

BARBARA JACENNIK
University of Warsaw

MEASURING CHILDREN'S ASSERTIVENESS FROM CONVERSATIONAL SAMPLES*

A communicative behavior protocol for assessing children's conversational assertiveness was designed based on discourse analytical concepts and tested to assess its psychometric characteristics. The protocol is to be used for scoring data collected during tape-recorded conversations. The aspects of validation work described here concern in particular the issues of inter-rater reliability and internal consistency. The test of the protocol was conducted on data from 21 child-adult conversations. The results show that the proposed protocol has good psychometric qualities and followed further may become a promising tool for assessment.

Assertiveness is a desirable behavioral trait closely related to communicative competence (Richmond and McCroskey, 1985). Most assertiveness measures comprise items assessing assertiveness across communicative situations, and, to a lesser extent, assertiveness as a conversational skill. Also, most measures are designed for adult populations, and cannot be used to monitor the development of assertiveness skills in children. The study reported here is part of a project to develop an assessment protocol measuring conversational assertiveness in children.

The first test of the protocol was conducted on data from 10 Polish and 10 Canadian children rated by two trained raters and was reported elsewhere (Jacennik, 1999). In this portion of the project, judgments made by two judges on data from two linguistic groups of children were consistent – inter-rater reliability was statistically significant (Kendall tau = 0.489, $p < 0.003$); quantitative measures of discourse events, such as repairs and conversational avoidance, were also calculated. Some significant findings pointing to cross-gender rather than cross-linguistic differences were obtained. Boys were more likely to use repairs and explicitly refuse answers to some questions. However, due to small sample sizes and

* This paper was first presented as a conference paper "Measuring children's assertiveness from conversational samples", Twenty-Eighth LACUS Forum, Montreal, Canada, July/August 2001. Address for correspondence: Barbara Jacennik, University of Warsaw, Faculty of Psychology, Stawki 5/7, 00-183 Warszawa, Poland. E-mail: jacennik@psych.uw.edu.pl

difficulties with the experimental control of cross-linguistic measures, these results had to be interpreted with caution.

A similar procedure was used in the current project. Again, judges were asked to judge tape-recorded conversations with children using the assessment protocol. This test of the protocol was performed on data from Polish children only in order to avoid difficulties arising in cross-linguistic comparisons. Statistical analyses conducted on the data examined inter-rater reliability and internal consistency of the protocol.

Method and procedure

Assertiveness as a psychometric concept

According to Richmond and McCroskey (1985), assertiveness “is the capacity to make requests; to actively disagree; to express positive or negative personal rights and feelings; to initiate, maintain, or disengage from conversations; and to stand up for oneself without attacking another.” Assertiveness is differentiated from aggressiveness in communication, the main difference being respecting the rights of others to communicate. Both assertiveness and aggressiveness are negatively correlated with communication apprehension when measured using psychometric instruments. Assertiveness is also differentiated by some authors from responsiveness. Fey (1986) defines a child’s conversational assertiveness as an ability or willingness (or both) to take a conversational turn when none has been solicited by a partner.

Fey differentiates assertiveness from responsiveness defined as the capacity to be sensitive to the communication needs of others, to be seen as a good listener. In Fey’s conception, the two dimensions intersect thus yielding four general types of children differing on one or both dimensions.

Rationale and background for the new protocol

Psychometric instruments measuring assertiveness usually refer to generalized life situations such as requesting help from a stranger or defending oneself in a verbal argument. These measures only indirectly refer to assertiveness as a conversational skill – for example Rathus (1973). There is a lack of assertiveness measures referring specifically to assertiveness in conversations despite the fact that speech situations and conversations, in particular, constitute a context in which a person’s assertiveness is especially apparent. In addition, the existing assertiveness measures are usually based on self-report and are not practical for young children.

