Identification and Correction of Syntactically Unacceptable Sentences

Kerstin Nauclér & Eva Magnusson

Introduction

Many researchers have shown linguistic awareness, especially phonological awareness, to be related to reading (for a review see for instance Ball 1993). Most studies are concerned with linguistically normal subjects and in many of them good and poor readers' phonological awareness is compared. Such data show that there is a relation between phonological awareness and reading, but they cannot be used for predicting reading.

In our earlier studies we have looked not only at linguistically normal children but also at language-disordered ones (Magnusson & Nauclér 1993). In these studies we have shown language-disordered preschool children's phonological awareness to be less developed than that of a matched group of linguistically normal children. We also found the language-disordered group to score significantly lower on reading tests both at the end of grade 1, 3 and 4 (Magnusson & Nauclér 1990, 1991a, b). It is thus possible to predict reading from metaphonological data in preschool, not only for normal children but for language-disordered children as well. Since many, but not all language-disordered children have reading problems at school, it is important to find reliable means for predicting their reading success or failure.

Purpose

The purpose of this study is to explore the possibility of using morpho-syntactic awareness for predicting reading by letting language-disordered and linguistically normal children identify and correct sentences in which morphosyntactic rules have been violated.

The reasons for comparing morphosyntactic awareness and reading are the following:

– There are studies showing that good and poor readers differ in morphosyntactic awareness (Bowey 1986, Elbro 1990, Flood & Menyuk 1983, Fowler 1988, Ryan & Ledger 1979), but they are not designed to predict reading from morphosyntactic awareness.

– Earlier we have predicted reading from phonological awareness. As we are more inclined to look upon reading as comprehension than as decoding, it is reading comprehension that we are concerned with. It seems plausible that reading comprehension would be predicted even better from morphosyntactic awareness than from phonological awareness.

The main purpose of the study is thus to predict reading from morphosyntactic awareness, but we are also going to compare the subjects ability to identify and correct sentences, and to look at how linguistic ability and linguistic awareness are related.

Subjects

The subjects are 78 children, a group of 39 linguistically normal children and a group of 39 language-disordered children, individually matched as to age, sex, and non-verbal cognitive level. By language-disordered children we mean children who are normally developed in all other respects; they have normal hearing, normal cognitive development, no neurological disturbances, no physical malformations, etc. Only monolingual Swedish-speaking children are included. These two groups are part of a longitudinal study in which language-disordered and linguistically normal children are followed from preschool (age six) to grade 4 (age eleven) (Magnusson & Nauclér 1987, 1989, 1991a). This was done in order to try to predict reading and spelling development at school from linguistic and metalinguistic preschool data.

Procedure

In the part of the study which we are going to present here we are focusing on the children's morphosyntatic awareness. This has been assessed in two types of tasks. First, we looked at the children's ability to identify morphologically and syntactically acceptable and unacceptable sentences, and then at their ability to correct the sentences successfully identified as unacceptable. Twelve sentences were presented orally. In eight of them morphosyntactic rules were violated in four different ways by manipulation of either agreement within the NP, word-order, negation, or wh-questions. The remaining four sentences were morphosyntactically acceptable. The

sentences were presented randomly with the exception of the wh-questions, which were presented together at the end of the test because they demanded an instruction that differed from the other three structures. The sentences are listed in the appendix.

In order to facilitate the first task, that is the identification task, the children did not have to answer orally but were to indicate whether a certain sentence had been uttered by a young child (violated sentence, morphosyntactic unacceptable) or by his/her mother (non-violated sentence, morphosyntactic acceptable) by pointing at a picture showing either a young child or a woman. In the second task, the correction task, they had to correct all the sentences successfully identified as unacceptable, i.e. as uttered by the young child. The instruction given to the child to correct a sentence was "to say the sentence as the mother would have said it, the mother who knows how to speak properly".

Results

Since the study consists of two parts, the identification task and the correction task, we first show the results from the identification task for the two groups and then the results from the correction task. After that we will consider the influence of different sentence structures and then the relation between linguistic awareness and linguistic ability. Finally, the correlations between morphosyntactic awareness and reading will be shown.

Identification

As can be seen in Table 1, the 37 children in the language-disordered group were expected to identify 444 sentences, 296 unacceptable sentences and 148 acceptable ones. In the linguistically normal group (N = 36) the expected number of identification is 432, 288 unacceptable sentences and 144 acceptable ones. The language-disordered group identified 273 sentences as unacceptable and 171 as acceptable. The corresponding results for the linguistically normal group are 256 unacceptable sentences and 176 acceptable sentences. This implies that both groups identified a smaller proportion of the sentences as unacceptable than expected.

