Published in final edited form as:

Int J Lang Commun Disord. 2015 ; 50(3): 312–321. doi:10.1111/1460-6984.12134.

Extra-Linguistic Influences on Sentence Comprehension in Italian-Speaking Children with and without Specific Language Impairment

P. Pettenati^{a,b}, E. Benassi^c, P. Deevy^d, L.B. Leonard^d, and M.C. Caselli^e

^aDipartimento di Neuroscienze, Università di Parma, Italy

^bAccademia di Neuropsicologia dello Sviluppo, Parma, Italy

^cDipartimento di Psicologia, Università di Bologna, Italy

^dPurdue University West Lafayette, IN

^eIstituto di Scienze e Tecnologie della Cognizione, Consiglio Nazionale delle Ricerche, Italy

Abstract

Background—Many children with specific language impairment (SLI) have deficits in sentence comprehension. These deficits are usually attributed to limitations in the children's understanding of syntax or the lexical items contained in the sentences. In this study, we examine the role that extra-linguistic factors can play in these children's sentence comprehension.

Aims—In this study, extra-linguistic demands on sentence comprehension were manipulated directly by varying the nature of the materials used.

Methods & Procedure—Forty-five Italian-speaking children participated: 15 with SLI (M age = 4;5), 15 typically developing children matched for age (TD-A, M age = 4;5), and 15 younger typically developing children matched according to language comprehension test scores (TD-Y, M age = 3;9). The children responded to sentence comprehension items that varied in their length and/or in the number and type of foils that competed with the target picture.

Outcome & Results—The TD-A children were more accurate than the TD-Y children and the children with SLI, but, for all groups, accuracy declined when task demands increased. In particular, sentences containing superfluous adjectives (e.g., *Il topo bello copre l'uccello allegro* "The nice mouse covers the happy bird" where all depicted mice were nice and all birds were happy) yielded higher scores than similar sentences in which each adjective had to be associated with the proper character (e.g., *Il cane giallo lava il maiale bianco* "The yellow dog washes the white pig" where foils included a yellow dog washing a pink pig, and a brown dog washing a

Corresponding author: Paola Pettenati, Department of Neuroscience, University of Parma, Via Volturno 39/E, 43125, Parma, Italy, Telephone: +39 0521-903945, Fax: +39 0521-903900, paola.pettenati@unipr.it.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Author contributions: LBL and PD designed the original material used in this study; MCC, PP, EB adapted the test to the Italian language; PP and EB collected the data, organized the database and analysed the data; all the authors discussed the results and wrote the paper.

white pig). Many errors reflected recency effects, probably influenced by the fact that adjectives modifying the object appear at the end of the sentence in Italian.

Conclusions & Implications—Differences between conditions were observed even when lexical content, syntactic structure, and sentence length were controlled. This finding suggests the need for great care when assessing children's comprehension of sentences. The same syntactic structure and lexical content can vary in difficulty depending on the number and types of foils that are used in combination with the target picture.

Keywords

specific language impairment; Italian; sentence comprehension

Introduction

Weaknesses in sentence comprehension have long been a part of the clinical profile of children with specific language impairment (SLI). Problems of this type have been reported for a wide range of languages. These include, among others, Romance languages such as Italian (Junyent, 2011) and Spanish (Grinstead, De la Mora, Vega-Mendoza, & Flores, 2009), Germanic languages such as Danish (Jensen de López, Sundahl Olsen, & Chondrogianni, in press) and Swedish (Sahlén & Nettelbladt, 1991), Semitic languages such as Hebrew (Friedmann & Novogrodsky, 2011), and English (e.g., Bishop, Bright, James, Bishop, & van der Lely, 2000; Leonard & Deevy, 2010; Robertson & Joanisse, 2010). Comprehension deficits are sometimes seen even when such deficits are described as expressive in nature (Bishop, 1979, 1982; Magnusson & Naclér, 1990).

Despite the frequent documentation of sentence comprehension problems in children with SLI, two factors have complicated interpretation of the available evidence. The first is that, often, not the same type of knowledge is tapped when children's sentence comprehension and production are assessed (Leonard, 2009). Often the receptive task can overestimate a child's comprehension of a sentence construction, because, for example, foils provide lexical cues and therefore do not isolate the child's syntactic knowledge. However, in other cases, comprehension tasks may require an ability that goes beyond the ability needed for production. For example, in a very well-designed picture-pointing task designed to assess children's comprehension of third person singular *-s*, Johnson, deVilliers, & Seymour (2005) found that typically developing children could not perform above chance until five years of age. Yet typically developing children produce this morpheme consistently well before this age. It seems possible that children have a sense that the *-s* inflection 'fits' with singular nouns during their sentence production without having this inflection formally marked as present tense third person singular. The latter might be required in comprehension tasks.

The second factor complicating interpretation is that children with SLI often display weaknesses that extend beyond syntax and vocabulary. For example, studies show weaknesses in tasks of sustained attention (e.g., Finneran, Francis, & Leonard, 2009; Spaulding, Plante, & Vance, 2008), speed of processing (e.g., Miller, Kail, Leonard, & Tomblin, 2001; Leonard et al., 2007;), information storage (e.g., Graf-Estes, Evans, & Else-

Quest, 2007), and measures of working memory designed to capture both storage and mental manipulation (e.g., Archibald & Gathercole, 2006; Bavin, Wilson, Maruff, & Sleeman, 2005; Mainela-Arnold & Evans, 2005; Marton & Schwartz, 2003).