Research from which the current work stems follows three different lines: psychological research on assertiveness, e.g. Rathus (1973); research on speech communication in adults – e.g., Richmond and McCroskey (1985); and speech pathology research on the application of linguistically-informed discourse measures to child language (e.g., Fey, 1986; Cole, Dale & Thal, 1996). Especially

useful for developing the statements of the protocol were the conversational analysis concepts developed by Brinton & Fujiki (1989).

Rating procedure

The self-report approach used in communication assessment tools for adults was not employed as it is not appropriate for young children. Instead, an observer-rating method was chosen. A set of rating scales was developed for use by trained adults who were asked to judge tape-recordings of children's conversations (Appendix 1).

Conversational samples, subjects and adult partners

The conversations used for the purpose of assessment were one-on-one and followed a semi-structured script. They were tape-recorded and transcribed (the transcriptions were not used in this phase of the project). The adult conversational partners were a psychologist and a teacher. The raters were two psychology students and a psychologist, all native speakers of Polish. The conversation script contained 8 topics which the adult partner could introduce into the conversation. These were topics likely to interest a child: food, pets, friends, movies. Any topics introduced by children were reinforced by the adult partners. The adult partner was instructed to use a conversational style that would encourage child participation – speak in a friendly, non-evaluative, peer-like, non-directive manner. A typical conversation lasted 10 minutes. The children with whom the conversations were tape-recorded were 21 primary school students from one Warsaw school. Age range was 7 to 11 years old.

Protocol design and instructions to raters

The statements in the protocol were descriptions of assertive behavior according to the definition of assertiveness. In particular, the items describe active, nonaggressive, balanced conversational behavior. Literature on child discourse analysis, particularly Brinton and Fujiki (1989), suggests a conceptual breakdown of discourse abilities into three sub-components: turn-taking, topic management and conversational repair. These concepts were used to formulate statements for the scale. In addition, some general items referring to the degree and quality of conversational participation were included. The protocol consists of a global assessment scale defined to the raters orally in terms of assertiveness, and 15 specific items which are to be used as the basis of the global assessment. Each rater was asked to make a global judgment of assertiveness and to justify it by making detailed judgments on items referring to specific conversational behaviors correlating with assertiveness. In addition, a few items referring to general aspects of interaction associated with assertiveness were included among the 15 subscales.

The term "assertiveness" was used only in the oral instructions to raters and not mentioned in the written version. The global scale on which the raters were

asked to judge the child's participation referred to "active participation in conversation" which was contrasted further with "being passive or withdrawing from conversation". After the general definition in the main scale, the raters were given an operational definition of assertiveness expressed in 15 statements. This was intended to prevent the raters from relying on their own idiosyncratic ideas about assertiveness. As they were asked to focus on individual assertive behaviors, it was hoped that they would not treat the task as an assessment of a child's personality in terms of "assertiveness". The scale was modified relative to the previous study (Jacennik, 1999) by removing unclear items and increasing the number of response options from 2 to 5. The raters were given instructions on how to fill out the form and asked to listen twice to each tape. They were asked to mark the form after the first listening and then listen for the second time to evaluate items they were not sure of. None of the raters had been familiar with the children whose conversations they were evaluating.

The results

Conversion of reverse items

Two items contained statements which should be answered negatively for an assertive child and positively for a nonassertive one. For the purposes of statistical analysis the scores on these two items had to be reversed. These were item K – "refused to answer some questions", and item O – "sometimes did not complete his/her utterances". Item J "sometimes expressed differing opinions" was not reversed as its correlation with the total score was negligible and will likely be changed or left out of the protocol in the future.

Inter-rater reliability

For the purpose of this analysis the individual scores were converted into ranks given by judges compared to the rest of the group for each item. The ranks were used as the scores are measured on ordinal scales, not interval or ratio – hence nonparametric statistics had to be used for tests of significance. Total scores were calculated as the sum of ranks for different items. These data were submitted to the Kendall W test of significance. Inter-rater reliability for the total score was relatively high (Kendall W = 0.701; Table 1).

