension difficulties need help in transferring these skills. For example, if learners cannot recognise main ideas in an L2 text, it is assumed they need to be taught this; advice on "teaching the specific reading skill of main idea comprehension" (Anderson 2002) is abundant in the literature (e.g. National Capital Languages Resources Center 2003, 2004; Cavallera and Leiguarda de Orué 2006; Han 2003–04; Sheorey and Edit 2004). However, if what happens is not transfer, but access, learners need not be taught this skill; rather, they need help in gaining access to it from the L2.

Another response to Alderson's question has focused on research into the relative contributions to L2 comprehension of L2 language proficiency and L1 reading comprehension ability. Bernhardt and Kamil (1995), in their own studies and by reviewing Markham (1985), Barnett (1986), and Shohamy (1984), established that upwards of 20% of the variance in L2 reading skill was accounted for by L1 reading skill, and upwards of 30% of the variance was accounted for by L2 proficiency. They also supported Carrell's (1991) finding that these proportions were not stable, with more variance accounted for by L1 reading ability and less by L2 proficiency as L2 proficiency increased. Taillefer (1996) went on to investigate different types of reading task and found that L2 proficiency accounted for more variance in more complex tasks.

LaBerge and Samuels (1974) proposed a limited attentional capacity for L1 reading: beginning L1 readers would use so much of their attentional

capacity for decoding that they could only periodically switch attention to comprehension. This hypothesis has been applied to the problems of L1-literate L2 readers. It is consistent with Bernhardt and Kamil's (1995) findings, and Bossers (1991) has characterised it as a 'short-circuit' phenomenon. Here I will once again support a limited capacity hypothesis, but I propose that the capacity in question is neither attentional nor intentional. Rather, I will argue that what constrains access to text comprehension skill is the capacity of L2-based verbal working memory, which is not accessible to consciousness.

Why access?

Models of reading comprehension commonly include the concept of building mental representations of text: some examples are Kintsch's (1998) Construction-Integration Model; van den Broek and colleagues' Landscape Model (van den Broek et al. 1998; Linderholm et al. 2004); and Just and Carpenter's (1992) Collaborative Activation-Based Production System. Among these, Gernsbacher's Structure Building Framework (henceforth SBF) is supported in detail by a large body of empirical work (Gernsbacher 1990, 1991, 1995, 1997, Gernsbacher et al. 2004). (For a detailed comparison of the SBF with other accounts of mental representations of text, see Gernsbacher 1995.)

The work of Gernsbacher and colleagues has supported the hypothesis that listening and reading comprehension are based on general cognitive processes and mechanisms. These are not specific to listening or reading (MacDonald and MacWhinney 1990), or even to language at all (Gernsbacher, Varner and Faust 1990). They are the same whether comprehension is based on reading, listening, looking at picture stories or watching silent filmed stories (Gernsbacher 1990). Comprehension is a general cognitive skill, working in the same way regardless of the mode of perception. Evidence for this amodality comes from Gernsbacher, Varner and Faust (1990), who gave 270 college students a battery of tests of reading, listening and picture-story comprehension: correlations for comprehension between the three modes were high, and factor analysis resulted in only one factor, suggested as general comprehension skill. So skill in comprehending texts is not a linguistic skill; rather, it is a general cognitive skill developing at the same time as the L1, but independently from it. It follows that the metaphor of 'transfer' of L1 comprehension skill to the L2 is misleading: what happens is more appropriately characterised as *access*, via L2 text, to the individual's already-established, amodal comprehension skill.

The next section provides a brief description of the SBF.

The Structure Building Framework

The three processes in the SBF are:

- 1) *laying a foundation* for a mental structure
- 2) *mapping* new information onto the developing mental structure
- 3) *shifting* to build a new substructure.

These processes are largely automatic and unconscious. The “building blocks” in the processes are what Gernsbacher calls “memory nodes” (1997: 267), in a parallel distributed processing approach (Rumelhart, McClelland and the PDP Research Group 1986). Memory nodes are activated by (1) information in the input, (2) the comprehender’s world knowledge and (3) (for written or spoken text) the comprehender’s language knowledge.

The first process, *laying a foundation*, occurs at the beginning of a unit, which in reading comprehension might be a clause, a sentence, a paragraph or an episode. Laying a foundation is triggered by the first element in the unit: for example, the first word in a sentence, whether written (Aaronson and Ferres 1986) or spoken (Cutler and Foss 1977); the first sentence in a story (Kieras 1981); or the first picture in a picture story (Gernsbacher 1990). Laying a foundation is qualitatively different from the following processes. Evidence for this is that the first element takes longer to process and is recalled better than later elements in the same unit.

The second process in the SBF is the *mapping* of new information onto a developing structure: if new information is coherent with the information in the ongoing structure, it activates the same as well as connected memory nodes. Coherence can come from matching in reference, time, location or cause (Gernsbacher 1997); matching elements take less time to process and to recall than elements that do not match.

The third process, *shifting*, takes place when, for the comprehender, the new information is not sufficiently coherent with ongoing structure information, so that other memory nodes are activated and a new substructure is initiated. This substructure is linked with the foundation but is not linked to the ongoing substructure. Gernsbacher (1985) demonstrates that this is a more adequate explanation for forgetting surface features of recently comprehended information than memory limitation or recoding; and Haenggi, Kintsch and Gernsbacher (1995) show that slowness in remembering objects in rooms left by a character in a narrative can be attributed to shifting prompted by location change (even change to an unspecified new location).

Skilled structure building results in networked, hierarchical structures with few main substructures. Less-skilled comprehenders, who fail to suppress irrelevant aspects of new information, initiate many new substructures, and so their mental structures are “bulkier [and] less cohesive” (Gernsbacher 1990: 213).

The SBF predicts that a major difference between the networked, hierarchical structures of skilled comprehension and the bulkier, less cohesive structures of less-skilled comprehension is the ease with which the comprehender can recall main ideas. This is not because of deliberate remembering, but because of the accessibility of information for recall. Accessibility depends on the

information's level of activation. Information in a current substructure is more highly activated than information in previous substructures; but outside a current substructure, information whose activation has been enhanced by repeated activation from related nodes is more accessible than information that has not received such enhancement. In a hierarchical structure built by a skilled comprehender, nodes representing the main ideas in a text will be frequently enhanced as new ideas are mapped onto the structure, and therefore will be more highly activated and more easily recalled than subsidiary ideas. Conversely, in the flat structures built by less-skilled comprehenders, few points outside the current substructure are very highly activated, so few will be easily accessible for recall.

Language-based comprehension uses the same mechanisms as comprehension based on other input modes. However, it also has some particularities: foundation laying, mapping and shifting are "influenced, or even to a large extent determined, by the linguistic characteristics of the text or discourse" (Sanders and Gernsbacher 2004: 80). This does not mean that all skilled comprehenders' mental representations of a given text are identical, but that there are resemblances between them, based on the text's linguistic characteristics.

