

## Easy-to-read texts for students with intellectual disability: Linguistic Factors affecting comprehension

### Abstract

*Background:* The use of “easy-to-read” materials for people with intellectual disabilities has become very widespread but their effectiveness has scarcely been evaluated. In this study, the framework provided by Kintsch’s Construction-Integration Model (1988) is used to examine (a) the reading comprehension levels of different passages of Spanish text that have been designed following easy-to-read guidelines (b) the relationships between reading comprehension (literal and inferential) and various linguistic features of these texts.

*Method:* Sixteen students with mild intellectual disability (ID) and low levels of reading skills were asked to read easy-to-read texts and then complete a reading comprehension test. The corpus of texts was composed of a set of forty-eight pieces of news selected from [www.noticiasfacil.es](http://www.noticiasfacil.es), a Spanish digital newspaper that publishes daily journalistic texts following international guidelines for the design of easy-to-read documents (IFLA, Tronbacke, 1997).

*Results:* Participants correctly answered 80% of the comprehension questions, showing significantly higher scores for literal questions than for inferential questions. The analyses of the texts’ linguistic features revealed that the number of co-references was the variable that best predicted literal comprehension but, contrarily to what previous literature seemed to indicate, the relationship between the two variables was inverse. In the case of inferential comprehension, the number of sentences was a significant negative predictor; i.e. the higher the sentence density, the lower the ability of these students to find relationships between them. The effects of the rest of linguistic variables, such as word frequency and word length, on comprehension were null.

*Conclusions:* These results provide preliminary empirical support for the use of easy-to-read texts but bring into question the validity of some popular design guidelines (e.g. augmenting word frequency) to optimally match texts and reading levels of students with intellectual disability. Two factors are suggested as contributing to the effect of sentence density on inferential comprehension: 1) long texts present higher conceptual density so there are more ideas to store, retrieve and integrate which increases the demand on inferential reasoning and 2) long texts are perceived as difficult which

affects reading motivation and, consequently, induce passive reading strategies. The need for further research to elucidate the origin of our main findings with a larger and more heterogeneous sample of students with ID is highlighted.

*Keywords:* Easy-to-read texts, reading comprehension, linguistics, readability measures.

## **Introduction**

Reading in general and text comprehension in particular have been identified as challenging skills for students with intellectual disability (ID), who usually do not achieve the level of their chronological and educational peers (Connors, 2003). The use of “easy-to-read”, or simplified, texts is one of the strategies used by educators to improve the reading comprehension performance of these types of students. However, the design or selection of “easy-to-read” texts is not a straightforward issue, so educators usually use recommendations made by international organizations such as the *Guidelines for Easy-to-Read Materials* by the International Federation of Library Associations and Institutions (IFLA) (Tronbacke, 1997) or the *Make it simple Guidelines* by the International League of Societies for Persons with Mental Handicap (ILSMH) (Freyhoff *et al.* 1998).

However, as guidelines designers admit (Nomura *et al.*, 2010), further research is needed in order to provide empirical support for the mentioned sets of guidelines, especially in the case of readers with ID. In addition, the extent to which a text fulfills a particular guideline or set of guidelines is not easily quantifiable. For instance, one example of the *Make it Simple guidelines* set is to “avoid difficult words”. How could educators objectively measure and compare texts that differ in this variable? The application of readability formulae based on text linguistic measures such as word frequency and word length (e.g. the classic Flesch’s formula by Flesch, 1948) would aid in this process but, to the best of our knowledge, they have scarcely been tested empirically for the purpose of selecting easy-to-read texts for readers with ID. Finally, easy-to-read guidelines designers and many readability formula developers do not explicitly state the theoretical models of reading and reading comprehension framing their work, if any, so it is difficult to establish their construct validity. Therefore, using

the framework provided by Kintsch's (1988) construction-integration model, the aim of this study was twofold: (1) To test intellectually disabled students' reading comprehension performance on a corpus of texts designed according to the *Guidelines for Easy-to-Read Materials* (Tronbacke, 1997) and (2) To examine the relationship between texts' linguistics variables, on the one hand, and reading comprehension performance at both literal and inferential level, on the other, in order to better establish how to measure text readability.

The following section presents a brief review of the reading process in intellectual disability, which will help to frame the objectives of our research. That section is followed by a description of classical and new linguistic measures applied in the fields of readability and easy-to-read text design. Finally, the method and results of our study are explained and interpreted.

### **Reading in students with intellectual disability**

A general language deficit or developmental delay is commonly identified as the core process affected in individuals with intellectual disability, especially in some etiologies like the Down syndrome (Rondal, 2001). This general language problem might be influencing the acquisition of reading literacy although, paradoxically, some reading processes seem to be acquired in spite of such a deficit. For instance, individuals with Down syndrome are able to recognize words without good phonological awareness considered a pre-requisite of reading for regular readers (Fowler, Doherty & Boynton, 1995). However, as Fowler et al. suggest, word recognition and listening comprehension does not guarantee reading comprehension.

Therefore, this type of findings seems to highlight that reading is a process which composed of multiple components each of which should be considered in order to adequately measure individuals' competence. In fact, defined as the process of extracting meaning from print, reading is far from being a simple process and can be roughly divided into two components: (1) lexical decoding (the process of transforming printed word into a sound) and (2) comprehension (the process of understanding the literal and implicit ideas of a text). With regard to the former component, previous research has shown that most children with ID show weak lexical decoding skills (e.g. Connors *et al.* 2001) and a small *mental lexicon*, knowledge of words (Aitchison, 1994)

or reduced vocabulary (e.g. Ferrer *et al.* [1999] observed that people with ID aged 14 presented an equivalent age of 4-5 years in receptive vocabulary). Although strongly related, the focus of the present work was not decoding but comprehension, which according to Kintsch's Construction-Integration Model (1988), occurs at two levels: (1) literal level: comprehension of the actual meaning of single propositions and (2) inferential level: integration between text segments or between text and prior knowledge. These two sentences may serve to illustrate the extraction of literal and inferential meanings:

*The politician said that petrol prices would decrease by 10% in 2010. She was wrong.*

If a teacher asked his/her students after reading these sentences "*How much would the prices decrease in 2010 according to the politician?*", the correct answer would be *10%*; a piece of information which is literally stated in sentence 1. If the question was "*Did petrol prices decrease by 10% in 2010?*", the correct answer would be "*No, prices did not decrease a by 10%*". This piece of information must be inferred from the text by integrating ideas from sentence 1 and sentence 2 (the politician was wrong, so prices did not decrease by 10% in 2010). In addition, a student could use their prior knowledge about petrol price changes in 2010 in order to answer the second question (e.g. his/her father drove less in 2010 so petrol prices probably increased, rather than decreased). The first and second answers constitute what Kintsch calls the *textbase model* and it is based on the "construction" of a mental representation of the text. The third answer requires the "integration" of the reader's knowledge and experience with the mental representation of the text and is part of the reader's *situation model*.