Table 1. Inter-rater consistency Kendall W for the total scale

N	3
Kendall W	0.701
Chi-square	42.036
df	20
Asymptotic significance	0.003

Table 2. Inter-rater consistency on individual items

Items	Inter-rater reliability Kendall W	asympt. significance
X. The child was an active participant in conversation	0.735	0.001*
A. elaborated the topic of conversation	0.705	0.003*
B. responded willingly to all questions	0.613	0.012*
C. initiated new topics	0.754	0.001*
D. sometimes completed the interlocutor's utterances	0.411	0.215
E. when interrupted resumed the topic of his/her utterance	0.593	0.017*
F. asked for confirmation or clarification when failed to understand something	0.486	0.085+
G. clarified some misunderstandings	0.300	0.586
H. interrupted the interlocutor when wanting to say something	0.452	0.133
I. corrected some of the interlocutor's utterances	0.346	0.413
J. sometimes expressed differing opinions	0.309	0.551
K. refused to answer some questions (reversed)	0.562	0.028*
L. was willing to participate in conversation	0.745	0.001*
M. sometimes cut into the interlocutor's utterances	0.413	0.211
N. enjoyed the conversation	0.786	0.001*
O. sometimes did not complete his/her utterances (reversed)	0.473	0.100+

* $p < 0.05$; + $p < 0.10$

Inter-rater reliability for each of 16 items (1 global and 15 specific) was also calculated. On 8 items the raters were consistent at a statistically significant level (Table 2). On two items the ratings showed a tendency towards statistical significance. On the remaining 6 items the raters were not significantly consistent with each other – which indicates that they had some difficulties with making judgments. These items may have to be changed in the future. Raters are known to vary so these results should not come as a surprise.

Internal consistency of the protocol

The second methodological question asked in this study addressed the issue of internal consistency of the protocol. The statistics used for this purpose was the Cronbach internal consistency alpha coefficient. This is a statistical measure often used to evaluate psychometric tests: it reflects the consistency of scores on individual items with the total score. The alpha coefficient was calculated on raw scores – ratings given by judges on a five point scale for each of 16 items. (The scores were not recoded to ranks within subjects, as such a situation would not be

Table 3. Correlations between the items and the total score on the protocol

Items	Corrected item-total correlation	Type of item
X. The child was an active participant in conversation	0.9577	G
A. elaborated the topic of conversation	0.8872	TM
B. responded willingly to all questions	0.8727	TT
C. initiated new topics	0.8663	TM
D. sometimes completed the interlocutor's utterances	0.7509	RP
E. when interrupted resumed the topic of his/her utterance	0.7248	TM
F. asked for confirmation or clarification when failed to understand something	0.7208	RP
G. clarified some misunderstandings	0.3826	RP
H. interrupted the interlocutor when wanting to say something	0.6995	TT
I. corrected some of the interlocutor's utterances	0.5639	RP
J. sometimes expressed differing opinions	-0.2247	TM
K. refused to answer some questions	0.7000	TT
L. was willing to participate in conversation	0.9433	G
M. sometimes cut into the interlocutor's utterances	0.7677	TT
N. enjoyed the conversation	0.8430	G
O. sometimes did not complete his/her utterances	0.7644	RP

consistent with a typical scoring situation – a psychometrist normally does not rank subjects). As the individual variation between raters is known to be usually greater than the intra-rater variability, the judgments from the three raters were averaged. The alpha of 0.9485 (standardized item alpha = 0.9426) indicated that the internal consistency of the protocol was high.

As shown in Table 3 the global activeness rating (item X) was the basic dimension used by all judges as intended by the protocol design. This item correlated most highly with the total score at a level of 0.9577 and therefore explained the greatest percentage of the total variance.