Significant correlations between weaknesses in these areas and performance on sentence comprehension tasks have been reported in several studies of SLI (e.g., Montgomery 2000; Montgomery & Evans, 2009; Norbury, Bishop & Briscoe, 2002; Robertson & Joanisse, 2010). Investigators have also attempted to measure this relationship more directly, by manipulating the presumed cognitive load in the sentences used in the sentence comprehension task (e.g., Montgomery, 2000; Robertson & Joanisse, 2010). These studies, too, have been informative, showing that children with SLI perform more poorly than typically developing peers when demands are increased in the sentences to be comprehended. However, in most studies, these demands have been increased by adding either syntactic complexity or by increasing sentence length through the inclusion of superfluous adjectives that are unnecessary for making a correct response. Unfortunately, the role of task demands independent of syntactic complexity or sentence length has not been clearly established in these studies.

In this study, we take a different task. To determine the role of factors other than syntax, sentence length, and vocabulary on children's sentence comprehension, we manipulate extra-linguistic factors such as the number and types of foils employed in the sentence comprehension task, while ensuring that the children have adequate command of the linguistic form and content employed in the sentences. Although these manipulations do not pinpoint the specific type of processing factor that is responsible for any effects on sentence comprehension accuracy, we view these manipulations as providing a critical first step in showing that task demands adversely affect the accuracy of children with SLI apart from any limitations they have in syntax, length, and vocabulary ability.

Our task is modeled after the work of Leonard, Deevy, Fey, and Bredin-Oja (2013), though we apply this task to children who are acquiring Italian. For reasons noted below, Italian differs from English in ways that could influence the interpretation of the findings. In the Leonard et al. study on English, sentences were limited to subject – verb – object (SVO) structures with adjectives (e.g., *The happy dog washes the little pig*) and without adjectives (e.g., *The dog washes the pig*). Leonard et al. first established that the children could comprehend the adjective meanings, and the SVO structure with and without adjectives. They then manipulated the amount of information that had to be detected, retained, and used in selecting the appropriate picture that corresponded to the stimulus sentence. The critical comparison was between a condition containing superfluous adjectives (e.g., *The happy dog washes the little pig*, where all dogs depicted were happy and all pigs depicted were little) and a condition in which all attributes had to be recalled with their appropriate nouns. For example, for the item *The yellow dog washes the white pig*, one of the foils depicted a yellow dog washing a pink pig, and another showed a brown dog washing a white pig. The reverse relationship with the appropriate adjective-noun pairings was also represented among the pictures. To respond to such items correctly, then, children had to grasp not only the proper SVO relationship but also remember the particular attribute associated with each character. Leonard et al. found that the accuracy of the children with SLI (*M* age 4;9)

dropped significantly in the latter condition relative to the superfluous adjective condition. The gap between these children and a group of typically developing same-age peers was also larger in this condition than in the superfluous adjective condition. The most common error in the high-demand condition was the selection of a picture reflecting the proper syntactic relationship (e.g., a dog washing a pig) but an incorrect attribute for the subject or object. A group of younger typically developing children (*M* age 3;5) performed much like the SLI group. These findings suggest that high task demands had an adverse affect on those children with less mature language skills, even when the children possessed the requisite lexical and syntactic knowledge to perform successfully. Age controls – whose language development level was more advanced – were not hindered when task demands were high.

In the present study, we employ the same type of manipulation in a sentence comprehension task designed for Italian. We compare Italian-speaking children with SLI and both same-age and younger typically developing children to determine whether such extra-linguistic demands affect sentence comprehension in Italian in the same way as in English. This issue is important to examine for three reasons. First, most adjectives in Italian follow, rather than precede the nouns they modify. Recall that the errors of English-speaking children in the high-demand condition most often involved the selection of a foil that reflected the appropriate SVO relationship but the wrong adjective associated with one of the nouns. Selections of an incorrect object adjective were just as likely as selections of an incorrect subject adjective. It is not clear if this pattern was related to the children placing equal importance on the subject noun phrase and object noun phrase, or if primacy effects (retention advantage for material near the beginning of a verbal string) and recency effects (retention advantage for material near the end of a verbal string) effects were playing an interacting role. In Italian, the adjective modifying the subject noun might have less of a primacy effect (as it appears closer to the middle of the sentence relative to English) whereas the adjective modifying the object should benefit more from a recency effect (because it appears at the end of the sentence). Based on this distributional information, we predict that more subject adjective errors than object adjective errors will be seen in the Italian data, thus showing a difference from English.

Second, by examining this issue in Italian, the generalizability of the contribution of extra-linguistic factors to children's sentence comprehension will be better understood. As this paradigm has been applied only to English, it is not yet clear if the Leonard et al. (2013) findings are applicable to other languages. In principle, replication is important, to ensure that the basic validity of the original findings. Replication across languages can have special importance. In the case of the Leonard et al. study, the children's ability to retain the particular adjective-noun pairings was crucial for success. In Italian, these pairings often have phonological cues, as in *coniglio piccolo* and *rana piccola*. These cues could render children's retention of the pairings somewhat easier.

Third, the assessment of sentence comprehension in Italian-speaking children with SLI is common clinical practice (e.g., Caselli, Monaco, Trasciani and Vicari, 2008). By discovering whether factors apart from lexical and syntactic comprehension affect these children's scores, practitioners will be better able to interpret the performance levels exhibited by children in their caseloads.