The question that arises here is 'what is known about these two comprehension processes (literal and inferential) in intellectual disability?'

#### *Literal and Inferential levels of reading comprehension*

Regarding literal levels of comprehension, the low levels of vocabulary knowledge mentioned above suggests that the actual meaning of single sentences may not be understood if they are composed of low frequency content words that are not in the lexical repertory of students with ID. Besides, these students may present problems with functional words such as pronouns and prepositions, which are also essential syntactic cues that help us to understand clauses. Perovic (2006) found that young adults with

Down syndrome and mild intellectual disability presented difficulties interpreting reflexives (e.g. herself, himself) but not pronouns (e.g. him, her) in an oral comprehension task. The author concluded that, as this is the opposite pattern of acquisition to that found in typically developed children (first reflexives, then pronouns), language acquisition in Down syndrome is not just delayed but qualitatively different at the specific level of the morphosyntactic processing.

Although the delay versus deficit debate is a research question expandable to intellectual disability in general, reading comprehension in Down syndrome presents particular features (e.g. specific deficits in phonological awareness and short term memory, see Fowler, Doherty & Boyton for a review) not presented in others intellectual disability etiologies. Due to this particularity, we did not include students with Down syndrome in the present study.

Although the understanding of pronouns like reflexives can influence the literal comprehension of single sentences, it can also affect the inferential level of comprehension since they usually work as linking devices, connecting parts of the text that need to be integrated in order to get implicit ideas. The ability of children with ID to infer pronouns and other types of linking devices, like connectives (e.g. however, in addition, etc.), was measured by Verhoeven & Vermeer (2006) with a cloze task. Children with ID were asked to complete text fragments in which those elements had been omitted. Their scores in this task were significantly lower than those of typically developed children. A possible interpretation of these findings is that readers with ID find it difficult to understand the type of relationship established between sentences (e.g. causality or temporality) and consequently they cannot correctly guess the connective that should accompany them. Sanders *et al.* (2007) hypothesise that connectives could actually help poor readers to integrate sentences. Effectively, they found that individuals with poor comprehension levels (non-ID), obtained higher comprehension scores after reading texts with connectives than the texts without them, so we could predict a similar facilitative effect of connectives in students with ID.

In order to measure literal and inferential comprehension of narrative texts, Wolman (1991) asked children with mild ID to recall and judge the relevance of recalled statements of cohesive and non-cohesive versions of two stories. Children who

were able to make inferences would understand the meaning of the story, differentiating relevant from irrelevant information, as shown in their judgements of relevance. Causal cohesion of stories was manipulated by varying the proportion of causal chain statements in the text; the higher the proportion of causal chain statements, the higher the casual cohesion. The group of children with ID recalled significantly less statements and episodes of the stories, mainly in the non-cohesive stories, than children with learning disabilities (children with reading difficulties and normal intelligence) and children without disability (children with normative levels of reading and intelligence). This finding demonstrates that literal comprehension of the stories was lower in children with ID than in the other two groups. In addition, children with ID judged relevant and irrelevant statements as equally important, especially in non-cohesive stories; i.e. children with ID find it difficult to infer the meaning of the story, which would help them to discriminate relevant from relevant information.

Finally, regarding specifically inferential comprehension in students with ID, Bos & Tierney (1984), asked children with and without ID to read narrative and expository texts and free-recall the text content. Narrative texts tell a story and have beginning, middle and end, characters, plot or conflict, and setting. Expository texts inform or describe and the information is organized in a logical manner. The analysis of free-recall revealed that the number and quality of inferences were identical for both groups in the narrative text condition. In the case of expository text, children without ID outperformed children with ID in the quality of inferences. The authors suggest that the narrative scripts are more accessible than expository scripts for children with ID, which would help them to make reasonable inferences about events in narrative texts. So far, the insight obtained from this finding is that children with ID are able to make inferences if the structure of the text is well known to them, as in the case of narrations.

In addition to specific linguistic and cognitive factors, there are subjective and motivational factors involved in reading comprehension as well. As detected in previous studies with ID youth readers, they usually do not read autonomously; feel frustrated or are reluctant to read long texts and texts with technical difficulty (e.g. academic texts). However, their attitude is highly receptive to the use of Internet reading environments, for example, digital newspapers (Gómez, 2011; Morgan & Moni, 2008) where in

addition to multimedia content, they can find personally interesting texts. That was the type of texts used in the present study.

This brief review of the topic confirms that children with ID present problems at both levels of Kintsch's Model of reading comprehension; i.e. literal and inferential. In particular, they can find it difficult to: understand the meaning of single sentences because of a lexical deficit; recall main statements of the text and identify main ideas; and interpret functional words and other syntactical cues that help to make inferences. We have also learned that findings from research in typically developed children must not be generalized for children with ID because their pattern of reading behaviour can be different, not just delayed (Perovic, 2006).

Now that the profile of readers with ID has been outlined, how can we know if a text fulfils the lexical, syntactical and cohesion constraints required for them? The following section presents a description of the most popular psycholinguistic measures in the fields of legibility and easy-to-read text design.

### **Easy to read text design guidelines and psycholinguistic measures**

As we mentioned above, there are sets of guidelines for the design of easy-to-read texts produced by several international organizations (e.g. IFLA and ILSMH Guidelines) but there is little empirical research that supports them, especially in the case of readers with ID.

Karreman *et al.* (2007) have conducted one of the few empirical studies testing the comprehension levels of individuals with ID after reading two different versions of a website: one adapted on the basis of ILSMH's easy-to-read guidelines (1998) and the other non-adapted. They found that both literal and inferential comprehension of the group of individuals with ID was higher in the adapted version than in the non-adapted version of the website. Although this finding is very useful, the drawback of the methodology used in this research is that the adapted and non-adapted websites differed in a number of linguistic elements (e.g. length of words and sentences, frequency and abstractness of words, tense of sentences, etc.), which makes it difficult to distinguish which elements or collection of elements were in effect facilitating reading comprehension.