Another special characteristic of language-based comprehension is that it can be more or less intentional: for L1 texts, "[i]nterpreting coherence clues can feel relatively unconscious or relatively deliberate" (Gernsbacher 1995: 278). I will now consider this point in regard to the methodology of the present study.

Probing structure building processes in reading comprehension

Because language-based comprehension processes can be more or less conscious, methods for studying them need to probe both conscious and unconscious activity. Think-aloud protocols and stimulated recall will only capture conscious activity, while anomaly detection (recognising when a later statement in a text contradicts an earlier one) captures the products of both types of activity. In addition, anomaly detection, unlike think-aloud protocols, shares with stimulated recall the advantage of minimal interference with the activity of reading.

Anomaly detection has been used to study the comprehension of L1 readers who decode well (reading aloud fluently) but who are poor comprehenders (e.g. they do poorly at standard measures of comprehension, as in Neale 1989). Garner and colleagues show that while good L1 comprehenders detect anomalies in narratives, poor comprehenders with no decoding problems systematically fail to do so (Garner 1980, 1981; Garner and Kraus 1981–82; Garner and Reis 1981). In Garner and Taylor (1982), 80% of poor comprehenders failed to mention problems with meaning, even when encouraged to criticise narratives containing gross anomalies. Yuill and Oakhill (1991: 125–35) gave

anomalous narratives to 7- to 8-year-old L1 readers who were skilled decoders, but either good or poor comprehenders. The good comprehenders detected anomalies more often than the poor ones. These results are consistent with the SBF: a coherent mental structure allows the reader to perceive when a new element contradicts the information in that structure.

Anomaly detection also indicates whether the comprehender distinguishes reliably between more and less important concepts in text. The SBF predicts that main-point anomalies, which contradict highly activated concepts in a structure, will be more easily detected than subsidiary-point anomalies, which contradict less highly activated concepts. This hypothesis may seem intuitively obvious. However, Yuill and Oakhill's L1 results (1991: 125–35) were inconclusive in this regard. Their L1 English 7- and 8-year-olds, who were all good decoders but either good or poor comprehenders, tended to score lower on subsidiary-point anomalies, but the effects were not significant for either group.

The present study will examine whether young people further on in their L1 reading careers distinguish between main-point and subsidiary-point anomalies. If these readers reliably detect main-point anomalies better than subsidiary-point anomalies in L1 texts, this confirms in detail the predictions of the SBF. Suppose that the same readers cannot reliably detect anomalies in level-appropriate L2 texts, and/or that their detection of main-point and subsidiary-point L2 anomalies does not differ. This cannot be because they are unskilled at structure building: they will have demonstrated this skill with L1 texts. However, it can indicate that they are failing to access their structure building skill when reading L2 texts.

Working memory and L2-based reading comprehension

If L2 learners demonstrate structure building skill when reading L1 texts but not when reading level-appropriate L2 texts, what is the reason? One explanation may be that the limited capacity of L2-based verbal working memory (WM) is highly solicited for sentence processing, with little spare capacity available for the higher-level process of structure building.

The involvement of WM in L1 reading comprehension is supported in the literature (Daneman and Carpenter 1980; Baddeley et al. 1985; Engle, Cantor and Carullo 1992; Ruffman 1996; Ericsson and Delaney 1999). Yuill and Oakhill (1991) working on L1 English and Cornoldi, De Beni and Pazzaglia (1996) working on L1 Italian compared children who were all good L1 decoders but either good or poor comprehenders as measured by standard tests; the poor comprehenders had lower WM capacity. Swanson and Berninger (1995), studying skilled and less-skilled L1 child readers, found that low WM correlated with poor reading comprehension. These L1 studies suggest the possibility of a WM capacity problem for the L2 reader when L2-based verbal WM capacity is highly solicited by sentence-level processing in

the L2. If this is the case, correlations between WM and structure building ability will be stronger for the L2 than for the L1, and stronger for learners with poor L2 reading comprehension.

In Walter (2004), a study of lower- and upper-intermediate French learners of English, I found no correlation between L1-based verbal WM and comprehension of L1 texts for either group. However, there was a significant correlation between L2-based verbal WM and comprehension of L2 texts for the lower-intermediate group. Small differences in L2-based WM in this group predicted significant differences in the ability to comprehend L2 texts.

The present study adopts the multiple-component model of WM (Baddeley and Logie 1999; Baddeley 2000). While detailed descriptions of WM differ, there is more consensus than disagreement on major issues, as shown in Miyake and Shah (1999). Not only for Baddeley and his colleagues, but also for other influential models (Anderson, Reder and Lebiere 1996; Conway and Engle 1994; Cowan 1995; Ericsson and Kintsch 1995; Howes and Young 1997; Kieras and Meyer 1995; O'Reilly, Braver and Cohen 1999; Schneider 1999; Teasdale and Barnard 1993), WM is a limited-capacity system of mental mechanisms for accessing, processing and temporarily storing the information needed for complex cognitive tasks. Commonly used measures comprise activities that progressively increase demands upon both processing and storage, putting these two into competition for the limited capacity of the system.

The study

Research questions

This study examined whether the L1 and L2 reading comprehension abilities of adolescents and young adults corresponded in detail to the predictions of the SBF, especially as regards the building of a hierarchical representation of a text and its effect on differential recall of main and subsidiary ideas; and whether performance in the L2 corresponded to L2-based WM capacity. The research questions were:

- 1) Do adolescent and young adult readers reliably distinguish between main points and subsidiary points in L1 narratives, as predicted by the SBF?
- 2) Does the pattern of sensitivity to main and subsidiary points in the L1 and the L2 support the hypothesis that lower-intermediate L2 learners lack access to structure building capacities that they demonstrate in their L1?
- 3) Is any main-point effect persistent over different story lengths?
- 4) Can L2-based verbal WM capacity predict relative success in L2-based structure building amongst lower-intermediate L2 learners?

Method

Participants

A cross-sectional study was designed, with two groups varying in L2-based comprehension skill. All participants had the same L1 and educational and sociocultural backgrounds: differences in age and educational attainment were taken into account in the design and analysis of the study.

The groups in the study were two English classes from middle and upper state secondary schools in a provincial French market town. The learners lived in the town or nearby villages; the range of socioeconomic categories was similar in the two groups. Four potential participants, two from each group, were eliminated after initial interviews, since they were bilingual from birth and spoke both languages daily. Thus, all participants were L1 French speakers from monolingual households who spoke no other languages regularly outside school.

The lower-intermediate (henceforth LowInt) group comprised 20 final-year middle school participants (12 girls, 8 boys) with a mean age of 14:7 years (SD 8 months, range 13:6 to 15:11 years; this age range is not unusual in France – see Royal 1999). The class was not streamed for English and included learners of all abilities, but not learners with special educational needs. The participants were in their fourth year of studying English (3 × 50 minutes per week) and in their second year of another foreign language.