Method

Participants

Forty-five children served as participants. Fifteen children, 12 males and 3 females, ranging in age from 4;1 to 5;11 ($M = 4;5$, $SD = 5$ months) constituted the SLI group. All children in this group were currently receiving or had qualified for therapy focusing on receptive and/or expressive language skills. A second group of 15 children, 11 males and 4 females ranging in age from 4;1 to 5;11 ($M = 4;5$, $SD = 5$ months) were typically developing children who served as age controls, hereafter referred to as the “TD-A” group. These children were very similar to the children with SLI in chronological age, $t(28) = 0.01$, $p = .991$. The remaining 15 children, 9 males and 6 females ranging in age from 3;1 to 5;1 ($M = 3;9$, $SD = 7$ months) were younger typically developing children, referred to here as the “TD-Y” group. These children were matched to the children with SLI on the basis of raw scores on the Italian standardization of the *Peabody Picture Vocabulary Test* (Stella, Pizzioli, & Tressoldi, 2000), a test of receptive vocabulary, $t(28) = 0.48$, $p = .630$. The large number of items, and hence potential raw scores on this test permitted closer matching than would be possible for other available receptive language measures. These children were significantly younger than the children with SLI, $t(28) = 2.25$, $p = .033$. The TD-A and TD-Y children were recruited from local preschool and kindergarten classes and were functioning at age level in all areas according to parental and teacher report.

The children with SLI scored at age level on the *Leiter International Performance Scale - Revised* (1979) ($M = 111.87$, $SD = 8.91$, range = 91–127), a nonverbal test of intelligence. However, they fell significantly below age-level expectations in receptive and expressive language. Their standard scores on the Italian standardization of the PPVT ($M = 79.67$, $SD = 9.28$) were significantly lower than those of the TD-A children ($M = 99.07$, $SD = 13.06$), $t(28) = 4.69$, $p < .001$. Likewise, their percentages correct on another language measure, the *Test di Comprensione Grammaticale per Bambini* (TCGB), a test of receptive grammar (Chilosi & Cipriani, 1995) ($M = 74.93$, $SD = 15.00$), were significantly lower than the percentages correct for the TD-A group ($M = 92.00$, $SD = 4.00$), $t(28) = 4.26$, $p < .001$. The expressive language skills of the children with SLI were especially weak. Whereas the TD-A children scored within age level on the *Test di Ripetizione di Frasi* (TRF), a sentence imitation task that assesses grammatical production (Zardini, Battaini, Vender, & D’Angelo, 1985) (M z-score = 0.66, $SD = 0.37$, M raw score = 16.73, $SD = 3.10$), all children with SLI earned a z-score lower than -2.00 (M raw score = 3.40, $SD = 3.36$), $t(28) = 11.30$, $p < .001$. The children with SLI also showed significant deficits in nonword repetition ability, a clinical marker of SLI in Italian with excellent sensitivity and specificity (Bortolini et al., 2006; Dispaldro, Leonard, & Deevy, 2013). On a nonword repetition test of 24 items varying from one to four syllables in length (Bortolini et al., 2006), the children with SLI produced fewer nonwords correctly ($M = 10.20$, $SD = 6.52$) than the TD-A children ($M = 21.87$, $SD = 2.77$), $t(28) = 6.38$, $p < .001$.

As noted above, the TD-Y children were matched to the children with SLI on the basis of the PPVT. In addition, these two groups did not differ significantly on another receptive measure, the TCGB, $t(28) = 1.75$, $p = .091$. However, the TD-Y children earned higher

scores than the children with SLI on the expressive measures, the TRF ($M=13.27$, $SD=4.20$), $t(28)=7.11$, $p<.001$, and the nonword repetition test ($M=19.60$, $SD=3.68$), $t(28)=4.86$, $p<.001$.

Procedure

Adjective screening task—To ensure that the children understood the meaning of the adjectives, a 20-item screening test was administered to all children. Each item consisted of a request (e.g., *Mostrami il cane giallo* “Show me the yellow dog”) to point to one of two drawings shown on a computer screen. One of the drawings depicted the target picture, the other the same character with a different attribute (e.g., a brown dog). All groups averaged over 90% accuracy on these items, with no group difference observed, $F(2, 42)=2.19$, $p=.124$.

Experimental task—Thirty items were employed in the experimental task, divided into three sets of 10 items. The three sets were administered in counterbalanced order. One set of items consisted of simple reversible SVO structures (e.g., *Il coniglio insegue il gatto* “The bunny chases the cat”); the child had to point to the one drawing of the two on the computer screen that served as a match. The other drawing on the screen depicted the opposite relationship (e.g., a cat chasing a bunny). Accuracy was expected to be high for these items. For this reason, we refer to them as the “low-demand” items.

The second set of 10 items consisted of similar reversible SVO sentences, but contained a post-nominal adjective that modified the subject noun, and a different post-nominal adjective that modified the object noun (e.g., *Il topo bello copre l'uccello allegro* “the nice mouse covers the happy bird”). All adjectives were superfluous, that is, they were non-contrastive. They served to add length to the test sentences relative to those used in the low-demand items. One drawing represented the target picture and the alternative picture foil showed the opposite relationship (e.g., a bird covering a mouse). Although the adjectives did not provide a clue to the correct picture, their inclusion added to the length of the test sentences which could have resulted in more errors than in the low-demand condition. We therefore regarded these superfluous adjective items as “intermediate-demand” items.

