Psychology of Language and Communication 1999, Vol. 3. No. 2

EMILIA ŁOJEK
University of Warsaw

DISCREPANCIES BETWEEN OBJECTIVE AND SUBJECTIVE MEASUREMENTS OF DISCOURSE IN APHASIC AND RIGHT-BRAIN-DAMAGED SUBJECTS

Script knowledge in discourse of 15 mild to moderate aphasics, 15 right-brain-damaged (RHD) subjects and 15 normal controls was studied. The experimental tasks consisted of retelling stories, describing daily life situations and providing plans of action in a given situation. Objective and subjective assessments of correctness of discourse content and order were used. The results revealed that aphasics' discourse was comparable to that of controls on the objective measures, whereas it was significantly impaired on the subjective ratings. In contrast, subjective ratings revealed fewer disorders than the objective ones in the RHD group. Various possible reasons of these differences are considered.

One of the most important findings from the pragmatic approach in studies of brain-damaged subjects is the demonstration of discrepancies between various measurements, which either include or exclude the natural context of communication. In the nineteen sixties Sarno (1969) identified the "clinic-nature gap". This is when there is an improvement in an impaired function under conditions arranged in a clinic, but this improvement is not reflected in natural circumstances.

Studies of discourse in the nineteen eighties revealed differences between impaired performance in aphasics at word/sentence levels and their preserved abilities to process information within a linguistic text. Despite their speech/language problems aphasics are also able to utilize information in the context of communication and can successfully obey the turn-taking rules in conversational discourse. On the other hand, studies on discourse in right-hemisphere-damaged subjects showed that, although purely linguistic features of their utterances are preserved, they have many problems with discourse production and comprehension (Joanette & Brownell, 1990).

Thus, the pragmatic orientation in neurolinguistic research has highlighted the great importance of context as a factor which significantly influences the performance of brain-

---

Requests for copies should be sent to the author at the University of Warsaw, Faculty of Psychology, Stawki 5/7, 00-183 Warszawa, Poland. E-mail: emilia@sci.psych.uw.edu.pl

damaged individuals. It has also demonstrated that different methods of assessment can produce conflicting pictures of subjects' abilities.

The main purpose of this paper is to report unexpectedly large discrepancies between objective and subjective discourse ratings used to estimate correctness of text contents and order. The discrepancies were revealed through analysis of discourse in aphasics and right-hemisphere-damaged patients.

Method

Subjects

Fifteen aphasics with left-hemisphere-damage (LHD) and 15 right-hemisphere-damaged patients took part in the investigation. Thirty healthy subjects (C1) were included for initial examination to elicit model discourse samples. A subgroup of 15 controls (C) was then selected from the 30 healthy subjects for the purpose of comparison with the experimental groups. All subjects were right-handed native speakers of Polish.

Table 1. Description of sample

Measure	LHD Mean (SD)	RHD Mean (SD)	C Mean (SD)	C1 Mean (SD)	Ratio
Age (years)	55.1(10.5)	54.4(11.1)	54.1(8.9)	55.1(8.1)	F(3,71)=0.28
Education (years)	14.0(2.5)	13.6(2.6)	14.3(2.5)	14.0(2.2)	F(3,71)=0.33
Months postonset	42.2(52.9)	40.5(42.6)			T(26) =0.93

As Table 1 shows, there were no significant intergroup differences in the mean age, education level, and duration of illness. The groups consisted of comparable numbers of male (M) and female (F) subjects (LHD, M=11, F=4; RHD, M=11, F=3; C1 M=22, F=8, C2, M=11 F=4).

According to the neurological data, the aetiology of brain damage was CVA in 14 aphasics and 13 RHD patients. The remaining patients suffered from a trauma. In 13 aphasics and 11 RHD subjects the lesions spread across more than one lobe of the left or the right hemisphere. In the remaining patients, lesions involved only the left of the right temporal lobe. Nine aphasics and 10 RHD patients exhibited left versus right hemiparesis. None of the patients in the study (except for one aphasic who had had a short ischaemic episode) had suffered previously from any other neurological condition. There were no clinically observable signs of dementia in either experimental group.

The Boston Diagnostic Aphasia Examination (BDAE) was used to classify the language impairments of the aphasic subjects. Eight aphasics were described as moderate and 7 cases as mild. From the speech disorders profile 7 subjects were classified as non-fluent, 8 as fluent, aphasics.

The general level of intellectual functioning measured by the full scale of the WAIS was average in the aphasic (Mean = 98,9, SD = 8,2) and RHD (Mean = 100,3, SD = 9,6)

patients and high in the C group (Mean = 111,1 SD = 7,6). The LHD and RHD patients scored significantly lower than the C group on the full scale ($F(2,42) = 9,16$, $p<.01$), and on nonverbal performance ($F(2,42) = 13,57$, $p<.001$). The aphasics obtained markedly lower results on the verbal scale than the RHD and C subjects ($F(2,42) = 7,96$, $p<.01$).