Alpha correlations by types of item were also looked at. The items in the protocol were arbitrarily sorted into 4 categories according to the kinds of conversational behaviors they referred to. The item-total correlations for each category are shown in Table 4: general participation (G), turn taking (TT), repair (RP),

Table 4. Item-total correlations averaged by category of items

Item category	Number of items	Item-total correlation
General participation	3	0.904
Repair	5	0.799
Topic management	4	0.905
Turn taking	4	0.876

Table 5. Correlations between item X and the remaining items

Items	Correlation with item X (Kendall tau b)	Asympt significance
X The child was an active participant in conversation	1.000	
A. elaborated the topic of conversation	0.872	0.000**
B. responded willingly to all questions	0.875	0.000**
C. initiated new topics	0.741	0.000**
D. sometimes completed the interlocutor's utterances	0.626	0.000**
E. when interrupted resumed the topic of his/her utterance	0.550	0.002**
F. asked for confirmation or clarification when failed to understand something	0.619	0.000**
G. clarified some misunderstandings	0.215	0.227
H. interrupted the interlocutor when wanting to say something	0.663	0.000**
I. corrected some of the interlocutor's utterances	0.435	0.013*
J. sometimes expressed differing opinions	-0.200	0.277
K. refused to answer some questions	0.604	0.001**
L. was willing to participate in conversation	0.898	0.000**
M. sometimes cut into the interlocutor's utterances	0.705	0.000**
N. enjoyed the conversation	0.791	0.000**
O. sometimes did not complete his/her utterances	0.650	0.000**

** $p < 0.01$ * $p < 0.05$

topic management (TM). The best item-total correlations were found for the items rating general participation in conversation.

Some items were found to correlate very poorly with the total score. These were: item J "sometimes expressed differing opinions" and item G "clarified some misunderstandings". Obviously, these statements were unclear for the judges; they were unsure whether these items reflected positive or negative conversational behavior – these items will have to be reworded in the next version of the protocol.

Correlations among items

As a final portion of the statistical analysis, tests of correlations on pairs of items using Kendall tau were also performed in order to check for similarity among items. The global item X correlated significantly with the majority of other items with the exception of item J "sometimes expressed differing opinions" and item G "clarified some misunderstandings" (Table 5). Only a few pairs of items produced higher correlations than their correlations with the scale X. The majority of items correlated significantly with almost every other item with the exception of three. The same items which have shown poor correlations with the general scale X, i.e.,

J “sometimes expressed differing opinions” and G “clarified some misunderstandings”, performed poorly on this analysis, consistently failing to correlate with other items. In addition, item I “corrected some of the interlocutor’s utterances” failed to correlate significantly with two other items (C “initiated new topics”, H “interrupted the interlocutor when wanting to say something”). These items will have to be dropped out or modified in the next version of the protocol.

Discussion

The purpose of this project was to test a working version of the conversational assertiveness protocol in order to verify its inter-rater reliability and internal consistency. The following psychometric properties have been established. The reliability analysis has shown that the raters generally agree with each other although consistently have problems with some items on the scale. The internal consistency analysis has shown that the majority of items were consistent, correlated with each other – and hence with the total score, with the exception of three items. On the whole, the reliability and consistency of scores demonstrate that the protocol, after further modifications, is likely to become a useful tool in clinical or educational settings. The work on assessing the validity of the protocol (whether it tests what it is designed to test) and its reliability (across sampling occasions, conversational partners and raters) will be continued in further studies.

The type of psychometric approach used in the protocol is likely to come under criticism from those stressing the fact that communicative performance is likely to change from moment to moment and one occasion to another. The following arguments can be brought in defense of this approach. The samples taken from children were approximately 10 minutes long. This constitutes a stretch of conversation which for a communicatively skilled child of this age contains a large number of communicative events (approximately 6 pages of transcripts). Conversation samples of 5-10 minutes have been shown to have fairly stable properties on various measures of discourse when compared with longer samples (Boles & Bombard, 1998). The conditions under which the conversations were conducted were very similar: school setting, separate room, same conversational script, conversation with an adult equally unfamiliar to all children. Most children in such a situation should show their skills in handling a mildly demanding social situation. Those who could not master anxiety or negative feelings are likely to fail in other similar social situations. The two persons talking to the children were instructed to behave in similar way, use the same conversational style. Of course, we cannot rule out personality differences, yet the interference of a tester’s personality holds for any testing situation. If the protocol were to be standardized in the future, the instructions on conversational style which is to be used with children would have to be defined precisely and the adult speakers could be trained as well as pre-tested. To be sure, however, that the assessment was not just the effect of a child