The remaining 10 items presented a greater challenge to the children; we therefore refer to them as “high-demand” items. In terms of their lexical content, length, and syntax, they were indistinguishable from the intermediate-demand items. However, they differed from the intermediate-demand items in two important ways. First, instead of choosing between two drawings, the children had to choose among four drawings. Second, the alternative drawings required the children to understand not only the SVO structure but to retain each adjective and associate it with the appropriate character. For example, for the test sentence *Il cane giallo lava il maiale bianco* “The yellow dog washes the white pig,” one foil depicted the opposite subject-object relationship (a white pig washing a yellow dog), another showed the named subject with the correct attribute acting on the named object with an incorrect attribute (e.g., a yellow dog washing a pink pig), and the third foil portrayed the named subject with an incorrect attribute acting on the named object with the correct attribute (e.g., a brown dog washing a white pig). Note that a correct response to items of this type required

the children to understand the syntax and retain each adjective, each noun, and their proper association as they searched for the correct drawing. The children's failure to retain any portion of this information could lead them to select the wrong drawing.

It is important to point out that the low- and intermediate-demand items used in the current study differed from those employed by Leonard et al. (2013) because they involved two, rather than four alternatives for the children to choose from. These two-alternative items were based on similar items used by Leonard et al. as screening items to ensure that the children understood the vocabulary and syntax of the items of interest, and could perform accurately even with the added length created by superfluous adjectives. We reasoned that by using two-alternative SVO sentences with superfluous adjectives and four-alternative SVO sentences with contrastive adjectives as experimental items, we could increase the relative extra-linguistic demands on the children. We assumed that this effect would be smaller on the TD-A children, because Leonard et al. found that the accuracy of the age controls in their study did not differ between their two-alternative screening items (92% for SVO, 93% for SVO with superfluous adjectives) and the four-alternative counterparts (93% SVO, 91% SVO with superfluous adjectives). For the children with SLI, however, such a difference between the intermediate- and high-demand conditions was expected to be quite large.

By controlling for lexical content, syntactic structure, and sentence length, our four-alternative contrastive adjective items and the two-alternative superfluous adjective items differed greatly in their extra-linguistic demands – demands that represented the central interest in this study. However, in retrospect, it is apparent that this design prevented us from distinguishing between the composition of the alternative drawings and the number of alternative drawings as the primary extra-linguistic factor at work in the four-alternative contrastive adjective condition. We return to this issue in the Discussion.

We scored the children's responses for accuracy. For the high-demand items, each errant response was coded according to the foil that was selected in place of the target picture.

Results

The children's performance on the experimental task was examined through a mixed-model analysis of variance (ANOVA) with participant group (SLI, TD-Y, TD-A) as a between-subjects variable and demand level (low, intermediate, high) as a within-subjects variable. The number of correct responses served as the dependent measure. A significant difference was observed for demand level, $F(2, 84) = 18.45$, $p < .001$, $\eta^2_p = .290$. Post-hoc testing revealed that accuracy on the high-demand items ($M = 7.066$, $SD = 2.489$) was significantly lower than accuracy on both the intermediate-demand items ($M = 8.622$, $SD = 1.849$, $p < .001$, $d = 0.710$) and low-demand items ($M = 8.466$, $SD = 1.878$, $p < .001$, $d = 0.635$).

A significant effect was also seen for participant group, $F(2, 42) = 5.31$, $p < .01$, $\eta^2_p = .197$, with significantly greater accuracy by the TD-A children ($M = 9.089$, $SD = 1.240$), than by the children with SLI ($M = 7.156$, $SD = 2.619$, $p = .003$, $d = 0.943$). The TD-Y children's accuracy fell numerically between that of the other two groups ($M = 7.911$, $SD = 1.509$), but

did not differ significantly from either (TD-A versus TD-Y $p = .166$, $d = 0.853$; TD-Y versus SLI $p = .640$, $d = 0.353$). The participant group x demand level interaction was not significant, $F(4, 84) = 0.681$, $p = .607$, $\eta^2_p = .023$. This finding was unexpected, as a significant interaction was found by Leonard et al. (2013). In that study, only the SLI and TD-Y groups were adversely affected in the high-demand condition; the performance of the TD-A children did not show a significant decline. A summary of the findings appears in Table 1.

These findings indicated that demand level had a significant bearing on the performance of all three participant groups, with greater overall accuracy by the TD-A children. However, it was important to determine if the nature of the children's errors on the high-demand items differed across participant groups. For example, comparable accuracy scores could mean different things if the predominant error type for one group of children was the selection of the opposite relationship (a syntactic error) whereas for another group most errors were selections of foils with the proper relationship but the wrong attribute associated with the subject. The former could reflect a syntactic breakdown when demands were high, whereas the latter might involve failure to retain an attribute term appearing near the middle of the sentence. Recall that we predicted a larger number of errors representing choices of the foil with an incorrect subject attribute than choices of the foil with an incorrect object attribute, owing to expected recency effects. Accordingly, we selected those children from each participant group who showed at least 80% correct performance on the low- and intermediate-demand items. This criterion ensured that problems on the high-demand items were not due to weaknesses in understanding the simple SVO structure (as assessed in the low-demand items) or problems coping with longer sentences (as occurred in the intermediate-demand items). Nine children in the SLI group met this criterion, as did 14 children in the TD-A group and 11 children in the TD-Y group. In Table 2, we present the percentage of errors that constituted selections of the opposite relationship (a syntactic error), the selection of the foil showing the subject with the wrong attribute, and the selection of the foil showing the object with the wrong attribute.