Materials

A battery of experimental tasks was designed to elicit discourse. The battery used in the present investigation consisted of the following tasks:

STORIES – retelling two short stories ("Robbery" and "Mr Kowalski");

SCRIPTS – describing two daily life situations – that is, scripts of borrowing money and asking the way in an unfamiliar town;

PLANS – providing two courses of action usually taken when money or information is lacking.

The experimental tasks were administered in one 60-minute session in hospitals or in the subjects' homes. The discourse responses were tape-recorded.

Methods for the objective analysis of discourse

For the purpose of objective discourse analysis the tape-recorded data were transcribed verbatim. All filler words and phrases (eg. "well", "you know") and false starts were excluded from the scoring. Discourse samples produced by 30 healthy subjects were used to develop sets of model units of information for each story, script and plan. The model units of information represented chunks of semantic information and were distinguished on the basis of sentences. In that sense the meaning of the term was concordant with the meaning of the term "proposition" (Frederiksen et al., 1990). However, because of wide variety between the types of discourses and number of information units produced, the amount of semantic content represented by a model unit of information had to be different in each type of text (Appendix 1). Thus, the usage of the term "model units of information" seemed to be more appropriate than "proposition".

A model unit of information for the stories reflected the content of the preliminary texts read by the experimenter. Some of the model units of information for scripts represented a general category of information (eg. "Other options if unsuccessful"). That kind of information was chosen in order to cope with analysis of various possible contents pertaining to scripts and plans provided by normal controls. Using the notion of more abstract category of information was absolutely necessary in the case of plans. The model units of information for plans were divided into two categories: abstract and concrete elements. Abstract elements included the most general steps taken in the course of action to attain a goal. Concrete elements described different concrete actions or ways of reaching a goal. Some of the subjects produced arguments for taking into account a given prerequisite of an action or gave short descriptions of particular methods of action. Such descriptions were treated as one unit of content referring to one abstract or concrete element. In both plans the concrete elements usually followed the abstract ones. The order of model information in all discourse tasks was established using the mean position of each information scored on the basis of production of 30 normal controls.

The aphasics' and RHD patients' discourses were subsequently compared with the model sets of information. The variables analyzed included:

– the number of information units from the model sets present in a discourse,

– the order of model information, measured by an Index of Wrong Order (IWO) – quantitative computation of the number of shifts in the sequential order of each unit of information in comparison with the model position irrespective of number of units of information produced. The idea of the IWO was based on Kendall's Tau Rang Coefficient.

– the impairment of discourse coherence reflected in production of additional commentaries and remarks, measured by number of inappropriate propositions.

Methods for the subjective analysis of discourse

Discourse variables which were evaluated subjectively included: appropriateness and order of utterances. The measure of appropriateness referred to the plausibility and conventionality of text information. It was rated in response to the question: "How adequate in terms of the raters' knowledge was the content of a given text as a response to a particular question?" The measure of order referred to the formal construction of a discourse. The raters evaluated the correctness of the order of provided information. Both subjective dimensions were estimated on five-point scales (from 1 to 5 points).

Four independent raters (3 females and 1 male) performed the subjective discourse analysis. All raters were qualified psychologists but none of them were speech pathologists. They were not familiar with the subjects.

The materials for the raters consisted of the original contents of the experimental discourse tasks, the audiocassettes with randomized texts, and the rating system containing the definitions of appropriateness and order of discourse. The raters were informed that some of the subjects might demonstrate language problems, such as difficulty in finding words or wrong pronunciation of speech sounds. It was also stressed that these kinds of language problems are distinct from the dimensions to be evaluated, which refer to the content and information expressed in a text. Additionally, sample data were provided for the raters.

The consistency of obtained subjective ratings was estimated using the Chi Test for each discourse sample. For 86% of the results, the consistency of ratings was high and very high, as well as statistically significant. The only cases where the consistency was relatively low and nonsignificant were a few ratings for scripts and plans produced by single control subjects. The raters showed unequivoced agreement in their ratings of aphasic and RHD patients' discourse.

Results

Results of the objective discourse analysis

Overall across the three task types, LHD and RHD patients produced significantly fewer model units of information than normal controls. However, taking each of the six tasks in turn, intergroup differences only reached significance in two cases (see Table 2).

As Table 2 summarizes, LHD patients produced the smallest number of model units of information in the second plan compared with normal controls. Aphasics tended to omit important information pertaining to more abstract considerations of a given action. In the other discourse tasks the scores of aphasics did not significantly differ from those of control subjects.

RHD subjects scored significantly lower than the remaining groups in the first story, but the other results in that group were comparable to the scores of the control group.

Table 2. Number of model units of information

Task	LHD Mean (SD)	RHD Mean (SD)	C Mean (SD)	Ratio F(2.42)	Intergroup Differences
Story 1	7.93(1.94)	5.53(2.03)	8.13(1.76)	8.54**	LHD,C>RHD
Story 2	7.53(1.84)	7.13(2.16)	8.40(1.80)	1.66	
Script 1	4.33(1.44)	4.53(1.68)	5.00(1.30)	0.79	
Script 2	4.53(1.50)	4.20(1.56)	4.73(1.33)	0.50	
Plan 1	2.73(1.58)	3.73(1.83)	3.86(1.59)	2.06	
Plan 2	3.00(2.30)	4.13(1.76)	5.13(1.84)	4.24*	C > LHD
Summary	29.46(6.52)	29.20(5.90)	35.80(5.06)	6.10*	C>LHD,RHD

*p<0.05
**p<0.001

Analysis of the sequential order of information using the IWO indicated no significant differences between the results of aphasics and control subjects (see Table 3).