However, not all identifications made by the subjects were correct. The linguistically normal group was able to correctly identify 83% of the unacceptable sentences and 88% of the acceptable sentences, whereas the language-disordered group was a little less successful and identified 77% of the unacceptable sentences and 70% of all the acceptable ones. The overall

identification rate was 84% in the linguistically normal group and 75% in the language-disordered one. This is shown in Table 1.

Table 1. Number of identifications of unacceptable and acceptable sentences in the language-disordered group (N=37) and in the linguistically normal group (N=36).

	Language-disordered group			Linguistically normal group		
	unacc.	acc.	total	unacc.	acc.	total
expected	296	148	444	288	144	432
given	273	171	444	256	176	432
correct	229	104	333	238	126	364
% correct	77%	70%	75%	83%	88%	84%
false alarms	44	67		18	50	
% alarms	16%	39%		7%	28%	

Not only did the groups differ in identification rate; the language-disordered group is also more inclined to accept ungrammatical sentences than the linguistically normal group, as they identify 39% of the ungrammatical sentences as acceptable. The linguistically normal group identifies only 28% of the ungrammatical sentences as acceptable. Furthermore, the tendency to reject grammatical sentences is stronger in the language-disordered group than in the linguistically normal group, the former regarding 16% of the grammatical sentences as unacceptable, the latter only 7%. This mirrors the language-disordered subjects' lower ability to correctly identify both unacceptable and acceptable sentences.

Correction

The correction task is more difficult for both groups than the identification task, but it is more problematic for the language-disordered group than for the linguistically normal one. The language-disordered group is able to correct 93 sentences out of 229 sentences identified as unacceptable, the linguistically normal group 183 out of 238 sentences, which is shown in Table 2. Thus, the linguistically normal group successfully corrects 77% of the sentences judged as unacceptable, the language-disordered group only 41%.

There is no difference between boys and girls in either of the groups in their ability to identify and correct sentences.

Table 2. Number of corrections of unacceptable sentences in the language-disordered group (N=37) and in the linguistically normal group (N=36).

	Language-disordered group	Linguistically normal group
attempted	229	238
successful	93	183
% successful	41%	77%

Sentence structure influence

When we examine the influence of the four different morphosyntactic sentence structures on identification ability, we find that the scores vary depending on the structure. This is so in both groups, but the ranking order is not identical. As can be seen in Table 3 the linguistically normal group found negation and agreement easier to identify than word-order and *wh*-questions. They correctly identified 94% of the sentences with negation, 92% of the sentences testing agreement, 81% of the sentences assessing word-order and 78% of the *wh*-questions. The language-disordered group found it extremely difficult to identify word-order (52% correct), but identified both negation, agreement and *wh*-questions at approximately the same level (83%, 85% and 80% respectively).

Table 3. Proportion of correct identifications and corrections of violated structures (negations, agreement, word-order and *wh*-questions) in the language-disordered group (N=37) and in the linguistically normal group (N=36).

	Language-disordered group		Linguistically normal group	
	Identification	Correction	Identification	Correction
negation	83	25	94	65
agreem	85	51	92	83
w-order	52	56	81	93
wh-?	85	29	78	59

The different morphosyntactic structures seemed to have a stronger impact on the results in the correction task. The scores vary considerably depending on sentence structure. The ranking order is similar in the two groups, although the result in the language-disordered group is much lower, as was said before. Both groups find it easiest to correct sentences with violated word-order. Correction of incongruent sentences comes in the second place. Finally, sentences with negation and *wh*-questions are the most difficult ones to correct. The ranking order of the structures in the

correction task is not the same as in the identification task (see Table 3), with the exception of *wh*-questions. They were the most difficult for the linguistically normal group both to identify and to correct. Thus, there is no obvious relation between how difficult or how easy a certain morpho-syntactic structure is to identify and correct. The language-disordered group found word-order the most difficult structure to identify but the easiest one to correct.

The lack of relation between the subjects' ability to identify certain sentence structures and to correct them was further confirmed when we compared the corrections made by the best identifiers from both groups (subjects who scored 100% correct or had only one misidentification). There were five subjects from the language-disordered group and 20 from the linguistically normal group. In spite of their identical ability to identify both acceptable and unacceptable sentences, they differed in ability to correct sentences; the correction scores of the language-disordered group varied between 20% and 80% (with a mean of 54%), those of the linguistically normal group between 55% and 100% (with a mean of 77%).