As can be seen in Table 2, the distribution of errors was uneven for each participant group. Although the children whose data are reflected here showed at least 80% accuracy on low- and intermediate-demand items, they nevertheless occasionally selected a foil that represented a syntactic (reversal) error on the high-demand items. However, for all participant groups, these errors were outnumbered by selections of a foil that had the wrong attribute associated with the subject. Errors reflecting selections of a foil with the wrong attribute associated with the object were not as abundant, and were especially low in frequency for the SLI group. To test our prediction that selections of the incorrect attribute of the object would be less frequent than selections of the incorrect attribute of the subject due to recency effects, we performed a chi-square analysis of the number of children showing more subject attribute errors than object attribute errors and the number of children showing the reverse pattern, across the three groups of children. The difference was significant, chi-square (1) = 6.37, $p < .02$. The children were clearly less likely select a foil with an object attribute error. An inspection of Table 2 suggests that the children with SLI, especially, may have been relying in the most recent information (the adjective at the end of

the sentence). However, the numbers for each participant group taken separately were not sufficiently large to permit statistical analysis.

Finally, an inspection of the children's accuracy on individual items revealed no clear tendency to rely on phonological cues that reflected gender agreement between the adjective and noun. For example, an inspection of the data revealed that the item *Il coniglio bello tocca la tartaruga malata* was no more accurate than the item *Il gatto arancione bacia la scimmia marrone* even though only the first of these two items provided transparent phonological agreement cues.

Discussion

Given the similar length, lexical content, and syntax of the intermediate-demand and high-demand items, the significantly poorer performance associated with the high-demand items suggests that the extra-linguistic factors of number of foils (four rather than two) and foil type (rendering adjectives contrastive rather than superfluous) played a significant role in the children's accuracy. This finding is consistent with the Leonard et al. (2013) finding for English, and indicates that extra-linguistic factors operate even when the word order rules of the language are somewhat different, and many adjective-noun pairs offer phonological cues that mark agreement (e.g., *rana piccola*). Such a finding increases the generalizability of the finding that demands related to the search for and selection of an appropriate picture has an important bearing on a child's performance independent of the specific linguistic material being assessed.

As was found by Leonard et al. (2013), the TD-A group in the present study were significantly more accurate than the children with SLI, and no differences were seen between the SLI and TD-Y groups. However, unlike in the Leonard et al. study, we did not find a participant group x demand level interaction. Leonard et al. found that the TD-A children in their study did not show an appreciable decline in accuracy from the intermediate- to the high-demand items. In contrast, in the present study, we found that the decline from intermediate- to high-demand items was just as detectable in the TD-A group as in the other two groups.

There are two possible explanations for the differences in the TD-A groups' performance in the two studies. The most obvious is that, in the present study, the children selected from two alternative drawings in the intermediate-demand items and from four alternative drawings in the high-demand items, whereas Leonard et al. (2013) used four alternative drawings for all items. We had not anticipated the across-condition difference in our TD-A participants given that, as noted earlier, Leonard et al. observed uniformly high performance by their TD-A participants regardless of whether two alternatives or four alternatives were used for their superfluous adjective items. We had assumed that, as in the Leonard et al. study, our TD-A participants would maintain their high levels of accuracy across all conditions. The fact that this was not the case allows for the possibility that these children were adversely affected by the change from two alternative drawings in the intermediate-demand items to four alternative drawings in the high-demand items as well as, or rather than the particular types of foils used. In retrospect, then, this must be viewed as a limitation

in our study. On the other hand, the fact that both the number of foils and the types of foils constitute factors other than lexical content, syntax, and sentence length, we can nevertheless conclude that the extra-linguistic forces played a significant role in the findings.

A second possible reason for the differences in the performance of the TD-A children in the present study and the Leonard et al. (2013) study is that our TD-A participants were, on average, approximately four months younger. Although one must be cautious in comparing these two groups' performance given the language differences, the levels of accuracy of the Italian-speaking TD-A group on the low- and intermediate-demand items (all means above 9.00) were very similar to those seen for the English-speaking TD-A children on the two-alternative screening items involving SVO sentences and SVO sentences with superfluous adjectives (all means above 9.00). Yet for the high-demand items, which involved four alternative drawings in both studies, the Italian TD-A children had somewhat more difficulty ($M = 8.47$) than did their English-speaking counterparts in the Leonard et al. study ($M = 9.00$). It is possible that the slight age differences contributed to these apparent differences in performance on the high-demand items.

The similar accuracy levels of the SLI and TD-Y groups coupled with the similar decline in accuracy when confronted with high-demand items suggests that the performance of the children with SLI – including their response to increasing task demands – was probably a reflection of their lower language development level rather than a characteristic unique to a condition of impairment. However, we cannot rule out the possibility that the extra-linguistic task demands operated somewhat differently in the SLI and TD-Y children. The search among four drawings while retaining the proper adjective-noun pairings as well as the proper syntactic relations might involve a variety of processes, such as attention, storage, and rehearsal, among others. It is very possible that some of these processes develop earlier than others and thus they may have played a somewhat different role in the SLI and TD-Y groups. For example, for younger children such as those in the TD-Y group, attention might have been especially challenged in our task, especially for the high-demand items. With further maturation, these children might have achieved greater attention ability, thus removing this obstacle in higher demand contexts. For the children with SLI, a different process might have been the chief obstacle, one that should have, but did not reach the expected maturational level.

Although it is not yet clear which specific extra-linguistic factors operated in these two groups of children, it does appear that memory could have played a role. When we selected children from each group who demonstrated good comprehension of the SVO structure with and without (superfluous) adjectives, we found that errors on the high-demand items were not evenly distributed across the different types of foils. For each group, the most frequent type of error was the selection of the correct syntactic relationship, but the named subject had the wrong attribute (e.g., choosing a drawing of a brown dog washing a white pig instead of a yellow dog washing a white pig). As predicted, this attribute association problem was much less striking for the object position. This pattern held statistically across the children taken as a whole. Although the smaller numbers did not warrant statistical analysis within groups, an inspection of Table 2 suggests that the children with SLI may have been especially tuned to information appearing at the end of the sentence.