Table 3. Results on the Index of Wrong Order

Task	LHD Mean (SD)	RHD Mean (SD)	C Mean (SD)	Ratio F(2.42)	Intergroup Differences
Story 1	0.00(0.01)	0.01(0.03)	0.01(0.02)	0.62	
Story 2	0.01(0.03)	0.01(0.03)	0.00(0.00)	1.59	
Script 1	0.07(0.09)	0.08(0.14)	0.01(0.05)	1.91	
Script 2	0.07(0.11)	0.09(0.08)	0.06(0.05)	0.50	
Plan 1	0.03(0.09)	0.16(0.16)	0.07(0.09)	4.66*	RHD>LHD
Plan 2	0.04(0.07)	0.12(0.11)	0.13(0.11)	3.75*	C,RHD>LHD
Summary	0.24(0.25)	0.51(0.22)	0.30(0.15)	7.04**	RHD>LHD,C

*p<0.05
**p<0.01

The largest number of shifts in the ordering of model units of information was observed in the RHD group. Significant differences between the results of the RHD subjects and the other groups were detected in the formulation of both plans but not of stories or scripts. However, the total amount of order errors produced by the RHD patients on all discourse tasks was significantly higher than that of the remaining groups.

Additional commentaries and remarks were present in the discourse of all groups. As Table 4 shows, no significant differences were found between aphasics and the control group on the number of inappropriate contents. LHD patients produced the highest number of additional commentaries in the first plan but the scores did not significantly differ from the results of the remaining groups.

Table 4. Number of inappropriate commentaries and remarks

Task	LHD Mean (SD)	RHD Mean (SD)	C Mean (SD)	Ratio F(2.42)	Intergroup Differences
Story 1	0.20(0.56)	0.60(0.91)	0.06(0.25)	2.87	
Story 2	0.20(0.77)	1.00(1.30)	0.20(0.41)	3.86*	RHD>C,LHD
Script 1	0.26(0.79)	0.33(0.61)	0.13(0.51)	0.36	
Script 2	0.53(1.24)	1.93(3.19)	0.33(0.48)	2.85	
Plan 1	2.00(3.44)	0.93(1.79)	1.06(1.33)	0.90	
Plan 2	0.53(1.55)	1.33(3.01)	0.26(0.59)	1.17	
Summary	3.73(5.24)	6.13(7.78)	1.73(1.66)	2.40	

*p<0.05;

The RHD patients produced more additional commentaries and remarks than the other two groups. They scored significantly higher that the LHD and control subjects in the second story. Additional analysis using T test also revealed significant differences between the RHD and the healthy subjects in the first story (T test(16)=2.18 p<0.05), and on the overall number of inappropriate contents in all discourse tasks (T test(15)=2.14 p<0.05).

Results of the subjective analysis of discourse

Results of the subjective discourse analysis revealed that the texts produced by the aphasics were usually rated as exhibiting midrange levels of both appropriateness and order of discourse.

Table 5. Subjective ratings of appropriateness of discourse

Task	LHD Mean (SD)	RHD Mean (SD)	C Mean (SD)	Ratio F(2.42)	Intergroup Differences
Story 1	17.26(2.28)	16.60(3.54)	18.53(1.30)	2.23	
Story 2	16.66(2.05)	17.26(2.81)	19.06(0.88)	4.42**	C>LHD,RHD
Script1	15.73(2.76)	17.60(3.08)	18.20(2.59)	3.11*	C>LHD
Script 2	17.20(3.59)	17.60(3.99)	18.86(2.26)	1.00	
Plan 1	13.66(3.17)	14.93(2.54)	14.06(1.79)	0.95	
Plan 2	12.53(2.99)	16.60(2.82)	15.33(1.44)	10.23**	RHD,C>LHD
Summary	93.07(9.79)	100.00(10.61)	103.87(5.32)	5.69**	C>LHD

*p<0.05
**p<0.01

As Table 5 indicates, no aphasic's utterance was estimated as absolutely inadequate or disordered by any of the four raters. However, the ratings of discourse in the remaining groups were usually evaluated significantly higher than in the group of aphasics. LHD patients obtained significantly lower results on appropriateness of discourse in stories, first script and second plan.