having a bad day, or a case of an interaction style mismatch between the child and the adult, it would be advisable to use two or more samples of conversations with different adult partners for each child.

Conclusion

Conversational assertiveness appears to be a particularly significant trait influencing how a child develops other social and intellectual abilities. The present study demonstrates an application of the conversational analysis to children's communicative assessment focusing on assertiveness. The current results of the protocol validation project show that the proposed tool has good psychometric qualities. In the next step, it is planned to relate the scale judgments to quantitative data extracted from the transcripts. It would be interesting to check whether the items on which judges are most consistent correlate well with the quantitative measures of linguistic events. The relationship between specific linguistic events and the perception of assertiveness is an area of inquiry interesting both from the communicative and developmental perspective.

Acknowledgements

The research reported here was supported by the grant from the State Committee for Scientific Research to the Warsaw University, Faculty of Psychology, grant No. 1462/5.

References

- Boles, L. & Bombard, T. (1998). Conversational discourse analysis: appropriate and useful sample sizes. *Aphasiology*, 12 (7-8), 547-560.
- Brinton, B. & Fujiki, M. (1989). *Conversational management with language-impaired children*. Rockville, Maryland: Aspen.
- Cole, K.N., Dale, P.S., & Thal, D.J. (Eds.) (1996). *Assessment of communication and language*. Baltimore: Brookes.
- Fey, M.E. (1986). *Language intervention with young children*. San Diego: College-Hill.
- Jacennik, B. (1999). Assertiveness markers in conversational speech. Paper presented at Pragma'99. International Conference on Pragmatics and Negotiation, Tel Aviv, Israel. June, 1999.
- Jacennik, B. (2001). Asertywność w konwersacji jako komponent sprawności komunikacyjnej dziecka (Assertiveness in conversation as a component of children's communicative abilities). *Kwartalnik Pedagogiczny*, 181-182, 163-176.
- Rathus, S.A. (1973). A 30-item schedule for assessing assertive behavior. *Behavior Therapy*, 4, 398-406.

- Richmond, V.P. & McCroskey, J.C. (1985). *Communication: apprehension, avoidance, and effectiveness*. Scottsdale, Arizona: Gorsuch Scarisbrick.
- Scherer, K.R. (1979). Personality markers in speech. In K.R. Scherer & H. Giles (Eds.), *Social markers in speech*. Cambridge: Cambridge University Press.

Appendix 1

- X. The child was an active participant in conversation
(as differentiated from being passive or withdrawing from conversation)
strongly disagree 1, disagree 2, can't decide 3, agree 4, strongly agree 5

*What was the basis of your assessment of the child's activeness?
Please evaluate the child's conversational behavior using the statements given below.*

- A. elaborated the topic of conversation (response options as above)
 - B. responded willingly to all questions
 - C. initiated new topics
 - D. sometimes completed the interlocutor's utterances
 - E. when interrupted resumed the topic of his/her utterance
 - F. asked for confirmation or clarification when failed to understand something
 - G. clarified some misunderstandings
 - H. interrupted the interlocutor when wanting to say something
 - I. corrected some of the interlocutor's utterances
 - J. sometimes expressed differing opinions
 - K. refused to answer some questions
 - L. was willing to participate in conversation
 - M. sometimes cut into the interlocutor's utterances
 - N. enjoyed the conversation
 - O. sometimes did not complete his/her utterances
-