We are unaware of studies focusing specifically on recency effects in children with SLI. However, numerous studies have reported weaknesses in these children on tasks requiring them to recall sequences of digits (e.g., Conti-Ramsden, 2003; Gray, 2003) or to recall a word from each of several consecutive sentences as they judge the sentences' truth value (e.g., Ellis Weismer, Evans, & Hesketh, 1999). Given these weaknesses, it is probably not surprising that the children with SLI in particular would show the classic sign of retention difficulties in our high-demand condition – that of responding more accurately to the most recently appearing information in the sentence.

This pattern of errors differs from the Leonard et al. (2013) findings for English. In that study, errors were equally divided between subjects with the wrong attribute and objects with the wrong attribute. We believe that the cross-linguistic differences between adjective location in English and Italian were probably responsible for this difference in error distribution. Because adjectives follow the noun in Italian, the adjectives modifying the subject were closer to the middle of the sentence than they are in English. This would seem to render subject adjectives in Italian to be less amenable to primacy effects than in English. In contrast, adjectives modifying the object appear in sentence-final position in Italian, making them a likely beneficiary of recency effects, to a greater degree than in English. These differences match up well with the distribution of error types shown in Table 2.

The findings of this study have implications for both clinical assessment and treatment. Our task employed the same type of (picture-pointing) task used in the routine clinical assessment of language comprehension in preschool-age children. Yet, we found that even when children demonstrated adequate comprehension of the lexical content and syntax to be employed, the children's accuracy nevertheless varied significantly depending on the types and number of pictures used as foils. In clinical assessment, clinicians usually do not yet know if the child possesses adequate comprehension of the lexical and syntactic details contained in the test – indeed the test was administered to make this determination. For this reason, it is especially crucial that clinicians have an a-priori understanding of the degree to which the test items assess lexical and syntactic knowledge independent of processing demands. Comprehension scores could potentially represent an underestimation of children's grasp of the lexicon and syntax of their language if, as in the high-demand condition of the present study, extra-linguistic task demands are heavy.

Treatment, too, is probably influenced by the role that extra-linguistic demands might play in sentence comprehension. Specifically, when new syntactic structures are introduced to the child, task demands should be simplified so that potentially compounding factors are less likely to interfere with the child's discovery of the syntactic rules or principles to be learned. Later in treatment, as the child demonstrates greater knowledge of the syntax itself, the child's ability to sustain accuracy in the face of extra-linguistic demands (e.g., the inclusion of additional foils and their increasing similarity to the target picture) can be assessed to ensure that adequate comprehension levels can be maintained even in more difficult communication contexts.

The fact that similar findings emerged from English (Leonard et al., 2013) and Italian (the present study, suggests that extra-linguistic factors probably operate in children's sentence

comprehension in other languages. An important question for future research is whether there are languages whose typology renders them less vulnerable to these factors (e.g., a language that marks structural case on nouns), or even more vulnerable (e.g., a language with a sparse morphology that does not have a dominant word order) than has now been shown for both Italian and English.

Finally, the significant involvement of extra-linguistic task demands during sentence comprehension raises the question of the degree to which language test scores can be improved through activities designed to promote sustained attention, or rehearsal of stored information. Previous reports of remarkable gains in language test scores – as in the Tallal et al. (1996) study of gains through FastForWord – have subsequently been interpreted as being the result of increased attention to detail that, in turn, might have given children the ability to demonstrate their actual language knowledge on formal language tests (Leonard, 1998). Although seemingly paradoxical, it is possible that efforts to help children in skills related to operations such as attention, storage, and rehearsal may allow them to show significant improvements on measures that were actually designed to reflect language knowledge.

Acknowledgments

Funding for research: The work reported in this paper was supported by Fondazione Monte of Parma (Centre for Study on Children's Motor and Language Development, University of Parma) and Research Grant DC009574 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health, USA.

We are very grateful to Azienda Sanitaria Locale of Reggio Emilia, Azienda Sanitaria Locale and Fondazione Madonna della Bomba Centro Scalabrini Onlus of Piacenza and Ambulatorio Clinico "Evolvendo" of Mantova for their collaboration in data collection. We acknowledge as well the very helpful municipal educational services of the cities of S. Ilario of Reggio Emilia that allowed the data collection in their school. We especially thank the children who participated in the study and their parents.

References

1. ARCHIBALD L, GATHERCOLE S. Short-term and working memory in specific language impairment. *International Journal of Language and Communication Disorders*. 2006; 41:675–693. [PubMed: 17079222]
2. BAVIN E, WILSON P, MARUFF P, SLEEMAN F. Spatio-visual memory of children with specific language impairment: Evidence for generalized processing problems. *International Journal of Language and Communication Disorders*. 2005; 40:319–332. [PubMed: 16195191]
3. BISHOP DVM. Comprehension of spoken, written and signed sentences in childhood language disorders. *Journal of Child Psychology and Psychiatry*. 1982; 23:1–20. [PubMed: 6174536]
4. BISHOP DVM, ADAMS C, ROSEN S. Resistance to grammatical impairment to computerized comprehension training in children with specific and non-specific language impairments. *International Journal of Language and Communication Disorders*. 2006; 41:19–40. [PubMed: 16272001]
5. BISHOP DVM, BRIGHT P, JAMES C, BISHOP S, VAN DER LELY H. Grammatical SLI: A distinct subtype of developmental language impairment? *Applied Psycholinguistics*. 2000; 21:159–181.
6. BORTOLINI U, ARFÉ B, CASELLI MC, DEGASPERI L, DEEVY P, LEONARD L. Clinical markers for specific language impairment in Italian: The contribution of clitics and non-word repetition. *International Journal of Language and Communication Disorders*. 2006; 41:695–712. [PubMed: 17079223]
7. CASELLI MC, MONACO L, TRASCIANI M, VICARI S. Language in Italian children with Down syndrome and with specific language impairment. *Neuropsychology*. 2008; 22(1):27–35. [PubMed: 18211153]