Table 6. Results on subjective ratings of discourse order

Task	LHD Mean (SD)	RHD Mean (SD)	C Mean (SD)	Ratio F(2.42)	Intergroup Differences
Story 1	13.93(3.10)	16.00(4.50)	18.13(1.99)	5.85*	C >LHD
Story 2	14.06(3.71)	17.13(1.76)	19.00(1.46)	14.66***	C>RHD>LHD
Script 1	14.33(2.99)	18.33(1.63)	19.06(1.48)	21.12***	C > LHD
Script 2	15.20(4.07)	18.60(1.20)	19.06(1.58)	8.80**	C,RHD>LHD
Plan 1	14.93(3.17)	18.00(1.60)	17.93(1.90)	8.48**	C,RHD>LHD
Plan 2	13.80(3.61)	17.60(1.76)	18.53(1.24)	15.99***	C,RHD>LHD
Summary	86.27(15.33)	105.87(9.83)	111.67(6.74)	21.15***	C,RHD>LHD

*p<0.05
**p<0.001
***p<0.0001

According to Table 6 on the subjective ratings of order, aphasics scored markedly lower than both remaining groups in all experimental tasks.

No significant differences were observed between non-fluent and fluent aphasics on either of the objective or the subjective discourse measures (Table 7).

Table 7. Comparisons between non-fluent and fluent aphasics on the objective and subjective discourse measures

Measure	Non-fluent Mean (SD)	Fluent Mean (SD)	T ratio	DF
Number of model units of information	26.29(5.82)	32.25(6.09)	-1.94	12
Index of Wrong Order	16.00(0.19)	0.31(0.24)	-1.36	12
Number of inappropriate contents	2.29(5.62)	5.00(4.90)	-0.99	12
Appropriateness – subjective measure	92.70(13.0)	93.38(6.80)	-0.12	8
Order – subjective measure	85.70(19.7)	86.70(11.6)	-0.12	9

Results of both aphasic subgroups were comparable when scores were analyzed separately on each discourse task as well as on all experimental tasks taken together.

Correlations between particular measurements

In order to examine the relationships between the various discourse measurements and between the discourse and language scores, correlations were carried out. Correlations between objective and subjective discourse measurements were not significant in most cases, particularly in the control subjects (Table 8).

However, in the experimental groups there were significant negative correlations between the number of inappropriate contents and both subjective ratings. There were also

Table 8. Correlations between objective and subjective discourse measures

Objective Measures	Subjective Measures	LHD	RHD	C
Number of model units of information	Appropriateness	0.29	0.53*	-0.08
Number of model units of information	Order	-0.04	0.52*	-0.03
IWO	Appropriateness	-0.08	0.05	-0.18
IWO	Order	0.23	-0.00	-0.08
Number of inappropriate contents	Appropriateness	-0.72**	-0.51*	0.14
Number of inappropriate contents	Order	-0.55**	-0.54*	0.06

*p<0.05
**p<0.01

strong positive correlations between the number of model units of information and the subjective estimates in the RHD patients.

Correlations between discourse ratings and scores on the BDAE in aphasics revealed very clear relationships (Table 9).

Table 9. Correlations between the total scores on the discourse measures and the total scores on the BDAE in Aphasics

Discourse measures	BDAE	Auditory comprehension	Oral expression
Number of model units of information	-0.06	0.39	-0.14
IWO	0.25	-0.02	0.31
Number of inappropriate contents	-0.41	-0.37	-0.40
Appropriateness – subjective measure	0.62**	0.75**	0.57*
Order – subjective measure	0.86**	0.58*	0.87**

*p<0.05
**p<0.01

There were no significant correlations between either of the objective discourse measures in relation to the scores on the BDAE. On the other hand, all subjective ratings of aphasics' discourse correlated positively with the results on the BDAE.

Discussion

The present experiment was designed to examine correctness of discourse content and order using objective and subjective methods in mildly to moderately impaired aphasics and RHD patients. Unexpectedly large discrepancies between the results on these measurements were found in both groups. The main point of this part of the paper will be to summarize and to attempt to interpret these contrasting results.

LHD group

The results of objective text measurements in the LHD group revealed that aphasics' discourse was comparable to that of normal controls. Although the LHD patients produced significantly fewer model units of information overall than the healthy subjects, the amount of model information in particular discourse tasks in aphasics was similar to that of the control group. The only exception was a markedly lower number of model units of information referring to abstract analysis of action produced by the LHD subjects in the second plan. Moreover, there were no significant differences between aphasic and control subjects in the ordering of model information and in the amount of inappropriate content.

The results of the objective discourse analyses are in accordance with the previous evidence indicating intact abilities of these patients to produce well-formed structure of discourse (Huber, 1990; also see Osiejuk, 1994, Łojek-Osiejuk, 1996). The method of discourse analysis used in all these studies was objective, that is, the experimenters measured the occurrence of specific contents and/or elements of text structure in utterances produced by aphasic patients.

A different picture of the aphasics' discourse emerged from the subjective ratings. Although, none of the aphasic utterances was judged completely inadequate or disordered, the texts of LHD patients were usually rated significantly lower on appropriateness and order of discourse than the texts of healthy persons. Thus, from the perspective of the subjective judges, aphasics' discourse was significantly less appropriate and less organized than that of healthy subjects. This finding is in conflict with the results of the objective discourse estimates for the LHD group.