Linguistic ability and linguistic awareness
When it comes to the relation between linguistic ability and linguistic awareness, it may seem plausible to assume that it is not possible to be aware of morphosyntactic structures, i.e. in this case to identify and correct structures, that one does not master linguistically. A child who is unable to produce a certain syntactic structure should not be able to identify such a structure as acceptable or unacceptable. However, the results of this study show that such an assumption is not true. Whether or not one is able to produce a certain structure does not seem to be decisive for being able to identify that structure. Language-disordered children who do not master a certain structure or only begin to do so, are able to identify that structure as easily as children who master it completely. This is shown in Table 4.

The picture for correction is different. Children who do not fully master a certain structure have many more problems in correcting such a sentence than children who master the structure entirely, as can be seen in Table 4. Surprisingly enough they are able to make *some* successful corrections.

Table 4. Number of correct identifications and corrections of unacceptable sentences in the language-disordered group made by children who do not master (–), who are beginning to master (+/–), and who master (+) the violated structures.

	Degree of mastering the violated structures		
	–	+/–	+
attempted identifications	57	45	282
correct identifications	45	36	205
% correct	79%	80%	73%
successful corrections	8	5	61
% successful	27%	20%	43%

Prediction of reading
In this study we did not find any correlations between preschool children's morphosyntactic awareness and their reading in grade 1 (see Table 5). We did not test the children in grade 2, but when they were tested again in grade 3 and 4, we found some correlations between reading ability and morphosyntactic awareness in preschool. This was so for the children from both the language-disordered and the linguistically normal group.

Table 5. Correlations between identification and correction tasks in pre-school and measures of reading and spelling in grade 1, 3 and 4 for the language-disordered (LD) group (N=37) and the linguistically normal (LN) group (N= 36).

	Identification		Correction	
	LD	LN	LD	LN
Word reading, gr 1	.15	.27	.04	.16
Sentence reading, gr 1	.27	.26	.25	.15
Word reading, gr 3	.21	.37*	.25	.12
Sentence reading, gr 3	.26	.12	.45**	.04
Text comprehension, gr 4	.39*	.28	.10	.39*
Spelling, gr 1	.32*	.42*	.55***	.22
Spelling, gr 3	.06	.25	.06	.13
Spelling, gr 4	.07	.22	.17	.15

*p≤.05, **p≤.01, ***p≤.001

The ability to *correct* morphosyntactically unacceptable sentences predicted reading better than *identification* ability, and the predictions were better for the subjects from the language-disordered group than for those from the linguistically normal one.

What is best predicted from morphosyntactic awareness is spelling in grade 1. In the language-disordered group spelling is best predicted from the correction task and in the linguistically normal group from identification.

Discussion

The purpose of this study was to develop better predictions for reading. Therefore, it was disappointing that we did not find any correlations in our data between morphosyntactic awareness in preschool and reading in grade 1. The reading tests that we have used measure comprehension of words and sentences. We have earlier shown that reading comprehension can be predicted from phonological awareness. Morphosyntactic awareness should therefore be an even better predictor. The reason for this not being true could be that reading in grade 1 can be performed without morphosyntactic awareness. Since we found correlations with reading comprehension in grade 3 and 4, it might be that this kind of awareness only predicts reading of the linguistically more demanding texts that are used later in school. We could speculate along the same lines about spelling as a more linguistically demanding task than reading. This would explain why we get correlations between morphosyntactic awareness in preschool and spelling in grade 1, but no correlations between awareness and reading in grade 1. The fact that we do not find any correlations with spelling in grade 3 and 4 is puzzling, as we would have expected spelling in higher grades to be an even more demanding linguistic task than spelling in first grade.

We have tried to predict reading from preschool data, i.e. there is at least one year between the awareness tests and the reading tests, whereas other studies use data on awareness and reading from the same point in time (e.g. Fowler 1988, Forrest-Pressley 1983). Therefore, one might argue like Ehri 1979 and Ryan & Ledger 1984 that morphosyntactic awareness is not a prerequisite for reading, but a consequence of it. However, since we find correlations between morphosyntactic awareness in preschool and reading in grade 3 and 4, morphosyntactic awareness cannot be entirely a consequence of reading.

Fowler 1988 found correlations between reading and correction tasks in grade 2, but no correlations for identification tasks. In a study where children's word-order awareness was tested at five to seven years of age, Bohannon et al. 1984 found significant correlations between identification of word-order and reading a year later. From our study we cannot conclude that either identification or correction is the better predictor. Both tasks predict spelling in grade 1, identification in the linguistically normal group and correction in the language-disordered one. Reading in grade 3 and 4 is predicted to some extent by both identification and correction.