8. CHILOSI, AM.; CIPRIANI, P. Test di Comprensione Grammaticale per Bambini. Tirrenia: Del Cerro; 1995.
9. CONTI-RAMSDEN G. Processing and linguistic markers in young children with specific language impairment (SLI). *Journal of Speech, Language, and Hearing Research*. 2003; 46:1029–1037.
10. DISPALDRO M, LEONARD L, DEEVY P. Real-word and nonword repetition in Italian-speaking children with specific language impairment: A study of diagnostic accuracy. *Journal of Speech, Language, and Hearing Research*. 2103; 56:323–336.
11. ELLIS WEISMER S, EVANS J, HESKETH L. An examination of verbal working memory capacity in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*. 1999; 42:1249–1260.
12. FINNERAN D, FRANCIS A, LEONARD L. Sustained attention in children with specific language impairment (SLI). *Journal of Speech, Language, and Hearing Research*. 2009; 52:915–929.
13. FRIEDMANN N, NOVOGRODSKY R. Which questions are most difficult to understand? The comprehension of Wh questions in three subtypes of SLI. *Lingua*. 2011; 121:367–382.
14. GRAFF ESTES K, EVANS J, ELSE-QUEST N. Differences in nonword repetition performance of children with and without specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research*. 2007; 50:177–195.
15. GRAY S. Diagnostic accuracy and test-retest reliability of nonword repetition and digit span tasks administered to preschool children with specific language impairment. *Journal of Communication Disorders*. 2003; 36:129–151. [PubMed: 12609578]
16. GRINSTEAD, J.; DE LA MORA, J.; VEGA-MENDOZA, M.; FLORES, B. An elicited production test of the optional infinitive stage in child Spanish. In: CRAWFORD, J.; OTAKI, K.; TAKAHASHI, M., editors. *Proceedings of the 3rd Conference on Generative Approaches to Language*. Cascadia Proceedings Project; Somerville, MA: 2009. p. 36-45.
17. JENSEN DE LÓPEZ K, SUNDAHL OLSEN L, CHONDROGIANNI V. Annoying Danish relatives: Comprehension and production of relative clauses by Danish children with and without SLI. *Journal of Child Language*. 2014; 41(1):51–83. [PubMed: 23200200]
18. JOHNSON V, DEVILLIERS J, SEYMOUR H. Agreement without understanding? The case of third person singular /s/. *First Language*. 2005; 25:317–330.
19. JUNYENT, A. Doctoral dissertation. Università di Padova; 2011. Individual differences in specific language impairment: Profiles of preschoolers exposed to Italian.
20. ROID, G.; MILLER, L. *Leiter International Performance Scale-Revised*. Chicago, IL: Stoelting; 1998.
21. LEONARD, L. *Children with specific language impairment*. Cambridge, MA: MIT Press; 1998.
22. LEONARD L. Is expressive language disorder an accurate diagnostic category? *American Journal of Speech-Language Pathology*. 2009; 18:115–123. [PubMed: 19029534]
23. LEONARD LB, DEEVY P. Tense and aspect in sentence interpretation by children with specific language impairment. *Journal of Child Language*. 2010; 37:395–418. [PubMed: 19698206]
24. LEONARD L, DEEVY P, FEY M, BREDIN-OJA S. Sentence comprehension in specific language impairment: A task designed to distinguish between cognitive capacity and syntactic complexity. *Journal of Speech, Language, and Hearing Research*. 2013; 56:577–589.
25. LEONARD L, ELLIS WEISMER S, MILLER C, FRANCIS D, TOMBLIN JB, KAIL R. Speed of processing, working memory, and language impairment in children. *Journal of Speech, Language, and Hearing Research*. 2007; 50:408–428.
26. MAGNUSSON E, NAUCLÉR K. Reading and spelling in language-disordered children. *Linguistic and metalinguistic prerequisites: Report on a longitudinal study*. *Clinical Linguistics and Phonetics*. 1990; 4:49–61.
27. MAINELA-ARNOLD E, EVANS J. Beyond capacity limitations: Determinants of word-recall performance on verbal working memory span tasks in children with SLI. *Journal of Speech, Language, and Hearing Research*. 2005; 48:897–909.
28. MARTON K, SCHWARTZ R. Working memory capacity and language processes in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*. 2003; 46:1138–1153.