There are very few studies on aphasics' discourse in which the method of subjective assessment was applied. One of these studies revealed results comparable with the present ones (Ulatowska et al., 1983a, b). Using the objective discourse measures, Ulatowska and co-authors demonstrated preservation of narrative and procedural superstructure in moderately impaired aphasics. On the other hand, the use of subjective rating of aphasics' discourse showed results significantly lower than those of normal controls.

As in the studies by Ulatowska and colleagues, a comparison between non-fluent and fluent aphasics did not demonstrate any evident differences in results on objective and subjective measures of discourse coherence. This may be due to a relatively low severity of language impairment in the aphasics and the fact that the patterns of reduction in discourse structure (that is, preservation of the most important elements and reducing optional information) are common in both subgroups of aphasics.

McDonald (1993) also used subjective and objective discourse ratings in the investigation of pragmatic language skills in two closed-head-injured non-aphasic subjects. McDonald obtained different results from those reported here. In his study the scores on objective propositional analysis corresponded with the results of text estimates performed by subjective judges. Lack of concordance between the present study and the research by McDonald may be due to the differences in the selection of subjects taking part in the experiments (aphasics versus nonaphasics) and methodological differences in usage of subjective methods of discourse analysis and measurements. A more detailed comparison of the two studies is carried out below.

Conspicuous contrasts in the findings from objective versus subjective discourse ratings in aphasics in the present study can be analyzed in terms of three main factors: deficits

at the discourse level, linguistic impairment, and methodological differences in discourse measurement and assessment.

Discourse deficits

The reduction of model units of information revealed by the objective analyses in the LHD group might be one of the variables that entailed a subjective judgement of aphasics' texts. Actually, in the three tasks where the aphasics obtained significantly lower results on appropriateness, they also produced relatively fewer model units of information. This result supports the data reported by Ulatowska and co-authors (1983 a). However, in contrast to their findings, there were no significant correlations between the scores on subjective ratings and the number of model units of information produced by the aphasic subjects. This evidence suggests that, although the raters assessed the amount of necessary information in a text, it was not the most important factor (or not the only one) influencing their estimates.

There is a possibility that the differences in the correlations might have been affected by differences in instructions given to the raters performing subjective evaluations. In the studies by Ulatowska and colleagues the raters were asked to assess three independent aspects of content pertaining to the main objects, actions and details of a discourse. Thus the raters took the preservation of particular contents of a discourse into account. On the other hand, in this research the estimate of text appropriateness was based on the answer to one question pertaining to the adequacy of text content as a whole. That kind of instruction did not require from the raters an analysis of occurrence of specific information.

There are also other discourse variables which could produce differences between results on objective versus subjective measurements. The materials for subjective estimates involved paralinguistic and some cohesion characteristics of discourse (eg. false starts) which were not present in the transcribed texts for objective analysis.

Since the main subject of the present study was correctness of discourse contents and order, other qualities of speech performance were not considered. However, according to the results of research on linguistic prosody (Baum et al., 1982; Heilman et al., 1984) as well as on coherence in aphasia, one could expect errors to occur on these dimensions of aphasic discourse. Consequently, these kinds of disorders should also be examined as potential factors that could cause differences in the subjective ratings of contents of aphasic discourse.

Linguistic impairment

The results of the study indicated strong associations between the results on subjective discourse ratings and the level of linguistic functioning in aphasics measured by the BDAE. Similar positive correlations were also reported in the earlier research (Ulatowska et al., 1983 a, b). According to the interpretation by these authors the correlations suggested that performance on a standardized language test could predict ability to produce connected language in the form of a narration and a procedure.

While agreeing with this conclusion, it also seems important to take into account relationships between the linguistic performance of aphasics and the estimates of their discourse content by subjective receivers. The results of these studies also suggested that the performance of aphasics on a standardized language test could paradoxically be a predictor of estimates of aphasics' discourse performed by subjective receivers. Moreo-

ver, these findings also seem to suggest that the occurrence of linguistic impairment may be one of the important variables that could influence perception of aphasics' discourse. Namely, in the subjective ratings of a healthy listener, language errors produced by aphasics might decrease correctness of discourse content and order, even if, objectively, aphasics' texts were comparable to that of normals.

It is also important to notice that in the study by Ulatowska and colleagues the raters were speech pathologists. They had probably more experience in comprehension of aphasics' discourse than the clinical psychologists taking part in the present research. However, the level of agreement between the two studies makes this eventuality unconvincing.

If the influence of linguistic deficits on perception of aphasics' discourse really exists, the relationships between results on subjective ratings and on non-standardized linguistic measures should also be observed. The studies by Ulatowska and others demonstrated that the scores on some of the selected linguistic measures correlated significantly with the subjective ratings of content of aphasics' discourse, but some of them did not. A broader selection of linguistic measures and deeper analysis of linguistic characteristics of aphasics' language for perception of their discourse are needed for further examination of this problem.