In addition to the indicated lack of exhaustive examination of the validity of specific IFLA or ILSMH guidelines, the problem with some guidelines is that they are formulated in very general terms, making it difficult to use them. For instance, two of the IFLA guidelines are “use a simple, straightforward language” and “use a clear and logical structure”. How can text designers implement or measure these types of features? Readability or legibility formulae may serve to aid in the measuring process if a text fulfils or does not fulfil some of them.

As suggested by Crossley *et al.* (2007), we can talk about shallow and deep readability formulae. In the former group, the measurement of text difficulty relies on superficial text variables, such as word and sentence length (e.g. Flesch Index by Flesch, 1948), or word frequency (CELEX by Baayen *et al.*, 1995). It is assumed that the shorter the word and sentence length and the higher the word frequency, the lower is the text difficulty. However, there is contradictory evidence concerning this. While the classical study of Doctorow *et al.* (1974) showed that children’s reading comprehension was increased when text with high frequency words was provided, other authors have found a null effect of increasing word frequency on the literal and inferential comprehension of texts (Freebody & Anderson, 1983; Ryder & Hughes, 1985).

The effect of word frequency is not clear either in the case of simpler lexical tasks, such as naming or lexical decisions. Although it has been classically reported that frequent words are read more quickly than infrequent words (Frederiksen & Kroll, 1976), null or even inverse effects of frequency have also been found (Adelman *et al.*, 2006; Monsalve & Cuetos, 2001).

In addition to the lack of clear empirical evidence about the effect of shallow readability measures, these types of formulae are insufficient for evaluating other text features, such as cohesion. In the attempt to overcome the problems with traditional readability formulae, the Coh-Metrix computational tool (Graesser *et al.*, 2004) measures cohesion features, such as the density of connectives and co-references, as it is assumed that those elements would help readers to link ideas and construct a coherent mental representation of the text content. For example, Crossley *et al.* (2007), measured the number of connectives in authentic texts (text written to fulfill a social purpose for native speakers within a language community, Lee, 1995) and simplified texts for second language students. They found that simplified texts tend to contain more



common connectives (e.g. and, but) and more complex and infrequent connectives (e.g. if, only if, on the condition that, etc.) than authentic texts.

However, Crossley et al. (2007) noted that the syntactic simplification may not necessarily be facilitating text comprehension because such a type of simplification could prevent readers from elaborating ideas that are needed to understand hypothetical situations. For instance, let's consider the following paragraph written by a journalist: "If politicians were sensible they would search out the economists that predicted the crisis". The hypothetical situation supposed by the journalist would be more difficult to infer in a paragraph like this "Politicians are not sensible and they did not search out the economists that predicted the crisis", where the syntactic complexity is lower than in the "authentic" text but the journalist's hypothesis about the relationship between politicians sensibility and the lack of prevention of the crisis is not evident.

Another cohesion element is the co-reference, defined as argument overlaps between sentences, which seems to improve text comprehension and reading speed (Kintsch & van Dijk, 1978). For instance, in the following two sentences "Anne ordered Claudia to sit down. She is very bossy", "she" works as a "pronominal co-reference" for "Anne". If the proper name "Anne" was repeated in the second sentence instead of "she", then it would be working as a "same noun phrase repetition" co-reference. McNamara *et al.* (2010) assessed the validity of a set of text cohesion indices of Coh-Metrix. The researchers checked the ability of these indices to discriminate between high and low-cohesion versions of texts. Texts were extracted from published experimental studies in which the variable "cohesion" had been manipulated between groups, that is, where different groups of participants had been asked to read more cohesive and less cohesive versions of the same texts. McNamara et al. found that, in effect, more cohesive texts contained more causal connectives and co-references (in particular, noun, arguments and stem overlaps) than less cohesive texts. The validation of these deep-level linguistic measures would suggest that they must be taken into account in order to design or adapt texts for students with ID.

Finally, it is important to note that the text adaptation approach for students with ID is frequently aided with the used of pictures, icons and symbols despite the fact that the benefits of them are not well established yet. As a matter of fact, Jones, Long and Finlay (2007), observed that the addition of symbols to texts improved reading comprehension of adults with mild and borderline ID. On the other hand, the study of

Poncelas and Murphy (2007) compared two versions of a simplified text: one text-based and the other symbol-based (with text). Adults with ID were asked a series of questions about the material to assess comprehension. Both versions produced low levels of comprehension. The group of readers with symbols (and text) showed no better understanding than the group with text only. The results of these two studies would suggest that the addition of symbols to simple texts does not necessarily improve readers' comprehension of it.

It is important to note that methodological differences between both studies (e.g. types of symbols, type of texts, age mean of the sample, etc.) may account for the difference in effectiveness of addition of symbols to texts. Therefore, as the former authors suggest, future research should focus on the control of those methodological parameters in order to elucidate the origin of symbols reading comprehension facilitation. While admitting the importance of graphic aids, the present study focused exclusively in the isolation of linguistics factors affecting literal and inferential reading comprehension for readers with ID.

This literature review makes it evident that the recommendations for the design or selection of easy-to-read texts are neither proven nor easily quantifiable for students with ID. It seems necessary to identify which linguistic features must be considered to make a text easy to read for students with ID. Therefore, we designed a descriptive study aimed at: (1) Testing the reading comprehension levels of a corpus of easy-to-read texts by intellectually disabled students, (2) Examining the relationship between texts' surface and deep linguistic measures, on the one hand, and literal and inferential reading comprehension performance on the other.

In particular, our research questions were: (1) What is the reading comprehension level of a corpus of easy-to-read texts by intellectually disabled students? (2) Do surface and deep linguistic measures predict literal and inferential reading comprehension in students with ID?

## **Method**

### **Participants**

Twenty-eight students with ID (17 males and 11 females) were recruited in 'Camí Obert', a centre of vocational training belonging to the Asociación Pro-

Discapacitados Psíquicos of Alicante. Students and students' parents signed an informed consent before participating in the study. The average age of the sample was 18.9 (SD=1.84, ranging from 16 to 22). No individuals with known genetic syndromes, in particular Williams' and Down's syndromes, were included in the sample. The reason for this exclusion criterion was that those specific syndromes present language particularities which could be not present in the rest of etiologies and they did not constitute a big enough group as to be studied as a separated group.