One participant from the upper-intermediate (henceforth UpInt) class left school for psychosocial reasons near the beginning of the study; this left 22 final-year upper school participants in the group (18 young women, 4 young men). The mean age was 17:10 years (SD 9 months, range 16:11 to 19:6 years). This was their seventh year of English classes, and they had studied/were studying at least one other foreign language. The class was not streamed for English. It was composed of students who had chosen to specialise in literary subjects for their final examinations.

Both groups took part in four other studies besides the present one. Participants and their bilingual classmates received presents at Christmas and book vouchers at the end of the series of studies.

Materials

The materials that participants would read were versions of the same narratives in English and French, with most containing different sorts of anomalies. Initially, 19 non-anomalous stories at the proficiency level of the LowInt group were written in English by the author (who has 20 years' experience of writing published materials for teaching English as a foreign language). The French versions of the stories, translations by the author from the English versions, were checked for accuracy and naturalness by two

educated French informants. Seven of the stories were approximately 100 words long (mean = 107 words, SD = 3), and six each were approximately 200 words (mean = 208, SD = 6) and 300 words (mean = 301, SD = 6). The topics of the stories were appropriate to an adolescent/young adult audience. Nine stories had female protagonists and ten had male protagonists. The stories did not favour girls' or boys' preferred narrative styles (Millard 1997). They were written to the lexical and syntactic level of the LowInt group, and the French versions maintained this simple level of lexis and syntax.

In order to determine what to consider as 'main' and 'subsidiary' points in each story, fourteen L1 English-speaking adults (8 women, 7 men, aged 19 to 62) and fourteen L1 English-speaking adolescents (9 girls, 5 boys, aged 13 to 16, who each received UK£5) read the stories and answered a questionnaire about each (Appendix 1). The adolescents also assessed the stories in terms of interest and unanimously rated all the stories as good to excellent. A story element was classified as a main point if it was spontaneously mentioned in summaries by at least 85% of the adults and 85% of the adolescents and had a mean rating of at least 4 (out of 5) in the second part of the questionnaire, with no rating lower than 3. A story element was classified as a subsidiary point if it received a mean rating of 2 or less.

Eighteen of the stories were altered to produce three anomalous versions of each (see Appendix 2 for the three anomalous L2 versions of one story):

- 1) a version with an anomaly contradicting a nearby main point (Main Near condition), separated from the contradicted precondition by seven or fewer clauses.
- 2) a version with an anomaly contradicting the same main point but occurring at least five clauses further from the contradicted point (Main Remote condition). The SBF predicts that good comprehenders will remember main points well wherever they occur, because of repeated activation from new points as these are integrated into the structure.
- 3) a version with an anomaly contradicting a subsidiary point remote from the contradicted point (Subsidiary Remote condition). The SBF predicts that good comprehenders will recall main points better than subsidiary points because of differences in activation. The Main Remote and Subsidiary Remote anomalies were located at approximately the same distance from the contradicted point.

No anomaly was introduced into the nineteenth story (the fourth story in the series for all participants) in order to discourage participants from false positive anomaly identification.

To control for level, the non-anomalised L2 versions and the extra sentences from the anomalous versions were read by a class of adolescents in a Paris school in the same year as the LowInt group, and they reported no difficulties in comprehension at any level. The English teacher of the LowInt group also judged the L2 versions appropriate to the proficiency level.

Shortly before the trial, the participants practised in class the few words that might cause problems for less proficient learners.

To summarise: there were 18 base stories, and each of these was configured in three different anomalous versions, for a total of 54 anomalous texts; additionally, there was one non-anomalous story. Each text was presented in English and in French. Thus, there were 110 texts in total.

The texts were entered into the Psyscope© programme (Cohen et al. 1993) on a Macintosh PowerBook, one clause or long clause constituent per screen.

Design and procedure

All participants read all 19 stories, but each participant read each anomalous story in only one anomaly condition and in only one language, plus the non-anomalous story in one language. Participants began either with an L2 story or with an L1 story. They read two stories in one language before changing to the other language, with changes clearly signalled in advance. All participants began with a series of seven stories: six anomalous 100-word stories plus the non-anomalous 100-word control story (always the fourth story), and then went on to the 200-word and 300-word series. The presentation of each series of six same-length stories was counterbalanced for language, for the three anomaly conditions, and for order. Each story appeared an approximately equal number of times in each condition across the whole study.

Trials were conducted with participants individually. After a spoken L1 explanation, including a sample anomalous story, and before the trials, each participant practised on the computer with an L1 version of the Little Red Riding Hood story containing each kind of anomaly. Participants were asked to watch out for contradictions in the story and to press a given key on the computer when they found one. The story also included a surprise for a character (Little Red Riding Hood sees the wolf) and the traditional in-character lie by the wolf. This was so that students understood that they were looking neither for surprises in the story nor for lies by characters acting consistently, but only for internal anomalies. Participants were told that most, but not all, of the stories would contain contradictions. So that these young participants did not finish the session with a sense of failure, they were also told that they would probably not detect every contradiction, as we humans have been telling stories for a long time, and our strong tendency is to want stories to make sense.²

Stories appeared on the computer screen one clause or one long clause constituent at a time. The participant controlled the passage to the next screen via a keystroke on the computer but could not return to a previous screen. If a double keystroke caused a skipped screen, the researcher said what had been on the skipped screen or showed it on paper; this happened for only four of the approximately 23,000 screens in the study. Before each text,

a screen announced which language it was in. Participants were encouraged to ask at any time the meanings of words or expressions that they did not understand. When this happened (six times, four participants), the L1 meaning was given; no questions were asked within an anomalous sentence.

Participants pressed a different key from the 'next screen' key when they detected an anomaly and told the researcher what the anomaly was. Answers were noted on a grid unseen by the participant. Participants who did not detect the anomaly while reading from the screen were invited to read the full text on paper. In order for them to complete the trial with perceived success, they had up to six chances to detect each anomaly; but only the first three types of response, on a decreasing scale of 3 points to 1, were scored.³ Participants could detect the anomaly:

- 1) on-screen (3 points)
- 2) while reading the following screen (2 points)
- 3) not while reading on-screen, but spontaneously when re-reading the text on paper (1 point)
- 4) not during initial re-reading on paper, but when shown the portion of the text containing it (0 points)
- 5) when given a clue (0 points)
- 6) not at all, in which case it was pointed out and explained (0 points).

This procedure approximated natural reading: in the first reading, participants read the text straight through without returning to earlier parts, as readers do when they do not encounter difficulties; in the second reading from paper, they were able to return to earlier parts of the text to check their understanding.