The phenomenon of possible influence of language disorders on subjective perception of aphasics' discourse in natural receivers can also be interpreted in terms of influence of context on an analyzed variable. The main problem of the prior discourse research was dependence of aphasics' speech–language performance on the influence of contextual variables, eg. non-verbal clues or redundancy of information in a text (Huber, 1990). In this study, correctness of discourse content and order were the key variables, whereas language performance became one of the contextual components (in the material for subjective analysis). The data obtained demonstrated once again the power of contextual influences. However, this time perception of the discourse content became a variable changing with regard to the contextual (linguistic) influences.

It is also interesting to note that both discourse measures, the objective and subjective, are considered as related to pragmatics. However, from the perspective of natural communication, the subjective ratings might be called "more pragmatic", i.e. more representative than the objective evaluations for the ways in which discourse is perceived in natural circumstances. Paradoxically, in contrast to the results of objective measures, the results of use of the "more pragmatic" method of discourse evaluations were very much related to the performance at the word/sentence level of language – usually separated from pragmatics. These observations suggest that the theoretical models of discourse should not only demonstrate the hierarchy of discourse representations, but also the interactions between various components of a text, including purely linguistic aspects.

The findings suggesting that language errors of aphasics render the proper perception of their discourse by normal receivers more difficult may be of potential importance for aphasia therapy. The limitations in perception of aphasics' discourse by other people might be an additional variable responsible for the "clinic-nature gap" described by Sarno (1969). In that case, to improve communication of aphasic persons, therapy should involve members of families and other significant persons, as it is postulated and practized in pragmatic oriented rehabilitation (Davis & Wilcox, 1985; Lesser & Milroy, 1993). However, significant persons should not only be trained in how to deal with language problems of aphasics, but also how to cope with their own limitations in perceiving aphasic discourse.

Methodological reasons

There were differences between the two methods of discourse analysis concerning ways of text perception and assessment. Examining written texts might be easier than listening to recorded data. In the objective type of analysis, in contrast to the subjective one, the receiver could control the speed of perceived discourse, could easily come back to the previous parts of a text, and was not affected by possible technical difficulties of tape recordings. High correspondence between the results of objective and subjective discourse measures in McDonald's study (1993) might be partially caused by the fact that the discourse material in the case of both measures was the same (transcribed data) and the perception of it was similar (reading of texts).

In the present study, as well as in the experiments by Ulatowska and colleagues (1983a, b) raters listened to tape-recorded productions. Nevertheless, one could expect that technical limitations of the subjective method in this research should not produce such marked differences between results in the LHD group since they were present during estimates of the texts of all the subjects, not only aphasics. However, co-existence of various technical restrictions and language deficits of the aphasic patients might render perception of the content of aphasics' discourse more difficult than that of the other subjects.

The types of discourse analysis and assessment were also different in both ratings. The objective method was analytical, concentrating on selected, precisely defined, elements of the text structure. The second type of estimates relied on subjective, more holistic, assessment of discourse. In that sense variables chosen for objective analysis were not completely compatible with those estimated by subjective raters. Difference between these two methods was clearly demonstrated by the lack of correlation between the results on objective and subjective discourse measures in the control group. Validity of both assessments would probably be higher if the definitions of text variables for objective discourse analysis were related to more specific discourse characteristics.

The results reported by Ulatowska and co-authors (1983a, b) and McDonald (1993) confirm the above assumption. In their studies, subjective ratings included judgement of separate text features. Thus, in the research by Ulatowska and others (1983a) the results on quality of discourse content (involving the assessment of occurrence of main objects, clarity of action, amount of detail) correlated positively with the total number of steps as well as the number of essential steps and target. According to McDonald (1993), the scores on the number of new and repeated propositions were related to the results of judgements of detail and repetitiveness, and the amount of errors in sequencing and inclusion of irrelevant proportions contributed to the ratings of text organization and clarity.

In contrast to prior investigations, characteristics of discourse for subjective ratings in the present research were less distinguished. The subjective rating of appropriateness was estimated in response to one question referring to the general adequacy of text content. This was probably one of the reasons why the number of model units of information did not correlate with the scores on appropriateness in the LHD and C groups. Under the subjective estimate the adequacy of a text may be acceptable, even if the text does not contain all model information.

RHD group

The results on objective and subjective discourse measures were also not entirely concordant in the second experimental group. However, the gap between the two kinds of results in the RHD patients was not as big as for the aphasics. Findings from both discourse ratings revealed

that the RHD subjects were generally able to produce adequate, relatively well organized discourse. They provided a number and chronology of model units of information comparable to that of normal controls on the majority of discourse tasks. On subjective ratings their utterances were estimated as similar to those of healthy people in almost all tasks.

Despite these generally preserved cognitive abilities, objective methods of analysis clearly revealed deficits of the RHD patients in performing discourse. The total scores on all discourse tasks showed that the RHD group produced significantly less model information, a higher number of order errors and more inappropriate content in comparison with the control group. The most conspicuous impairment was on the tasks of recalling content of stories and on the ordering plans. These results are in agreement with studies showing impaired abilities to perform discourse in RHD subjects (Joanette & Brownell, 1990).