In order to establish the baseline of reading skills, the participants were asked to complete a series of reading processes and reading-related linguistic tests: Spanish Peabody Picture Vocabulary Test (PPVT-III) (Dunn *et al.* 2006), syntactic processing subtest from the PROLEC-SE test (Ramos & Cuetos, 1999) and the ECL1 Reading Comprehension Test (De la Cruz, 2005). The PPVT-III measures the individual's receptive vocabulary by asking participants to identify words based on a series of pictures. The examiner states a word describing 1 out of 4 pictures and asks the individual to point to or say the number of the picture that the word describes. The syntactic subtest of the PROLEC-SE consists of a set of pictures that represents scenes (for instance, a policeman arresting a burglar). Participants are asked to select one written sentence out of four that correctly describes the scene. The syntactic complexity of the sentences increases as trials proceed so that the syntactic competence of participants can be established. The ECL1 measures reading comprehension in children aged 6 to 8 years and consists of 3 short texts (approx. average sentence length of 100 words) accompanied by literal and inferential questions. In addition, the participants' IQ was measured with the Kaufman Brief Intelligence Test (K-BIT) (Kaufman & Kaufman, 1997).

Twelve participants were removed from the original sample because they did not fulfil the criterion of mild to moderate mental retardation according to the DSM-IV-TR Manual (American Psychiatric Association, 2000). The average IQ score measured with the K-Bit test for the remaining 16 participants was 64.9 (SD=10, min=43, max=82). Corrected scores for each measure were obtained from the normative data of the higher, either educational level, or chronological age provided by each particular test. The remaining 16 participants fulfilled the criterion of having a reading comprehension level of at least second year primary school, measured with the ECL test, although only

56.2% of the sample reached the level of third year primary school (raw score,  $M=9.1$ ;  $SD=2.4$ ). For the rest of the reading tests, the participants performed below their chronological age (Peabody's raw score,  $M=110.9$ ;  $SD=22$ ; Equivalent age= 9.3) or educational level (PROLEC-SE's raw score,  $M=12.9$ ;  $SD=4.3$ ; percentile of fourth-grade secondary school= 5).

### Task and materials

Participants were asked to read easy-to-read texts and then complete a reading comprehension test. This study was framed by a project aimed to increase the usage, usability and validity of a Spanish digital newspaper, [www.noticiasfacil.es](http://www.noticiasfacil.es), which publishes daily journalistic texts. These news or journalistic texts are directly designed (written) in an easy-to-read format following the IFLA's guidelines. The newspaper uses this particular set of guidelines because they form the standards for easy-to-read text adaptation and design more commonly found internationally and the only one adapted to Spanish. Therefore, the corpus of easy-to-read texts was composed of a set of 48 pieces of news selected from this website. The genre of texts was narrative or expository. There were 6 to 10 pieces of news for each of the newspaper's six main sections: economy, sport, national, international, culture and miscellaneous. For each piece of news, three levels of linguistic variables were measured: lexical and sentence variables (shallow measures), and textual cohesion variables (deep measures).

*-Lexical variables:* number of words, average syllables per word and average word frequency. Word frequency was obtained from the Dictionary of Spanish Linguistic Units (Alameda & Cuetos, 1995), a Spanish data base of two million words in which word frequency is measured as the number of occurrences per million words.

*-Sentence variables:* number of sentences per text, average words per sentence and Flesch-Szigriszt Index (Szigriszt, 1992), a version of the classical Flesch index for texts in Spanish. The Flesch-Szigriszt Index is expressed as follows:  $206,835 - (62.3 \times S/P) - P/F$ , where S is the number of syllables, P is the number of words in the text and F is the number of phrases (defined as the content between two punctuation marks).

*-Textual cohesion variables:* referring to elements affecting the integration between sentences and paragraphs. In particular, in this research, we measured: (1) the number of co-references – *noun overlap* (where the same noun is common to a pair of sentences;

e.g. child-child), *argument overlap* (where the word stem is repeated among sentences without varying the lexical category; e.g. child-children) and *stem overlap* (where the shared form is just the word stem, allowing word category variation; e.g. child-childish); (2) the number of ellipses (omissions in a clause of one or more required words or elements that are implicit and can be understood anyway); and (3) the number of connectives per text (linking words, such as conjunctions between ideas or clauses; e.g. *however, even though, after that*).

The reading comprehension test was composed of multiple-choice questions that tested readers' comprehension of the news content. These questions were compiled according to the two levels of processing (literal and inferential) of the Kintsch's Construction-Integration Model (1988). Literal questions: the answer to this type of question was either explicitly stated or restructured within a single sentence. Inferential questions: required the integration of inter-sentence information and integration of textual information with prior knowledge. The number of literal and inferential questions varied per text because it was not always possible to extract them from the information provided for the texts (see examples in Appendix 1). In particular, for 3 pieces of news there were no inferential questions so at this level only 45 pieces of news were analysed. One piece of news did not have any literal questions, so the analysis of this type of question was performed for 47 texts only. The total number of questions was 178 (100 literal questions and 78 inferential questions).

Multiple choice questions were used for several reasons: 1) To prevent the potential masking effects of writing competence and memory or verbalization difficulties of our students which could appear with open questions such as retelling or another type of free recall technique (for a discussion about this issue see Fletcher & Clayton, 1994). In order to prevent the additional difficulty of the questions themselves, both questions and choices were short and written in easy-to-read format; 2) To facilitate answer analysis and interpretation due to the high number of questions we had (178 in total); 3) To ensure that readers had to activate a specific type of process in order to correctly answer our questions since we were interested in measuring two different types of reading processes: literal vs. inferential. The correct choice for multiple choice questions can be selected by researchers in such a way that readers necessarily had to activate one of the referred processes to find it. However, the process

activated by open questions could be more dependent on the reader's question answering satisfaction threshold than on the information comprehension itself.

### Procedure

Each week, for a period of 16 weeks, participants were asked to read three pieces of news selected randomly from [www.noticiasfacil.es](http://www.noticiasfacil.es) and to complete a reading comprehension test. Thus, there were 48 pieces of news in total. Each week, participants were given a booklet with the three texts, each followed by comprehension questions on it. The order of presentation of the texts was the same for every participant since, as pieces of news, the relevance of the topic for each participant could have varied if the presentation order had been varied throughout the days. The order of presentation of literal and inferential question options was randomized for each text. Although participants were instructed to first read each text and then answer the questions, they were allowed to go back to the text if necessary. When the questions for a text had been answered, the participants were instructed to read the next text. There was no time limit for completing the task. Text reading and question fulfilment processes were supervised by the participants' educators ensuring that their doubts about the question answering process were solved before starting.