Participants were eager to know what the anomalies were; progressively more explicit prompts gave them several chances to discover the anomalies for themselves and complete the session with perceived success. There were only five (out of a possible 798) instances where a participant had to be told what the anomaly was.

Working memory measure

The verbal working memory (WM) measure used in this study was identical to the measure described in detail in Walter (2004), based on Waters and Caplan (1996). Trials were conducted individually before the anomaly detection trial, using the Psyscope© programme (Cohen et al. 1993) on a Macintosh PowerBook. Participants read sets of successively longer series of short simple sentences ending in concrete nouns, beginning with series of two sentences and continuing until they could no longer perform the concurrent tasks demanded of them. These concurrent tasks were: (1) judging whether each sentence was 'logical' (in 'illogical' sentences, the argument

requirements of the verbs were violated, e.g. *The sofa jumped on the dog*); and (2) remembering the last word in all the sentences in a series until the end of that series. Participants attempted a set of five series in each language at each series length: five two-sentence series, five three-sentence series and so on. They stopped when they could not recall all the sentence-final words in at least three series of a set. Waters and Caplan (1996) argue that, while the classic Daneman and Carpenter (1980) reading span measure *requires* both processing and storage, it only *measures* storage; an instrument that measures both is more valid. In this study, to minimise differences attributable to some students adopting a rehearsal strategy and others not, all participants rehearsed the list of sentence-final words subvocally between sentences. After a practice session, each participant performed the WM trial in both languages, alternating two series of sentences in one language with two series in the other. There were two versions of the materials with the same sentences appearing in different series; trials were counterbalanced for version and for order of languages.

WM scores were calculated using three components: (a) percentage of correct logicity judgments, (b) mean time taken to judge all correctly judged sentences, and (c) highest number of sentence-final words remembered correctly, in order, in at least three out of five attempts, with a half-point given for two correct out of five attempts. A set of standard scores was calculated for each component, using both L1 and L2 raw scores. The standard scores for mean time were multiplied by -1 (to align these with the other two components). Each participant's L1-based WM score was then calculated as the mean of the three L1 component standard scores, and each L2-based WM score as the mean of the three L2 component standard scores. This resulted in L1- and L2-based WM scores comparable on the same scale.

Results

Anomaly detection

The means and standard deviations for each anomaly detection condition are given in Table 1. Figure 1 shows the anomaly detection scores by anomaly condition. Reliability tested with Cronbach's alpha was .75.

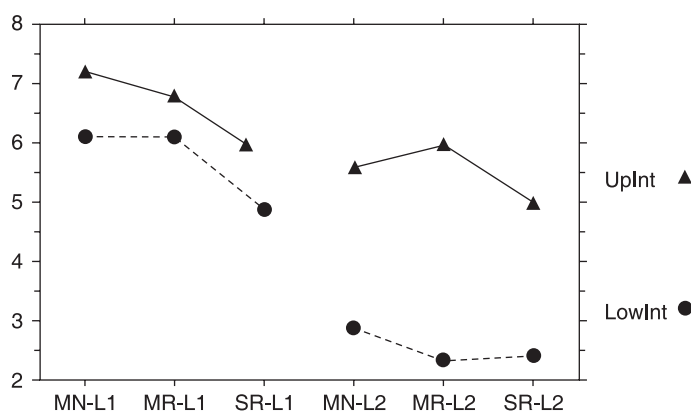
A $2 \times 2 \times 3 \times 3$ mixed-design ANOVA was carried out on the anomaly detection scores, with one between-subjects factor: group (LowInt, UpInt), and three within-subjects factors: language (L1, L2), story length (100, 200 or 300 words), and anomaly condition (Main Near, Main Remote, Subsidiary Remote). Confidence intervals in all ANOVAs = 95%.

All four main effects were significant. There was a main effect for group, with the UpInt participants detecting anomalies more efficiently than the LowInt participants: $F(1,41) = 26.76$, $p < .001$. There was a main effect for

Table 1. Mean anomaly detection scores, by anomaly condition, length and group. Scores: out of 3 for each story length, out of 9 for all lengths

| Anomaly condition | Story length (words) | Lang | LowInt Mean | SD | UpInt Mean | SD |
|-------------------|----------------------|------|-------------|-----|------------|-----|
| Main Near | 100 | L1 | 2.2 | 1.2 | 2.6 | 0.8 |
| | | L2 | 1.4 | 1.3 | 2.0 | 1.2 |
| | 200 | L1 | 1.9 | 1.3 | 2.0 | 1.2 |
| | | L2 | 0.6 | 1.0 | 1.7 | 1.2 |
| | 300 | L1 | 1.8 | 1.1 | 2.7 | 0.7 |
| | | L2 | 0.9 | 1.2 | 1.9 | 1.3 |
| | All | L1 | 6.0 | 2.6 | 7.3 | 1.7 |
| | | L2 | 3.0 | 2.5 | 5.6 | 2.0 |
| Main Remote | 100 | L1 | 2.6 | 0.8 | 2.6 | 0.8 |
| | | L2 | 1.2 | 1.2 | 2.2 | 1.1 |
| | 200 | L1 | 1.7 | 1.2 | 2.0 | 1.1 |
| | | L2 | 0.5 | 1.1 | 1.9 | 1.3 |
| | 300 | L1 | 2.0 | 1.0 | 2.1 | 1.1 |
| | | L2 | 0.6 | 1.1 | 1.9 | 1.2 |
| | All | L1 | 6.3 | 2.1 | 6.7 | 1.6 |
| | | L2 | 2.4 | 2.1 | 6.0 | 2.2 |
| Subsidiary Remote | 100 | L1 | 2.0 | 1.0 | 2.2 | 1.2 |
| | | L2 | 1.1 | 1.2 | 2.1 | 1.1 |
| | 200 | L1 | 1.5 | 1.4 | 1.7 | 1.3 |
| | | L2 | 0.8 | 1.2 | 1.9 | 1.3 |
| | 300 | L1 | 1.5 | 1.2 | 1.9 | 1.0 |
| | | L2 | 0.4 | 0.9 | 1.0 | 1.1 |
| | All | L1 | 5.0 | 2.3 | 5.7 | 2.0 |
| | | L2 | 2.2 | 2.4 | 5.0 | 1.8 |

LowInt = Lower Intermediate, UpInt = Upper Intermediate

**Figure 1.** Anomaly detection scores (out of 9), by language, anomaly condition and group (MN = Main Near, MR = Main Remote, SR = Subsidiary Remote)

language, with L1 anomalies being detected more efficiently than L2 anomalies: $F(1,41) = 73.85$, $p < .001$. There was a main effect for story length: $F(2,82) = 15.41$, $p < .001$; post-hoc Scheffé tests showed that anomalies in 100-word stories were detected more efficiently than those in 200- or 300-word stories ($p < .001$ in both cases), but that there was no difference in detection between the 200- and 300-word stories. Lastly, there was a main effect for anomaly condition: $F(2,82) = 6.00$, $p < .01$; post-hoc Scheffé tests showed that both Main Near and Main Remote anomalies were detected more efficiently than Subsidiary Remote anomalies ($p < .01$ and $p < .05$ respectively), but that there was no significant difference between the two main-point anomaly conditions.