The results of subjective discourse ratings did not indicate such conspicuous discourse disorders in the RHD group. The only exception was the task of recalling the second story. Texts provided by the RHD subjects in that task were judged to be significantly impaired on both discourse characteristics measured subjectively. These results corresponded with that on the objective discourse measure, revealing a markedly high number of inappropriate propositions in the RHD group. There were no other differences between the RHD and control subjects which could be revealed by either discourse measurement.

Regardless of the lack of consistency in demonstrating intergroup differences, the other results showed associations between both kinds of measurements. The scores of both subjective measurements correlated positively with the number of model units of information. Significant negative correlations were also obtained between the ratings on the subjective assessment and the number of inappropriate commentaries. These correlations indicated that, estimating the discourse of the RHD subjects, the raters had a tendency to give higher ratings for the utterances with a higher number of model information units. They also tended to reduce the points for the texts with a higher amount of inappropriate content. The occurrence of the above tendencies suggests that the raters actually noticed deficits on the adequacy of information in the RHD patients' texts. However, from the point of view of the subjective raters, these deficits were acceptable and did not exceed the standards of normal discourse production.

Neither did the raters perceive any difficulties in the ordering of information in the RHD subjects, even though the results on the total IWO score and on the ordering of both plans showed significant deficits. This may be partially related to the fact that the evaluation of order was not completely separated from the ratings of adequacy and that the order errors produced by the RHD patients were relatively rare in a particular utterance. However, in the case of plans described by the RHD subjects (which were correct in terms of the adequacy of propositions and the only disturbance was the ordering), the raters did not exhibit any significant deficits in the chronology of information. Interestingly, the order of information produced by the healthy subjects in both plans also contained differences in comparison with the model order. Nevertheless, neither the plans of the RHD patients nor the plans of the healthy people were estimated as disturbed by the subjective raters. It seems that, from the perspective of the subjective receivers, the changes in ordering of information made by the RHD and C subjects in both plans were not noticed as errors. This tolerance in perceiving the sequence of information produced in spontaneous speech shown by the subjective raters might be another reason why the subjective ratings, in contrast to the objective measures, did not reveal difficulties of the RHD patients.

Lack of explicit reports about discourse disturbances in the RHD patients from clinicians and relatives is probably due to the fact that in natural circumstances inadequacy of RHD subjects' discourse can be compensated for by receivers and not perceived as exceeding standards of normal communication (see Tonkonovich, 1989).

The results of objective and subjective measures in the RHD group evoke once again the problem of flexibility of both methods. The findings indicate that higher pliancy and tolerance in comparison with objective discourse measures characterize the subjective estimates of text coherence. In that sense, certain deficits on adequacy and order of discourse in the RHD patients shown with the use of objective measurements are not easily decoded by the method of subjective rating. Paradoxically, the subjective discourse estimates turned out to be much more flexible in the case of the RHD patients than in that of the aphasics.

Thus, the perception of discourse content disturbed by purely linguistic deficits seemed to be more difficult than perception of texts disturbed (to some extent) by inadequate information and errors of order. This finding may suggest that there is a specific hierarchy of importance of text components, which influence the perception of the content of discourse, and the role of the linguistic aspect of an utterance in that hierarchy may be more important than is suspected.

Conclusions

To summarize, the results of this study revealed that two methods of analysis of discourse, objective and subjective, are not fully compatible. Aphasics' discourse was estimated as less appropriate and less organized by subjective raters that it appeared from objective analysis. On the other hand, subjective ratings in the RHD group did not indicate those discourse deficits that were shown by the use of objective methods.

On the basis of the present experiment it is impossible to draw final conclusions about factors which accounted for the results. Significant correlations between subjective ratings and the scores on the BDAE suggested that language problems of aphasics could render more difficult the proper perception of their discourse by a healthy listener. However, other variables such as: discourse errors and methodological differences between both discourse ratings should also be taken into account. The results in the RHD group also suggested that a certain degree of tolerance for changes in discourse content and order in subjective raters may influence text estimates.

More studies are needed to confirm the above findings and tentative conclusions. Both ratings should be more carefully examined to establish various psycholinguistic factors responsible for the observed differences in discourse measurement. Surprisingly, there are a lot of experiments on discourse production and comprehension in aphasic and RHD patients but very few on perception of their discourse by normal receivers.

References

Baum, S. R., Daniloff, J. K., Daniloff, R., & Lewis, J. (1982). Sentence comprehension by Brocka's aphasics effects of some suprasegmental variables. *Brain and Language, 17,* 261-271.

Davis, G. A., & Wilcox, M. J. (1985). *Adult aphasia rehabilitation. Applied pragmatics.* San Diego: College-Hill Press.

Frederiksen, C. H., Bracewell, R. J., Breuleux, A., & Renaud, A. (1990). The cognitive presentation and processing of discourse: function and dysfunction. In Y. Joanette, & H. H. Brownell (Eds.), *Discourse ability and brain damage: Theoretical and empirical perspectives*. New York: Spinger-Verlag.

Heilman, K. M., Bowers, D., Speedie, L., & Coslett, H. B. (1984). Comprehension of affective and nonaffective prosody. *Neurology, 34,* 917-921.