### Results

*What is the reading comprehension level of a corpus of easy-to-read texts by intellectually disabled students?*

Our first goal was to verify if easy-to-read news texts were, in effect, comprehended by students with ID. Therefore, an ANOVA test was performed for each item (piece of news), with the percentage of correct answers for literal and inferential questions used as measures ( $N$  was 47 and 45 respectively). The percentage of correct answers to literal and inferential questions was .87 ( $SD=.12$ ) and .76 ( $SD=.22$ ) respectively, with the difference between two types of questions being significant,  $F(1, 43) = 11.26$ ;  $MSE = .248$ ;  $p < .0017$ .

*Do surface and deep linguistic measures predict literal and inferential reading comprehension in students with ID?*

### Correlation analyses

In order to address the second goal of the study (i.e. to examine the relationship between the texts' surface and deep linguistic measures and reading comprehension performance), several Pearson correlation analyses were performed between each block of linguistic measures on the one hand and comprehension measures on the other. The descriptive data for each linguistic variable is provided in Table 1.

*Table 1 about here*

With regard to lexical variables, we found significant negative correlations between the number of words and both literal and inferential comprehension questions; i.e. the longer the texts, the lower the literal and inferential comprehension of the students. Neither word length (number of syllables per word) nor word frequency correlated significantly with comprehension scores.

With regard to sentence variables, we found significant negative correlations between the number of sentences and both literal and inferential comprehension; i.e. the higher the number of sentences, the lower the literal and inferential comprehension of the texts. However, no significant correlations were found between comprehension scores on the one hand and average words per sentence and Flesch-Szigriszt index on the other.

Finally, with regard to textual cohesion variables, we found significant negative correlations between the number of connectives and literal but not inferential comprehension. The total number of co-references correlated significantly and negatively with literal comprehension; i.e. the higher the number of co-references contained by texts, the lower the readers' scores for literal questions. This correlation was not replicated in the case of inferential comprehension questions. Similarly, the number of ellipses did not correlate significantly with comprehension scores.

#### Multiple regression analysis

Those factors that correlated significantly with comprehension were considered for two forward, stepwise, multiple regression analyses performed with the aim of investigating if they were collectively predicting comprehension. In the first multiple regression analysis, the number of words, sentences, connectives and co-references were used as predictors and the percentage of correct answers in literal comprehension questions used

as a dependent variable. The best regression model was obtained in the first step where only the variable “number of co-references” was entered, accounting for 14% of the variance in the literal comprehension of our students with ID (Adjusted  $R^2 = .14$ ,  $F(1, 45) = 8.42$ ,  $p < .001$ ; Beta =  $-.40$ ,  $p < .01$ ). As the Beta coefficient is negative, this result means that the more co-references there are in the text, the lower the literal comprehension of the students.

In the second multiple regression analysis, the number of words and sentences were used as predictors and the percentage of correct answers in inferential comprehension questions used as a dependent variable. Again, we obtained a significant 1-step regression model where the number of sentences accounted for 12 % of the variance in inferential performance (Adjusted  $R^2 = .12$ ,  $F(1, 43) = 6.74$ ,  $p < .01$ ; Beta =  $.37$ ,  $p < .01$ ). The negative Beta value indicates an inverse relationship between the two factors; i.e. the more propositions there are in the text, the lower the inferential comprehension of students.

*Do prior reading skills and IQ levels predict literal and inferential reading comprehension in students with ID?*

Finally, in order to examine the relationship between prior reading and IQ levels of students and comprehension performance, several Pearson correlation analyses were performed between the direct scores in each of the standardized reading tests on the one hand and the percentage of correct answers in literal and inferential comprehension of news on the other. Only direct scores in the standardized reading comprehension test (ECL-1) positively predicted news comprehension at both literal and inferential levels (literal,  $r(N=16) = .74$ ,  $p = .001$ ; inferential,  $r(N=16) = .53$ ,  $p = .003$ ), that is, the higher the ECL-1 score, the higher the literal and inferential comprehension of the news. The rest of correlations were not significant (K-BIT, literal;  $r(N=16) = .41$ ,  $p = .12$ ; inferential  $r(N=16) = .39$ ,  $p = .14$ ; PEABODY, literal,  $r(N=16) = .29$ ,  $p = .27$ ; inferential,  $r(N=16) = .05$ ,  $p = .85$ ; and PROLEC-SE, literal,  $r(N=16) = .13$ ,  $p = .62$ ; inferential,  $r(N=16) = .05$ ,  $p = .84$ ).

## **Discussion**

*Reading comprehension level of a corpus of easy-to-read texts by intellectually disabled students*



Regarding our first research question, the results indicate that the corpus of easy-to-read texts tested is actually comprehended by our sample of intellectually disabled students, especially at a literal level, with 87% of questions correctly answered. These data are optimistic regarding the use of easy-to-read texts as inputs for students with ID but also suggest that the design method applied to simplify this corpus of texts does not guarantee deep or inferential levels of reading comprehension for such students. In other words, although the average performance in inferential questions was high (76%), 24% of inferential questions were answered incorrectly by our students despite the apparent simplicity of the texts in our corpus, with an average of 94 words. These results agree with those of Karreman *et al.* (2007) who observed that participants with ID who used a website adapted following easy-to-read guidelines (Freyhoff *et al.*, 1998) answered more literal and inferential questions correctly than when using a non-adapted version of the website. Although the difference between the number of correct literal and inferential answers was not statistically tested in the study of Karreman *et al.*, they also observed higher scores in literal questions than in inferential questions. Our congruent results support the idea of conducting theoretically driven research that, in our case, made it possible to distinguish between different levels of reading comprehension, based on Kintsch's model.

*Surface and deep linguistic measures as predictors of literal and inferential reading comprehension*

With regard to the second research question, that is, the relationship between linguistic variables and comprehension performance, we found that only one linguistic variable (number of co-references), predicted literal comprehension, explaining just 14% of its variance. However, the relationship between these two variables was inverse, in contrast to what previous literature seemed to indicate. For instance, Crossley *et al.* (2008) found that the number of co-references (measured by noun overlaps) was a good predictor of text difficulty ( $R^2=.63$ ); i.e. the higher the number of co-references, the better the comprehension scores in a cloze test. Similar results were found by McNamara *et al.* (2010): noun co-reference was highly discriminative between low and high-cohesion texts which led to low and high comprehension, respectively. It could be hypothesized that in our texts the number of co-references co-varied with text length, which, in turn, could be explaining the literal comprehension variance. Alternatively, as noted by

Meisel (1980), the redundancy introduced in simplified texts by means of noun overlaps would make the sentences' grammar more complex and unnatural. As was shown in the reading skills baseline description, our students with ID present a low syntactic proficiency which prevents them from understanding unnatural sentence grammar.