There was a significant two-way interaction between language and group, with LowInt participants' anomaly detection deteriorating much more seriously in L2-based comprehension than did their UpInt counterparts': $F(1,40) = 19.57$, $p < .001$. No other interactions were found.

Working Memory measure

The WM (standard) scores are given in Table 2, and Figure 2 shows the WM scores by language and by group.

A 2×2 mixed-design ANOVA was carried out, with one between-subjects factor: group (LowInt, UpInt), and one within-subjects factor: language (L1, L2). Both main effects were significant. There was a main effect for group, with UpInt participants having higher WM scores than LowInt participants: $F(1,41) = 50.89$, $p < .0001$; this is not unexpected, as working memory develops at least until around age nineteen (Siegel 1994) and perhaps as late as age 45 (Swanson 1999). There was a main effect for language, with L1-based WM scores higher than L2-based scores: $F(1,41) = 256.92$, $p < .0001$. There was a significant two-way interaction between language and group, indicating that the LowInt participants' WM scores in the two languages were much further apart than the UpInt participants': $F(1,41) = 25.21$, $p < .0001$.

Pearson product-moment correlations were carried out between the French and English WM scores for each group. Significant correlations were found: LowInt $r = 0.45$, $p < .05$ and UpInt $r = 0.70$, $p < .001$. The correlations were compared using Fisher's (1921) method for testing the difference

Table 2. Verbal working memory scores (based on standard scores for the three components)

| | LowInt Mean | SD | UpInt Mean | SD |
|-------------|-------------|-----|------------|-----|
| L1-based WM | .35 | .25 | .73 | .26 |
| L2-based WM | -1.09 | .66 | -.02 | .30 |

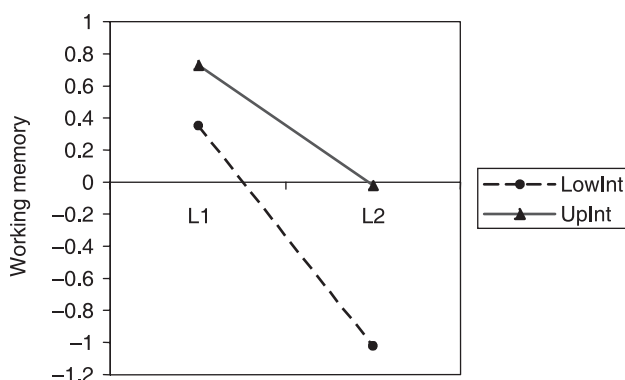


Figure 2. Working memory scores (composite standard scores) by language and by group

Table 3. Pearson product-moment correlations between overall verbal working memory scores and overall anomaly detection scores

| | L1 scores | L2 scores | LowInt | UpInt |
|-------------------|-----------|-----------|--------|-------|
| Main Near | .34 | .48 | .47 | .57 |
| Main Remote | ns | .57 | .64 | ns |
| Subsidiary Remote | ns | .60 | .55 | ns |

($p < .001$ for numbers in *bold italics*; $p < .05$ for others)

between two independent r s and were shown to be significantly different ($p < .001$), indicating that the correlation between L1 and L2 scores was significantly weaker for the LowInt group than for the UpInt group.

In order to examine the overall relationship between WM and anomaly detection, both L1- and L2-based WM scores were plotted on the x axis of the same graph, with the means of the corresponding anomaly detection scores on the y axis. The resulting Pearson product-moment correlations are given in Table 3; they are stronger for L2 than for L1, and stronger for the LowInt than the UpInt group.

Discussion

L1 anomaly detection

Both groups performed well in L1 anomaly detection, with no significant difference between their scores. The scores were not at ceiling: participants did not detect all anomalies, indicating that the cognitive level of the task

was appropriate. These results confirm the finding in Walter (2004) that the LowInt learners were competent comprehenders in their L1. They also mean that any differences in L2-based comprehension cannot be attributed to maturity or intellectual development.

Performance in the L1 also confirms the non-significant tendency found in Yuill and Oakhill's (1991) study of younger readers for main-point anomalies to be more efficiently detected than subsidiary-point anomalies. In the present study, the main-point effect was significant for both groups, regardless of the distance between the contradicted precondition and the anomaly. These data provide additional support for the SBF and confirm the hypothesis based upon it: that competent structure building allows adolescents and young people reliably to differentiate between main and subsidiary points in their L1.

L2 anomaly detection

Both groups detected main-point anomalies better than subsidiary-point anomalies in the L2. However, both groups' performance in anomaly detection deteriorated in the L2; and the LowInt participants' performance deteriorated much more seriously than their UpInt counterparts'. It is unlikely that this effect was due to sentence-level comprehension problems for the LowInt group: great care was taken to suit the level of the L2 texts to their proficiency level, and there is no reason to believe from piloting, the judgment of their teacher, or their behaviour during the trial that the LowInt participants had sentence-level problems. Nor do the members of this group have less well-developed comprehension skills than the UpInt group, as the L1 results demonstrate. Their performance cannot be attributed to conceptual difficulty, since the same stories were read in the L1 and the L2 (each participant reading each story in one language only). Rather, the results are consistent with the hypothesis that lower-intermediate learners, when reading texts in their L2, are unable to access the structure building skill that they deploy well in identical circumstances in their L1. It is a problem of access.

Story length

The hypothesis that longer narratives will pose more problems than shorter ones is not confirmed in a straightforward way. While anomalies in 100-word stories were detected more often than those in 200- and 300-word stories, by both groups and in both languages, there was no difference in detection rate between the 200- and 300-word stories, and the pattern of detection remained the same, with main-point anomalies being easier to detect, regardless of distance from the to-be-contradicted precondition. It may be that, even for skilled comprehenders, it is easier to cope with a very short text containing

only a few concepts. This would not be inconsistent with the SBF and might explain why the advantage of brevity disappears after a certain point. It would be interesting to compare texts of between, say, 100 and 500 words, to see whether the discontinuity between 100 and 200 words is unique.

Working memory and anomaly detection

In the present study, both groups showed evidence of efficient L1-based structure building. In the L2, on the other hand, there were significant differences in structure building between the groups. This cannot be accounted for by differences in the ability to distinguish main points (demonstrated by both groups); nor by differences based on story length (observed for both groups). However, differences in verbal L2-based WM scores, and differences in correlations between verbal L2-based WM and anomaly detection performance, correspond to the hypothesis that the capacity of verbal L2-based WM is inordinately taken up by lower-level processing for LowInt learners. In the UpInt group, even learners at the lower end of their group's L2-based WM scale had sufficient capacity to build hierarchical structures based on the L2 texts. In the LowInt group, in contrast, the L2-based WM scale was lower, and for this group individual differences in WM corresponded to differences in structure building. Additional evidence of the difference between L1-based and L2-based WM for the LowInt group was provided by the weaker correlation between the two WM measures for this group than for the UpInt group.