Morphosyntactic awareness is not a simple concept. This is shown by the fact that the two tasks we have used in this study obviously tap different degrees of morphosyntactic awareness – the correction task being much more demanding than the identification task. This is in accordance with for instance Ryan & Ledger 1984. Fowler 1988, however, found no difference between identification and correction tasks. However, in her study the subjects were not asked to correct sentences judged by themselves as ungrammatical, but their task was to correct ungrammatical sentences randomly selected from the identification task. In a study of language-disordered children conducted by Rubin et al. 1990, error identification was found to be a more difficult task than correction. This was explained by the authors in the following way: error identification requires more explicit awareness or a higher level of analysis of linguistic knowledge than correction. They seem to view correction as a more or less automatic process which can be run intuitively without necessarily being aware of the grammatical error or without explicit linguistic knowledge. This is not in accordance with our data, which clearly show correction to be more demanding than identification. One difference between our study and that of Rubin and her co-workers is that our language-disordered subjects are considerably younger, and also that our linguistically normal subjects found correction more difficult than identification.

It is often assumed that linguistic knowledge precedes linguistic awareness. This view is for instance taken by Kamhi & Koenig 1985, Leonard et al. 1977, Liles et al. 1977, who claim that children do not exhibit metalinguistic awareness of a language structure before they use the structure expressively. By comparing language-disordered children who do or do not give indications in their production of mastering the structures they have been asked to identify and correct, we have been able to show that identifications may also be made by children who do not master these

structures, whereas the more demanding task, i.e. correction, is managed more seldom by these children. This does not necessarily imply that they are lacking in awareness, only that their linguistic restrictions make corrections impossible. According to Fowler 1988:75 "it is not clear whether failure on sentence-level tasks derives from fundamental deficits in acquiring and processing language structures, or whether such failure is a direct reflection of an inability to access and manipulate available structures in playing metalinguistic games". Another explanation is given by Ljubesic & Kovasevic 1992:41 who argue that the reason for language-disordered children's poor performance is "their inability to implement their language knowledge in the sentence context".

Children's performance on identification and correction tasks is influenced not only by their language ability and degree of linguistic awareness, but also by what kind of syntactic and morphologial structures they are asked to deal with. In both the language-disordered group and the linguistically normal one we see an effect of sentence structure on the results. In the correction task, where this influence was more evident the ranking order was similar in the two groups, although the language-disordered children performed at a much lower level. Thus, it is not only the language-disordered children's deviant or less developed linguistic ability which is behind this variation.

There are many questions that remain to be answered about the interrelations between linguistic ability and linguistic awareness as well as between the knowledge of different syntactic structures and levels of linguistic awareness. Some of the questions have not yet been asked. Furthermore, methodological issues have to be considered carefully before we are able to make more reliable predictions for reading from preschool data.

References

Ball, E.W. 1993. 'Phonological awareness. What's important and to whom?' *Reading and Writing* 5, 141-159.

Bohannon, J.N., A. Warren-Leubecker & N. Hepler. 1984. 'Word-order awareness and early reading'. *Child Development* 55, 1541-1548.

Bowey, J.A. 1986. 'Syntactic awareness and verbal performance from pre-school to fifth grade'. *Journal of Psycholinguistic Research* 15, 285-308.

Ehri, L. 1979. 'Linguistic insight: Threshold of reading acquisition'. In T.G. Waller & G. E. MacKinnon (eds.), *Reading research: Advances in theory and practice (Vol 1)*, 63-114. New York, NY: Academic Press.

Elbro, C. 1990. *Differences in dyslexia. A study of reading strategies and deficits in a linguistic perspective.* Copenhagen: Munksgaard.

Flood, J. & P. Menyuk. 1983. 'The development of metalinguistic awareness and its relation to reading achievement'. *Journal of Applied Developmental Psychology* 4, 65-80.

Fowler, A.E. 1988. 'Grammaticality judgements and reading skill in grade 2'. *Annals of Dyslexia* 38, 73-94.

Forrest-Pressley, D.L. 1983. *Cognitive and meta-cognitive aspects of reading.* New York, NY: Springer-Verlag.

Kahmi, A.G. & L.A. Koenig. 1985. 'Metalinguistic awareness in normal and language-disordered children'. *Language, Speech and Hearing Services in School* 16, 199-210

Leonard, L., J. Bolders & R. Curtiss. 1977. 'On the nature of children's judgment of linguistic features: Semantic relations and grammatical morphemes'. *Journal of Psycholinguistic Research* 6, 233-245

Liles, B., M. Shulman & S. Bartlett. 1977. 'Judgments of grammaticality in normal and language-disordered children'. *Journal of Speech and Hearing Disorders* 42, 199-209

Ljubesic, M. & M. Kovasevic. 1992. 'Some insights into specific language impairment in Croatian'. *Scandinavian Journal of Logopedics & Phoniatrics* 17:1, 37-44

Magnusson, E. & K. Nauclér. 1987. 'Language disordered and normally speaking children's development of spoken and written language. Preliminary results from a longitudinal study'. *Reports from Uppsala University, Department of Linguistics (RUUL)* 16, 35-63.