Concerning inferential comprehension, the only linguistic feature that was a good predictor was the number of sentences. The higher the sentence density, the lower the ability of our students with ID to find relationships between sentences. This result could be related to the lower capacity of ID students to store and retrieve information from memory. Actually, Wolman (1991) found that children with ID recalled significantly less statements and episodes from the stories than normally developed children. Therefore, if students with ID have problems in retaining information, increasing the number of information units could hinder the process of making inferences between parts of the text or between the text and prior knowledge. In the end, the impossibility of making inferences, which would concurrently reduce the amount of information units to work with, could cause cognitive overload, leading to students with ID performing a superficial processing of texts.

Alternatively, text length could have affected our students' self-efficacy and reading motivation. Several studies have found that students with ID tend to be reluctant to read long texts (Gómez, 2011; Morgan & Moni, 2008) maybe because text length works as a superficial difficulty cue for them. Therefore, their perception of low self-efficacy with respect to comprehension of difficult texts may have led them to actively avoid reading and at best become passive readers (Guthrie & Davis, 2003). If students with ID did not invest enough effort reading long sentences in our study, they probably answered inferential multiple choice questions randomly. Consequently, as suggested by Guthrie and Davis, a primary challenge for the literacy improvement of students with ID would be to re-engage them. These authors propose that, in addition to providing a wide range of texts focused on content that deals with real-life problems (like the journalist texts used in this research) and the selection and design of texts matched to their reading ability, a pathway to the development of reading motivation is to give direct instruction for important reading comprehension strategies such as questioning, searching, summarizing, using graphic organizers, comprehension monitoring and critical evaluation. Guthrie and Davis (2003)'s rationale for this solution is that:

As students learn the strategies, they gain the confidence in their own capabilities. They increase their view of themselves as competent in reading. Engagement is fueled by the self-perception of ability to perform meaningful reading tasks. Consequently, strategy instruction fulfills the motivational need for self-perceived competence (Ryan & Deci, 2000) as well as the cognitive need for possessing skills that are central to text comprehension. (p. 77)

Another interesting result is that related to the effect of connectives: the number of connectives was negatively correlated with literal comprehension, although it did not contribute significantly to our regression model. It means that, probably, its effect is explained as well by a co-variation with other variables such as the length or conceptual density of the text. That could explain the negative effect on comprehension; i.e. the longer the text, the higher the number of connectives and also the number of ideas contained in the text (since connectives are supposedly added to connect ideas). Sanders *et al.* (2007) did find a facilitative effect of connectives on comprehension, but the conceptual complexity of original and experimental texts was identical in their study; i.e. the only aspect which varied between the two conditions was the presence/absence of connectives. In other words, connectives can facilitate text comprehension but, simultaneously, when they are present in a text it means that such a text contains ideas that must be integrated. In turn, that would be good news for the education of students with ID because, according to that interpretation, our results would indicate that designers of easy-to-read texts do not entirely renounce using texts with a certain level of conceptual richness. From a motivational point of view, the selection of texts personally relevant and connected to students' needs may compensate for the conceptual richness and complexity of long texts.

It is also worth noting that the effects of word frequency and the Flesch-Szigriszt Index on our measures of comprehension were null; i.e. they did not correlate significantly with either literal or inferential comprehension scores. These results are surprising because these two measures have traditionally been used to select and design texts in educational contexts. According to our data, more complex measures are needed

at the sentence and textual cohesion level to ensure that educational text will be understood by ID students.

### *Individual differences in reading skills of the students with ID*

In addition to linguistic variables, the individual differences in reading skills of the students with ID could also be affecting their comprehension performance. Amongst the range of individual measures evaluated (vocabulary knowledge, syntactic skills and comprehension), only the reading comprehension levels of our students positively predicted both literal and inferential comprehension; i.e. those students with better prior levels of reading comprehension took more advantage of the easy-to-read material than students with lower levels. This result suggests that, when adapting texts for this type of students, individual difference in reading comprehension, have to be taken into account.

On the other hand, the lack of predictive power of vocabulary and syntactic skills on reading comprehension could be explained by a floor effect. That is, the levels of vocabulary and syntactic skills were as low in general in our students as not to have an effect on comprehension. Therefore, we cannot discard the idea that given a particular level of performance, those linguistics variables can predict reading comprehension in students with ID. Anyhow, our students' reading profile is similar to the obtained by other researchers (e.g. Aitchison, 1994; Moni & Jobling, 2001) which signals the necessity to study it in depth in order to elucidate its origin: acquisition delay, deficit in a particular reading process or lack of reading motivation and engagement. Given the scarcity of specific theories and previous hypotheses about intellectual disability reading comprehension, a more qualitative and inductive methodology such as the case study or the narrative research may help to come up with ideas to study and develop research questions.

### **Conclusions: methodological shortcomings and future research**

In conclusion, we have achieved the two goals of our study; i.e. to validate a corpus of easy-to-read texts for students with ID and to broaden the knowledge of the relationship between texts' linguistics variables, on the one hand, and reading comprehension performance at both literal and deep/inferential level on the other.

However, although our research questions have been answered, further research is needed in order to tackle the several limitations of our study. One of our drawbacks is related to the small sample size ( $N=16$ ) which, together with the highly selected sample (as opposed to a heterogeneous population of students with ID), is limiting the generalization of our findings. Therefore, before it can be generalized our study should be replicated with a larger sample size and other levels of intellectual disability. In order to answer the question of qualitative difference versus acquisition delay of particular reading processes suggested by the Perovic (2006)'s findings about syntactic processing in Down syndrome, we would also need a control group formed by typically developed students (without ID) but similar reading comprehension. In fact, the comprehension-age match design has become a popular tool for investigating the causes of reading comprehension failure (Cain, Oakhill & Bryant, 2000). Let's imagine that we are interested in the role of casual versus temporal connectives (e.g. because vs. when) in texts comprehension. If the typically developed group showed difficulties with temporal but not casual connectives and the reverse pattern was observed in the group of students with ID matched in reading comprehension age, then it could be concluded that connective processing is not delayed in the group of students with ID but qualitatively different with regard to the control group.