Further research is needed to clarify how L2-based verbal WM develops and what its relationship is to L2-based structure building. One possibility is that sentence parsing might not yet be sufficiently automatised to allow for higher-level comprehension processes to occur. This would correspond to observations for L1s that readers with lower WM spans have problems with some of the elements of skilled comprehension, for example the use of elaborative inferences (Oakhill and Yuill 1996; Perfetti, Marron and Foltz 1996).

Another possibility is that problems may arise from an insufficiently elaborated representation of the L2 phonological system, making it difficult to retain recently decoded material in the phonological loop. This is because, for alphabetic and syllabic languages, readers hold the most recently read portion of text in the phonological loop (Gathercole and Baddeley 1993). (Note that this is probably not the case for L1 readers of logographic languages: see e.g. Holm and Dodd 1996; Koda 1998; Jackson et al. 1999.)

Implications for the second language classroom

This article has argued for a change of metaphor for the progression from L1 to L2 reading comprehension, from a metaphor of transfer to one of access.

It has demonstrated that readers' performance in their L1 and L2 corresponds to the predictions of Gernsbacher's Structure Building Framework, and that lower-intermediate learners' problems with comprehension reflect problems of access to the comprehension skills that these learners possess.

The study has implications for L2 classrooms around the world where already-literate learners are embarking upon reading in a second language. It throws into serious question the advisability of spending classroom time in the teaching of comprehension skills to L1-literate learners.

So how might that classroom time be spent? If effortful sentence parsing is at the root of some of the difference between L1-based and L2-based WM at lower-intermediate level, activities which promote automatization of parsing may be part of the answer. There is also the intriguing possibility that texts of around 100 words may be the most beneficial ones for lower-intermediate extensive reading materials. The value of extensive reading has been demonstrated (see the excellent review in Day and Bamford 1998). The advantage for 100-word texts in the study suggests that texts of around this length might be used to help lower-intermediate learners to bootstrap their way to better L2 comprehension. Lastly, if the development of L2-based working memory is limited by the elaboration of a reliable mental representation of the L2 phonological system, pronunciation instruction may be part of the answer in helping L1-literate learners to deploy their comprehension skills in L2. Further research is needed in all these areas.

Notes

1. I am most grateful to Gillian Brown for introducing me to Morton Gernsbacher's work; to Michael Swan and John Williams for their continuing advice and support; to Dominique Flandin-Granget and Catherine Cobès for serving as linguist informants; to the members of the Working Memory Workshop, especially Alan Baddeley and Robert Logie, for their advice on working memory matters; and to Matthew Saxton, Jan Blommaert, Alison Mackey, Ken Hyland and two anonymous reviewers for their helpful comments on earlier drafts of this article. Thanks are also due to the young people who gave their time for the empirical study I have reported.
2. Note that even good comprehenders in L1 do not tend to achieve perfect scores in anomaly detection: comprehenders tend to adhere to Grice's (1975: 45) maxim of relevance, maintaining an assumption that successive parts of a text fit together as a coherent whole, even when these comprehenders are expecting inconsistencies (Garner 1980, 1981; Garner and Kraus 1981–82; Garner and Reis 1981).
3. Stricter and more liberal scoring systems gave similar results.

References

- Aaronson, D. and S. Ferres (1986) Reading strategies for children and adults: a quantitative model. *Psychological Review* 93: 89–112.

- Alderson, J.C. (1984) Reading in a foreign language: a reading problem or a language problem? In J.C. Alderson and A.H. Urquhart (eds.), *Reading in a foreign language*. Burnt Mill, Harlow: Longman. 1–24.
- Anderson, J.R., L.M. Reder and C. Lebiere (1996) Working memory: activation limitations on retrieval. *Cognitive Psychology* 30: 221–256.
- Anderson, N.J. (2002) The role of metacognition in second language teaching and learning. *CAL Digest* EDO-FL-01-10. <http://www.cal.org/resources/digest/0110anderson.html> (accessed 22 April 2006).
- Baddeley, A. (2000) The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences* 4.11: 417–23.
- and R. Logie (1999) Working memory: the multiple-component model. In A. Miyake and P. Shah (eds.), *Models of working memory: mechanisms of active maintenance and executive control*. Cambridge University Press. 28–61.
- R. Logie, I. Nimmo-Smith and N. Brereton (1985) Components of fluid reading. *Journal of Memory and Language* 24: 119–31.
- Barnett, M.A. (1986) Syntactic and lexical/semantic skill in foreign language reading: importance and interaction. *Modern Language Journal* 70: 343–9.
- Bernhardt, E.B. and M.L. Kamil (1995) Interpreting relationships between first language and second language reading: consolidating the linguistic threshold and the linguistic interdependence hypotheses. *Applied Linguistics* 16.1: 15–34.
- Bossers, B. (1991) On thresholds, ceilings and short-circuits: the relation between L1 reading, L2 reading and L2 knowledge. In J.H. Hulstijn and J.F. Matter (eds.), *Reading in two languages [AILA Review 8]*. Amsterdam: Free University Press. 45–60.
- Carrell, P.L. (1991) Second language reading: reading ability or language proficiency? *Applied Linguistics* 12.2: 159–79.
- Cavallera, I.V. and A.M. Leiguarda de Orué (2006) Developing the reading brain: promoting reading. *Modern English Teacher* 15.1: 5–13.
- Clarke, M.A. (1988) The short circuit hypothesis of ESL reading – or when language competence interferes with reading performance. In P.L. Carrell, J. Devine and D.E. Eskey (eds.), *Interactive approaches to second language reading*. Cambridge University Press. 114–24.
- Cohen, J.D., B. MacWhinney, M. Flatt and J. Provost (1993) Psyscope: a new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments and Computers* 25: 252–71.
- Conway, A.R.A. and R.W. Engle (1994) Working memory and retrieval: a resource-dependent inhibition model, *Journal of Experimental Psychology: General* 123: 354–73.
- Cornoldi, C. and J. Oakhill (1996) *Reading comprehension difficulties: processes and intervention*. Mahwah, NJ: Lawrence Erlbaum.
- R. De Beni and F. Pazzaglia (1996) Profiles of reading comprehension difficulties: an analysis of single cases. In C. Cornoldi and J. Oakhill (eds.), *Reading comprehension difficulties: processes and intervention*. Mahwah, NJ: Lawrence Erlbaum. 113–36.
- Cowan, N. (1995) *Attention and memory: an integrated framework*. New York: Oxford University Press.
- Cutler, A. and D.J. Foss (1977) On the role of sentence stress in sentence processing. *Language and Speech* 21: 1–10.
- Daneman, M. and P. Carpenter (1980) Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior* 19: 450–66.
- Day, R.R. and J. Bamford (1998) *Reading in the second language classroom*. Cambridge University Press.