Huber, W. 1990. Text comprehension and production in aphasia: analysis in terms of micro- and macroprocessing. In Y. Joanette, & H. H. Brownell (Eds.), *Discourse ability and brain damage. Theoretical and empirical perspectives.* New York: Springer-Verlag.

Joanette, Y., & Brownell, H. H. (Eds.)(1990). *Discourse ability and brain damage. Theoretical and empirical Perspectives.* New York: Springer-Verlag.

Lesser, R., & Milroy, L. (1993). *Linguistics and aphasia. Psycholinguistic and pragmatic aspects of intervention*. London: Longman.

Łojek-Osiejuk, E. (1996). Knowledge of scripts reflected in discourse of aphasics and right-brain-damaged patients. *Brain and Language*, 53, 58-80.

McDonald, S. (1993). Pragmatic language skills after closed head injury: ability to meet the informational needs of the listener. *Brain and Language, 44,* 28-46.

Osiejuk, E. (1994). *Problematyka dyskursu w neuropsychologii poznawczej* (Problems of discourse in cognitive neuropsychology) Warszawa: Oficyna Wydawnicza Wydziału Psychologii UW.

Sarno, M. (1969). *The functional communication profile. Manual of directions.* New York: New York University Medical Center – The Institute of Rehabilitation Medicine.

Tonkonovich, J. D. (1989). Managing pragmatic communication deficit associated with right hemisphere damage. *Seminars in speech and language, 10,* 343-354.

Ulatowska, H. K., Freedman-Stern, R., Weiss Doyel, A., Macaluso-Haynes, & S., North, A. J. (1983 a). Production of procedural discourse in aphasia. *Brain and Language, 18*, 315-341.

Ulatowska, H. K., Freedman-Stern, R., Weiss Doyel, A., Macaluso-Haynes, & S., North, A. J. (1983 b). Production of narrative discourse in aphasia. *Brain and Language, 19,* 317-334.

Appendix 1. Sets of model units of information for stories, scripts, and plans

The first story – "Robbery"	**Core information**
1. It was night.	(night)
2. There was a man going down the street.	(man, go down, street)
3. He stopped behind the doors of a shop.	(he, stop, shop)
4. He broke the lock.	(he, break, lock)
5. He entered the shop.	(he, enter, shop)
6. He approached the cash-box.	(he, approach, cash-box)
7. He broke into the cash-box.	(he, break, cash-box)
8. He put all the money into a sack.	(he, put, money, sack)
9. He put the sack under his coat.	(he, put, sack, coat)
10. He ran from the shop.	(he, run, shop)
11. He disappeared into the darkness.	(he, disappear)

The second story – "Mr Kowalski"	**Core information**
1. Mr Kowalski was coming home from work.	(come, home, Mr. Kowalski)
2. He bought several newspapers.	(he, buy, newspapers)
3. He arrived home.	(he, arrive, home)
4. He put on his home clothes.	(he, put, clothes)
5 He had his dinner.	(he, have, dinner)
6. He started to read the newspapers.	(he, read, newspapers)
7. He quickly flicked through the first magazine.	(he, flick through, magazine)
8. He took the second magazine.	(he, take, second)
9. He read the headlines on the first page.	(he, read, headlines)
10. He finally found the article he was looking for.	(he, find, article)
11. He became engrossed in it.	(he, become, engrossed)

The first script – "Borrowing money"

1. Prerequisite activating the script – general labelling of the situation (eg. "When you do not have money and you have to borrow it you should...").
2. Choosing the person who would lend the money.

3. Contacting the person.
4. Describing the situation to the person.
5. Asking to borrow money.
6. Discussing the amount (sum) of the loan.
7. Establishing conditions for returning the money (percentage, reimbursement procedure).
8. Establishing a time limit for repayment.
9. Borrowing or not borrowing the money.
10. Returning the money.
11. Other options if unsuccessful.

The second script – "Asking the way in an unfamiliar town"

1. Prerequisite activating the script – general labelling of the situation (eg. "When you do not know the way in a town you can...").
2. Considering various methods of action in the situation, different from asking the way.
3. Asking somebody, as a general way of acting.
4. Establishing the purpose (needed localization).
5. Choosing a person who might know the way.
6. Asking the chosen person.
7. Getting or not getting information.
8. Going in the shown direction or asking for further directions.
9. Reaching the destination.

The first plan – "Getting money"

Examples of possible content of elements:
1. AE – Considering the amount of money.
2. AE – Considering how quickly money is needed.
3. AE – General considerations on the possible methods of an action.
4. AE – Considering the possible consequences of an action.
5. AE – Making the decision to choose a particular method.
6. AE – Making the decision to start an action or not.
7-12. CE – Concrete methods of acting.

The second plan – "Getting information"

Examples of possible content of elements:
1. AE – Considering the kind of information needed.
2. AE – Considering the possible sources of the information.
3. AE – Considering the possible consequences of an action.
4. AE – Making the decision to choose a particular method.
5. AE – Making the decision to start an action or not.
6-11. CE – Concrete methods of acting.

AE – an abstract element
CE – a concrete element