A second important methodological limitation that needs to be overcome in future research is the exclusive use of multiple-choice questions as a measure of reading comprehension which could be preventing students from making and reporting their own inferences. Actually, Karreman *et al.* (2007) observed that open questions were more sensitive to comprehension differences between students with ID than multiple-choice questions. However, the risk of using open questions is that memory and comprehension skill can be confounded. In addition to the memory issue, as it was highlighted in the material section, multiple-choice questions were considered more suitable for this particular study than open questions due to several reasons: the prevention of the potential masking effects of writing competence or verbalization difficulties of our students (Fletcher & Clayton, 1994); the facilitation of readers' answer coding (of a total of 178 questions); and the activation of specific types of reading processes (literal vs. Inferential) in our readers. If these constraints are not present in the assessment situation (e.g. writing and verbalization difficulties have been

discarded), then we encourage educators to use open questions instead of close ones in order to detect unexpected inferences and elaborations made by students.

A third methodological issue is related to the correlational design of our study which has allowed us to describe the reading comprehension behaviour of a group of students with ID and to identify several text linguistic variables related to it but it does not allow us to answer causal questions. Therefore, in order to study in depth the origin of our findings such as co-reference or text length effects on literal and inferential comprehension respectively, process measures like eye movements could be added to output measures (e.g. comprehension questions). Eye movement measures could help us to detect if readers are actually reading long sentences or neglecting them and paying especial attention (longer fixation times and revisits) to particular elements of long texts such as connectives or co-references. This type of process measure, combined with an inductive methodology like a case study, could provide us with an account of causal hypotheses to be tested by means of comprehension-age match designs.

Finally, with the aim of augmenting reading engagement, future research could focus on the use of electronic texts since the Internet and multimedia environments have been found to be highly motivating for youth students with ID (Gómez, 2011). However, we wanted to note that the use of electronic texts is not an un-risky engagement solution. Electronic reading not only involves reading ability but other kind of skills such as navigation and computer operating skills (e.g. operating a mouse, typewriting, etc., for a revision see Van Deursen, Van Dijk & Peters, 2011) which paradoxically might make more difficult the reading comprehension process for our target students.

Despite the limitations of our methodology, we think that our findings throw light on the validity of superficial and deep measures of text difficulty for students with ID. From an applied point of view, we expect our insights to assist professionals in the field of education (teachers, text designers, etc.) in the task of more optimally matching texts and ID students' reading levels.

## References

- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders*, (4th ed., text revision). Washington, DC: American Psychiatric Association.
- Adelman J. S., Brown G. D. A., & Quesada J. F. (2006) Contextual diversity, not word frequency, determines word naming and lexical decision times. *Psychological Science* **17**, 814–823.
- Aitchison, J. (2002). *Words in the Mind: An Introduction to the Mental Lexicon*. 3rd ed. Blackwell Publishing, Oxford
- Baayen R. H., Piepenbrock R. & Gulikers L. (1995) *The CELEX lexical database*. University of Pennsylvania, Linguistic Data Consortium, Philadelphia.
- Bos C. & Tierney R. (1984) Inferential Reading Abilities of Mildly Mentally Retarded and Nonretarded Students. *American Journal of Mental Deficiency* **89** (1), 75-82.
- Cain, K., Oakhill, J. V. & Bryant, P. E. (2000). Investigating the causes of reading comprehension failure: The comprehension-age match design. *Reading and Writing*, *12*, 31-40.
- Connors F. A. (2003) Reading skills and cognitive abilities of individuals with mental retardation. In: *International review of research in mental retardation, Vol 27* (ed L. Abbeduto), pp 191–229. Academic Press, London.
- Connors F.A., Atwell J.A., Rosenquist, C.J., & Sligh, A.C. (2001) Abilities underlying decoding differences in children with intellectual disability. *Journal of Intellectual Disability Research* **45**(4): 292-299.

- Crossley S. A., Dufty D. F., McCarthy P. M., & McNamara D. S. (2007) Toward a new readability: A mixed model approach. In: *Proceedings of the 29th Annual Conference of the Cognitive Science Society*, (eds D.S. McNamara & J.G. Trafton), pp: 197-202. Cognitive Science Society, Austin.
- Crossley S. A., Greenfield J., & McNamara D. S. (2008) Assessing text readability using cognitively based indices. *Tesol Quarterly* **42**(3), 475-493.
- De la Cruz M.V. (2005) Evaluación de la Comprensión Lectora (ECL).TEA, Madrid.
- Doctorow, M., Wittrock, M.C. & Marks, C. (1974) Generative processes in reading comprehension. *Journal of Educational Psychology* **70** (2), 109-118.
- Dunn Ll. M., Dunn Ll. M. & Arribas D. (2006) *PPVT-III Peabody, Test de Vocabulario en Imágenes*. TEA, Madrid.
- Ferrer A., Brancal M. & Asensi M<sup>a</sup>.C. (1999) La evaluación de la producción léxica en las personas con retraso mental: alternativas a la medición psicométrica. Paper presented at the III Congreso Internacional de Psicología Evolutiva y de la Educación, Santiago de Compostela.
- Flesch R. (1948) A new readability yardstick. *Journal of Applied Psychology* **32**, 221–233.
- Fowler, A.E., Doherty, B.J., & Boynton, L. (1995). The basis of reading skill in young adults with Down syndrome. In L. Nadel & D. Rosenthal (Eds.), *Down Syndrome: Living and Learning in the Community* (pp. 182-196). New York: Wiley-Liss.
- Freebody P. & Anderson R.C. (1983) Effects of vocabulary difficulty, text cohesion and schema availability on reading comprehension. *Reading Research Quarterly* **18** (3), 277-94.



- Frederiksen J. R. & Kroll J.F. (1976) Spelling and sound: approaches to the internal lexicon. *Journal of Experimental Psychology: Human Perception and Performance* 2 (3), 361-379.
- Freyhoff, G., Hess G., Kerr L., Menzel E., Tronbacke B., & Van der Veken K. (1998) *Make It Simple, European Guidelines for the Production of Easy-to-Read Information for People with Learning Disability*. ILSMH European Association, Brussels.
- Gómez, P. (2011) Identificación de dificultades de comprensión lectora en el uso de Internet en jóvenes con discapacidad intelectual para el desarrollo de un periódico digital. Poster presented to the *XV Congreso Nacional y I Internacional de Modelos de Investigación Educativa*, Madrid, 2011, September.
- Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, and Computers* 36, 193-202.
- Guthrie, J. T., & Davis, M. H. (2003). Motivating struggling readers in middle school through an engagement model of classroom practice. *Reading and Writing Quarterly*, 19, 5985.
- Jones, F.W., Long, K., & Finlay, W.M.L. (2007), Symbols can improve the reading comprehension of adults with learning disabilities. *Journal of Intellectual Disability Research*, 5, 545-550.
- Karreman, J., Van der Geest, T., & Buursink, E. (2007) Accessible Website Content Guidelines for Users with Intellectual Disabilities, *Journal of Applied Research in Intellectual Disabilities* 20(6), pp. 510-518.
- Kaufman A.S. & Kaufman N.L. (1997) *Test Breve de Inteligencia de Kaufman*. TEA, Madrid.
- Kintsch W. (1988) The use of knowledge in discourse processing: A construction-integration model, *Psychological Review*, 95, 163-182.