- Engle, R.W., J. Cantor and J.J. Carullo (1992) Individual differences in working memory and comprehension: a test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory and Cognition* 18: 972–92.
- Ericsson, K.A. and P.F. Delaney (1999) Long-term working memory as an alternative to capacity models of working memory in everyday skilled performance. In A. Miyake and P. Shah (eds.), *Models of working memory: mechanisms of active maintenance and executive control*. Cambridge University Press. 257–97.
- and Kintsch, W. (1995) Long-term working memory. *Psychological Review* 102: 211–45.
- Fisher, R.A. (1921) On the probable error of a coefficient of correlation deduced from a small sample. *Metron* 1: 3–32.
- Garner, R. (1980) Monitoring of understanding: an investigation of good and poor readers' awareness of induced miscomprehension of text. *Journal of Reading Behavior* 12: 55–63.
- (1981) Monitoring of passage inconsistency among poor comprehenders: a preliminary test of the "piecemeal processing" explanation. *Journal of Educational Research* 74: 159–62.
- and C. Kraus (1981–82) Good and poor comprehender differences in knowing and regulating reading behaviors. *Educational Research Quarterly* 6: 5–12.
- and R. Reis (1981) Monitoring and resolving comprehension obstacles: an investigation of spontaneous text lookbacks among upper-grade good and poor comprehenders. *Reading Research Quarterly* 16: 569–82.
- and N. Taylor (1982) Monitoring of understanding: an investigation of attentional assistance needs at different grade and reading proficiency levels. *Reading Psychology* 3: 1–6.
- Gathercole, S.E. and A.D. Baddeley (1993) *Working memory and language*. Hove, UK and Hillsdale, NJ: Lawrence Erlbaum.
- Gernsbacher, M.A. (1985) Surface information loss in comprehension. *Cognitive Psychology* 17: 324–63.
- (1990) *Language comprehension as structure building*. Hillsdale, NJ: Lawrence Erlbaum.
- (1991) Cognitive processes and mechanisms in language comprehension: the Structure Building Framework. In G.H. Bower (ed.), *The psychology of learning and motivation*. New York: Academic Press. 217–63.
- (1995) The Structure Building Framework: what it is, what it might also be, and why. In B.K. Britton and A.C. Graesser (eds.), *Models of text understanding*. Hillsdale, NJ: Lawrence Erlbaum. 289–311.
- (1997) Two decades of Structure Building. *Discourse Processes* 23: 265–304.
- , R.R.W. Robertson, P. Palladino and N.K. Werner (2004) Managing mental representations during narrative comprehension. *Discourse Processes* 37.2: 145–64.
- , K.R. Varner and M. Faust (1990) Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory and Cognition* 16: 430–45.
- Grice, H.P. (1975) Logic and conversation. In P. Cole and J. Morgan (eds.), *Syntax and semantics* 3. *Speech acts*. New York: Academic Press. 41–58.
- Haenggi, D., W. Kintsch and M.A. Gernsbacher (1995) Spatial situation models and text comprehension. *Discourse Processes* 19: 173–99.
- Han, J.-M. (2003–04) A reading problem in secondary schools. *IATEFL Issues* 176: 9.
- Holm, A. and B. Dodd (1996) The effect of first written language on the acquisition of English literacy. *Cognition* 59: 119–47.
- Howes, A. and R.M. Young (1997) The role of cognitive architecture in modelling the user: SOAR's learning mechanisms. *Human-Computer Interaction* 12: 311–43.

- Jackson, N.E., H. Chen, L. Goldsberry, A. Kim and C. Vanderwerff (1999) Effects of variations in orthographic information on Asian and American readers' English text reading. *Reading and Writing: An Interdisciplinary Journal* 11: 345–79.
- Just, M.A. and P.A. Carpenter (1992) A capacity theory of comprehension: individual differences in working memory. *Psychological Review* 99: 122–49.
- Kieras, D.E. (1981) Component processes in the comprehension of simple prose. *Journal of Verbal Learning and Verbal Behavior* 20: 1–23.
- and D.E. Meyer (1995) Predicting human performance in dual-task tracking and decision making with computational models using the EPIC architecture. In D.S. Alberts, D. Buede, T. Clark, R. Hayes, J. Hofmann, W. Round, S. Starr and W. Vaughan (eds.), *Proceedings of the International Symposium on Command and Control Research and Technology*. Washington, DC: National Defense University. 314–25.
- Kintsch, W. (1998) *Comprehension. A paradigm for cognition*. New York: Cambridge University Press.
- Koda, K. (1988) Cognitive processes in second language reading: transfer of first language reading skills and strategies. *Second Language Research* 4.2: 133–56.
- (1998) The role of phonemic awareness in second language reading. *Second Language Research* 14: 194–215.
- (2005) *Insights into second language reading: a cross-linguistic approach*. Cambridge University Press.
- LaBerge, D. and S. Samuel (1974) Towards a theory of automatic information processing in reading. *Cognitive Psychology* 6: 293–323.
- Lee, J.-W. and D.L. Schallert (1997) The relative contribution of L2 language proficiency and L1 reading ability to L2 reading performance: a test of the threshold hypothesis in an EFL context. *TESOL Quarterly* 31.4: 713–39.
- Linderholm T., S. Virtue, Y. Tzeng and P. van den Broek (2004) Fluctuations in the availability of information during reading: capturing cognitive processes using the Landscape Model. *Discourse Processes* 37.2: 165–86.
- MacDonald, M.C. and B. MacWhinney (1990) Measuring inhibition and facilitation effects from pronouns. *Journal of Memory and Language* 29: 469–92.
- Markham, P. (1985) The rational deletion cloze and global comprehension in German. *Language Learning* 35: 423–30.
- Millard, E. (1997) *Differently literate: boys, girls and the schooling of literacy*. London: Falmer Press.
- Miyake, A. and P. Shah (eds.) (1999) *Models of working memory: mechanisms of active maintenance and executive control*. Cambridge University Press.
- National Capital Language Resources Center, Washington, DC. (2003, 2004) Teaching reading: developing reading activities. *The Essentials of Language Teaching*. <http://www.nclrc.org/essentials/reading/developread.htm> (accessed 22 April 2006).
- Neale, M.D. (1989) *The Neale analysis of reading ability – revised British edition*. Windsor: NFER-Nelson.
- Oakhill, J. and N. Yuill (1996) Higher order factors in comprehension disability: processes and remediation. In C. Cornoldi and J. Oakhill (eds.), *Reading comprehension difficulties: processes and intervention*. Mahwah, NJ: Lawrence Erlbaum. 69–92.
- O'Reilly, R.C., T.S. Braver and J.D. Cohen (1999) A biologically based computational model of working memory. In A. Miyake and P. Shah (eds.), *Models of working memory: mechanisms of active maintenance and executive control*. Cambridge University Press. 375–411.