29. MONTGOMERY J. Verbal working memory and sentence comprehension in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*. 2000; 43:293–308.
30. MONTGOMERY J, EVANS J. Complex sentence comprehension and working memory in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*. 2009; 52:269–288.
31. NORBURY CF, BISHOP DVM, BRISCOE J. Does impaired grammatical comprehension provide evidence for an innate grammar module? *Applied Psycholinguistics*. 2002; 23:247–268.
32. ROBERTSON E, JOANISSE M. Spoken sentence comprehension in children with dyslexia and language impairment: The roles of syntax and working memory. *Applied Psycholinguistics*. 2010; 31:141–165.
33. SAHLÉN, B.; NETTELBLADT, U. Patterns of vulnerability of language in children with severe developmental language disorders. In: SAHLÉN, B., editor. *From depths to surface: A case study approach to severe developmental language disorders*. Lund: Lund University; 1991. p. 149-163.
34. SPAULDING T, PLANTE E, VANCE R. Sustained selective attention skills of preschool children with specific language impairment: Evidence for separate attentional capacities. *Journal of Speech, Language, and Hearing Research*. 2008; 51:16–34.
35. STELLA, G.; PIZZIOLI, C.; TRESSOLDI, PE. Peabody – Test di Vocabolario Recettivo – PPVT-R Peabody Picture Vocabulary Test – Revised. Torino: Omega Edizioni; 2000.
36. TALLAL P, MILLER S, BEDI G, BYMA G, WANG X, NAGARAJAN S, SCHREINER C, JENKINS W, MERZENICH M. Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science*. 1996; 271:81–84. [PubMed: 8539604]
37. ZARDINI G, BATTAINI S, VENDER C, D'ANGELO A. Un test di ripetizione di frasi. *Analisi delle performances di un gruppo di bambini con disfasia di evoluzione*. *Giornale di Neuropsichiatria in Età Evolutiva*. 1985; 5:235–244.

Appendix A

SVO Control

1. The girl tickles the boy
La bimba tocca il bimbo
2. The cat kisses the monkey
Il gatto bacia la scimmia
3. The dog washes the pig
Il cane lava il maiale
4. The bunny chases the cat
Il coniglio insegue il gatto
5. The mouse covers the bird
Il topo copre l'uccello
6. The horse kicks the cow
Il cavallo calcia la mucca
7. The turtle splashes the frog
La tartaruga bagna la rana
8. The monkey feeds the tiger

La scimmia imbocca la tigre

9. The bunny touches the turtle

Il coniglio tocca la tartaruga

10. The bear pulls the elephant

L'orso tira l'elefante

Superfluous Adjectives

1. The happy turtle splashes the wet frog

La tartaruga felice bagna la rana bagnata

2. The little horse kicks the brown cow

Il cavallo piccolo calcia la mucca marrone

3. The brown monkey feeds the little tiger

La scimmia marrone imbocca la tigre piccola

4. The gentle bunny touches the nice turtle

Il coniglio buono tocca la tartaruga bella

5. The pretty cat kisses the nice monkey

Il gatto allegro bacia la scimmia bella

6. The happy bunny chases the orange cat

Il coniglio felice insegue il gatto arancione

7. The nice girl tickles the little boy

La bimba bella tocca il bimbo piccolo

8. The happy dog washes the little pig

Il cane felice lava il maiale piccolo

9. The nice mouse covers the pretty bird

Il topo bello copre l'uccello allegro

10. The nice bird pulls the little elephant

L'uccello bello tira l'elefante piccolo

Contrastive Adjectives

1. The yellow dog washes the white pig

Il cane giallo lava il maiale bianco

2. The little bunny chases the happy cat

Il coniglio piccolo insegue il gatto felice

3. The brown mouse covers the red bird
Il topo marrone copre l'uccello rosso
4. The white horse kicks the brown cow
Il cavallo bianco calcia la mucca marrone
5. The happy turtle splashes the little frog
La tartaruga felice bagna la rana piccola
6. The happy bear pulls the sad elephant
L'orso felice tira l'elefante triste
7. The brown monkey feeds the orange tiger
La scimmia marrone imbocca la tigre arancione
8. The nice bunny touches the sick turtle
Il coniglio bello tocca la tartaruga malata
9. The orange cat kisses the brown monkey
Il gatto arancione bacia la scimmia marrone
10. The nice girl tickles the happy boy
La bimba bella tocca il bimbo felice

What this paper adds'

Many children with specific language impairment (SLI) present deficits in sentence comprehension. However, studies also show that these children can have weaknesses in areas related to attention, memory, and processing speed. For this reason, it is important to determine whether such extra-linguistic factors can affect the sentence comprehension of these children. In this study, we examined the contribution of extra-linguistic factors (the number and types of foils used in a picture-pointing task) in sentence comprehension while keeping lexical content, syntactic complexity, and sentence length constant.

We found that accuracy in sentences comprehension declined when extra-linguistic task demands increased. This finding suggests that, in clinical assessment, it is especially crucial that clinicians have an a-priori understanding of the degree to which the test items assess lexical and syntactic knowledge independent of processing demands. In fact, comprehension scores could potentially represent an underestimation of children's grasp of the lexicon and syntax of their language if the extra-linguistic demands of test items are not minimized.

Table 1

Descriptive Statistics

Condition	SLI	TD-Y	TD-A
Low Demand	7.667 (2.380)	8.467 (1.767)	9.267 (0.961)
Intermediate Demand	7.733 (2.492)	8.600 (1.549)	9.533 (0.640)
High Demand	6.067 (2.790)	6.667 (2.350)	8.467 (1.685)

Mean Number of Correct Responses (and Standard Deviations) for the Three Groups of Children on the Low-Demand, Intermediate-Demand, and High-Demand Items.

Table 2

Distribution of Error Types in the High-Demand Condition. Values are the Percentage of Errors that Each Error Type Represents.

Error Type	SLI	TD-Y	TD-A
Reversal ^a	34%	25%	25%
Incorrect Subject Attribute ^b	58%	50%	42%
Incorrect Object Attribute ^c	8%	25%	33%

^aSelection of a foil depicting the syntactic relationship opposite that of the target sentence.

^bSelection of a foil depicting an appropriate character with an inappropriate attribute in subject position.

^cSelection of a foil depicting an appropriate character with an inappropriate attribute in object position.