- Kintsch W. & Van Dijk T.A. (1978) Toward a model of text comprehension and production. *Psychological Review*, **85** (5), 363-394.
- Lee, W. Y. (1995). Authenticity revisited: Text authenticity and learner authenticity. *ELT Journal* *49*, 323–328.
- McNamara D.S., Louwse M.M., McCarthy P.M., & Graesser A.C. (2010) Coh-Metrix: Capturing linguistic features of cohesion. *Discourse Processes* **47**, 292 - 330.
- Meisel J. (1980) Linguistic simplification. In: *Second language: Trends and issues* (ed S. Felix), pp. 13– 140. Gunter Narr, Tubingen.
- Moni, K. B., & Jobling, A. (2001). Reading related literacy learning of young adults with Down syndrome: findings from a three year teaching and research program. *International Journal of Disability, Development and Education*, *48* (4), 377-394.
- Monsalve A. & Cuetos F. (2001) Asimetría hemisférica en el reconocimiento de palabras: efectos de frecuencia e imaginabilidad. *Psicothema* **13** (1), 24-28.
- Morgan, M. F., & Moni, K. B. (2008). Meeting the challenge of limited literacy resources for adolescents and adults with intellectual disabilities. *British Journal of Special Education*, *35*(2), 92-101.
- Nomura, M. & Nielsen, G.S.. & Tronbacke, B. I. & International Federation of Library Associations and Institutions (2010) *Guidelines for easy-to-read materials / rev.* IFLA Headquarters, The Hague.

- Perovic A. (2006) Syntactic deficit in Down syndrome: More evidence for the modular organisation of language. *Lingua* **116** (10), 1616-1630.
- Poncelas, A. & Murphy, G. (2007) Accessible information for people with intellectual disabilities: Do symbols really help? *Journal of Applied Research in Intellectual Disabilities*, *20*, 466-474.
- Ramos J.L. & Cuetos F. (1999) *Evaluación de los procesos de lectura en alumnos del tercer ciclo de educación primaria y educación secundaria obligatoria (PROLEC-SE)*. TEA, Madrid.
- Rondal, J.A. (2001) Language in mental retardatio;: Individual and syndromic differences, and neurogenitic variation. *Swiss Journal of Psychology*, *60*, 161-178.
- Ryan, R. M., & Deci, E. L. (2000) Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, *25*, 547-67.
- Ryder R., & Hughes M. (1985) The effect on text comprehension of word *frequency*. *Journal of Educational Research* **78**, 286-291.
- Sanders T., Land J. & Mulder G. (2007) Linguistic markers of coherence improve text comprehension in functional contexts. *Information Design Journal* **15**(3), 219-235.
- Szigriszt F. (1992) *Sistemas Predictivos de Legibilidad del Mensaje Escrito* (unpublished masters' thesis). Facultad de Ciencias de la Información, Madrid.
- Tronbacke B. (1997) *Guidelines for Easy-to-Read Materials*. IFLA, The Hague.

- Van Deursen, A.J.A.M., Van Dijk, J.A.G.M. & Peters, O. (2011). Rethinking Internet skills. The Contribution Of Gender, Age, Education, Internet Experience, And Hours Online To Medium- And Content-Related Internet Skills. *Poetics*, 39, 125–144.
- Verhoeven L. & Vermeer, A. (2006) Literacy achievement of children with intellectual disabilities and differing linguistic backgrounds. *Journal of Intellectual Disability Research* 50 (10), 725-738.
- Wolman C. (1991) Sensitivity to causal cohesion in stories by children with mild mental retardation, children with learning disabilities, and children without disabilities. *Journal of Special Education* 25, 135–154.

**Tables**

Table 1. The first and second columns show Ms, SDs for each level of linguistic measures of the text corpus (N48) selected for this study. The third and fourth columns show Pearson correlation coefficients between linguistic measures and literal and inferential comprehension questions.

	M	SD	r (n)	
			Literal questions (47 texts)	Inferential questions (45 texts)
<b>Lexical variables</b>				
N° of Words	93.8	40.9	-.30*	-.32*
Average syllables per word	1.9	0.1	-.14	.05
Average word frequency	21361	4427	-.16	.05
<b>Sentence variables</b>				
N° of sentences	11.6	5.6	-.30*	-.37*
Average words per sentence	8.3	1.5	-.03	.11
INFLESZ index	74.7	6.7	.08	-.11
<b>Textual cohesion variables</b>				
N° of connectives	4.0	3.3	-.35*	-.13
N° co-references (noun, argument and stem overlap)	12.5	6.2	-.39**	-.24
N° of ellipsis	2	2	-.22	-.24

## **Appendices**

Appendix 1. Example of one of the journalist texts used in this study and two of the literal and inferential questions which accompanied it (originals in Spanish).

### **Tickets for the match Real Madrid-Alcorcón are sold out**

The football match between the Real Madrid and the Alcorcón which will take place next Tuesday has piqued the curiosity of everybody and all tickets were sold out today. The stadium of the Real Madrid will be full. The result of the last match in the Alcorcón stadium, where Real Madrid lost 4-0, is the reason why this match is so interesting.

#### **Literal Question:**

***When will the match Real Madrid-Alcorcón take place?***

- 1) *Next Month.*
- 2) *Next Tuesday.*
- 3) *Today.*

The correct answer (Option 2) is explicitly stated within a single sentence of the text.

#### **Inferential Question:**

***Why are the tickets for the match Real Madrid-Alcorcón sold out?***

- 1) *Because the Alcorcón won Real Madrid 4-0 in the last match. (correct answer)*
- 2) *Because the camp of the Real Madrid will be full.*
- 3) *Because the match will be on Tuesday.*

The correct answer (Option 2) requires the integration of inter-sentence information.