Magnusson, E. & K. Nauclér. 1989. 'Om projektet SPRUT – Läsning och stavning i språkligt utvecklingsperspektiv'. *Rapport från Andra Nordiska Barnspråkssymposiet,* Lund 1–2 december 1988, 119-126.

Magnusson, E. & K. Nauclér. 1990. 'Reading and spelling in language-disordered children – linguistic and metalinguistic prerequisites: report on a longitudinal study'. *Clinical Linguistics and Phonetics* 4, 49-61.

Magnusson, E. & K. Nauclér. 1991a. 'Hur ser läsutvecklingen ut hos normalspråkiga och språkstörda barn under de första skolåren?'. In B. Hene & S. Wahlén (eds.), *Barns läsning och läsutveckling,* 118-136. Rapport från ASLAs Höstsymposium november 1990, Stockholm.

Magnusson, E. & K. Nauclér. 1991 b. 'On the development of reading in good and poor readers'. *International Journal of Applied Linguistics* 1, 174-185.

Magnusson, E. & K. Nauclér. 1993. 'The development of linguistic awareness in language-disordered children'. *First Language* 13, 93-11.

Rubin, H., M. Kantor & J. Macnab. 1990. 'Grammatical awareness in the spoken and written language of language-disabled children'. *Canadian Journal of Psychology* 44 (4), 483-500.

Ryan, E.B. & G.W. Ledger. 1979. 'Grammaticality judgments, sentence repetitions and sentence corrections of children learning to read'. *International Journal of Psycholinguistics* 6, 23-40.

Ryan, E.B. & G.W. Ledger. 1984. 'Learning to attend to sentence structure: Links between metalinguistic development and reading'. In J. Downing & R. Valtin (eds.), *Language awareness and learning to read*, 149-172. New York, NY: Springer-Verlag.

Appendix

Acceptable and unacceptable sentences used in identification and correction tasks.

Acceptable sentences	Unacceptable sentences	Corrected sentences
Agreement		
Vår bil är ny.	*Min katten är snällt.	Min katt är snäll.
	*Våra huset är liten.	Vårt hus är litet.
Word-order		
Där ligger katten.	*Här vi bor.	Här bor vi.
	*Nu slut glassen.	Nu är glassen slut.
Negation		
Flickan vill inte leka.	*Katten nej sova.	Katten vill inte sova.
		Katten sover inte.
	*Kalle nej äta.	Kalle vill inte äta.
		Kalle äter inte.
Wh-questions		
Vad ska katten äta?	*Katten göra?	Vad gör katten?
		Vad ska katten göra?
	*Pojken heter?	Vad heter pojken?
		Vad ska pojken heta?

Automatic Translation of Knitting Instructions – KNITTRA

Anders Nordner

Introduction

The purpose of this paper is to present the results from the work with automatic translation of knitting instructions, KNITTRA – a part of the SWETRA project. SWETRA has earlier treated weather and stock-market reports (cf. Sigurd et al. 1990, 1992).

The languages treated in KNITTRA are Danish and Swedish. Swedish is the target language and the system is not bi-directional. KNITTRA deals with a domain which is as normalised as weather reports, but the vocabulary and the phraseology are different. The sentence types are mainly imperatives and passives apart from many special abbreviated and elliptic sentences.

The translation is carried out by a source and a target language module, each consisting of a lexicon, morphological rules which expand the lexicon, and grammar rules. The source module delivers a syntactic-semantic representation which is used as an interlingua and is fed into the target language module.

KNITTRA first tries to do a complete syntactic analysis of the sentence in order to make a sentence-by-sentence translation. If this does not work, KNITTRA uses a phrasal approach. Finally, a word-for-word translation is also implemented, which is utilised if neither the sentence-by-sentence nor the phrasal approach succeed. A time variable can be set, to three minutes for instance, allowing the program to try sentence-by-sentence translation for that amount of time before the phrasal translation is started. If the time allowed for the phrasal translation is not enough, a word-for-word translation is started. The word-for-word method will always produce a translated sentence or phrase. The quality of the translation depends on the method used. The best results emerge from the sentence-by-sentence translation. The phrasal approach works well with the languages treated,