- Perfetti, C.A., M.A. Marron and P.W. Foltz (1996) Sources of comprehension failure: theoretical perspectives and case studies. In C. Cornoldi and J. Oakhill (eds.), *Reading comprehension difficulties: processes and intervention*. Mahwah, NJ: Lawrence Erlbaum. 137–65.
- Robertson, D.A., M.A. Gernsbacher, S.J. Guidotti, R.R.W. Robertson, W. Irwin, B.J. Mock and M.E. Campana (2000) Functional neuroanatomy of the cognitive process of mapping during discourse comprehension. *Psychological Science* 11: 255–60.
- Royal, S. (1999) La mutation des collèges pour tous et pour chacun: vers un collège plus juste. Conférence de presse. Documents. *Débat national sur le collège*. <http://www.hebergement.ac-poitiers.fr/college/etudes/confpres.htm> (accessed 05 April 2005).
- Rumelhart, D.E., McClelland, J. L., and The PDP Research Group (1986) *Parallel distributed processing: explorations in the microstructure of cognition*. London: MIT Press.
- Ruffman, T. (1996) Reassessing children's comprehension-monitoring skills. In C. Cornoldi and J. Oakhill (eds.), *Reading comprehension difficulties: processes and intervention*. Mahwah, NJ: Lawrence Erlbaum. 33–67.
- Sanders, T.J.M. and M.A. Gernsbacher (2004) Accessibility in text and discourse processing. *Discourse Processes* 37.2: 79–89.
- Schneider, W. (1999) Working memory in a multilevel hybrid connectionist control architecture (CAP2). In A. Miyake and P. Shah (eds.), *Models of working memory: mechanisms of active maintenance and executive control*. Cambridge University Press. 340–74.
- Sheorey, R. and S. Edit (2004) Perceptions of reading strategy use among Hungarian EFL students. *Research News: The Newsletter of the IATEFL Research SIG* 14: 19–23.
- Shohamy, E. (1984). Does the testing method make the difference? The case of reading comprehension. *Language Testing* 1: 147–70.
- Siegel, L.S. (1994) Working memory and reading: a life-span perspective. *International Journal of Behavioral Development* 17: 109–24.
- Swanson, H.L. (1999) What develops in working memory? A life span perspective. *Developmental Psychology* 35: 986–1000.
- Swanson, H.L. and V. Berninger (1995) The role of working memory in skilled and less skilled readers' comprehension. *Intelligence* 21: 83–108.
- Taillefer, G.F. (1996) L2 reading ability: further insight into the short-circuit hypothesis. *Modern Language Journal* 80: 461–77.
- Teasdale, J.D. and P.J. Barnard (1993) *Affect, cognition and change: re-modelling depressive thought*. Hove, UK: Lawrence Erlbaum.
- van den Broek, P., M. Young, Y. Tzeng and T. Linderholm (1998) The landscape model of reading: inferences and the on-line construction of a memory representation. In H. Van Oostendorp and S.R. Goldman (eds.), *The construction of mental representations during reading*. Mahwah, NJ: Lawrence Erlbaum. 71–98.
- Walter, C. (2004) Transfer of reading comprehension skills to L2 is linked to mental representations of text and to L2 working memory. *Applied Linguistics* 25.3: 315–39.
- Waters, G.S. and D. Caplan (1996) The measurement of verbal working memory capacity and its relationship to reading comprehension. *The Quarterly Journal of Experimental Psychology* 49A.1: 51–79.
- Yuill, N. and J. Oakhill (1991) *Children's problems in text comprehension*. Cambridge University Press.

email: c.walter@ioe.ac.uk

[Final version received Nov. 31, 2006]

Appendix 1

The Tony story: a sample of the preliminary questionnaire for determining main points.

Story 100d

Read this story.

Tony didn't have a suit to wear to work that morning. His suit was at the cleaner's, and he had wanted to wake up early to go and get it. But he had not heard his alarm clock, and now it was late. As he looked round the kitchen, he saw his brother's car keys. Tony's brother had just bought his first car. Surely he wouldn't mind if Tony borrowed the car to go to the cleaner's. Tony took the keys and went to the car. He got in and started it. As he reversed he heard the sound of metal crunching against metal.

What is the story about? Write up to three sentences.

Now turn the page.

[next page]

How important are these elements in the story?

For each sentence, circle a number from 1 (not very important) to 5 (central to the story).

You can look back at the story.

If you are unsure, ask yourself: if this were changed, would it change the story much?

| | Not very important | | | Central to the story | |
|---|--------------------|---|---|----------------------|---|
| | ↓ | | | ↓ | |
| Tony's brother had just bought his first car. | 1 | 2 | 3 | 4 | 5 |
| Tony needed to wear a suit to work. | 1 | 2 | 3 | 4 | 5 |
| The story happened in the morning. | 1 | 2 | 3 | 4 | 5 |
| Tony's only suit was at the cleaner's. | 1 | 2 | 3 | 4 | 5 |

Appendix 2

The three anomalous L2 versions of the Tony story; in each version here (but not in the texts in the study), the anomaly is in italics.

Main Near

Tony had to wear a suit to work, but his only suit was at the cleaner's. He had wanted to wake up early to go and get it. But he had not heard his alarm clock, and now it was too late. *As he was putting on his suit*, he saw his brother's car keys. Tony's brother had just bought his first car. Surely he wouldn't mind if Tony borrowed the car to go to the cleaner's. Tony took the keys and went to the car. He got in and started it. As he reversed he heard the sound of metal crunching against metal.

Main Remote

Tony had to wear a suit to work, but his only suit was at the cleaner's. He had wanted to wake up early to go and get it. But he had not heard his alarm clock, and now it was too late. As he looked round the kitchen, he saw his brother's car keys. Tony's brother had just bought his first car. Surely he wouldn't mind if Tony borrowed the car to go to the cleaner's. *Tony put on his suit* and took the keys. He went to the car and started it. As he reversed he heard the sound of metal crunching against metal.

Subsidiary Remote

Tony had to wear a suit to work, but his only suit was at the cleaner's. He had wanted to wake up early to go and get it. But he had not heard his alarm clock, and now it was too late. As he looked round the kitchen, he saw his brother's car keys. Tony's brother had just bought his first car. Surely he wouldn't mind if Tony borrowed the car to go to the cleaner's. Tony took the keys and went to the car. He got in and started it. He wanted to get to the cleaner's before they closed *for the evening*.