Journal of Childhood Communication Disorders
1986 Vol. 9 No. 2 125-132

The Quality of a Language Sample as a Function of Its Placement in an Assessment Battery

SUE TONN, M.A. and ANNE VAN KLEECK, PH.D.
The University of Texas at Austin

Language evaluations in a clinical setting typically involve both standardized testing, in which a child is often required to merely point to correct responses, and the collection of a spontaneous language sample that is later subjected to various analyses. The present study sought to determine if beginning an evaluation with standardized tests that require only a pointing response might not predispose a child to talk less and use less complex language, since such tests require little or no verbal response. The results indicated that the placement of the elicitation of the spontaneous language sample in the initial language evaluation did not influence the length, complexity or spontaneity of the samples obtained from the children studied. These findings suggest that the clinician can exercise flexibility in choosing the order of presentation of diagnostic procedures in an initial language assessment.

The spontaneous language sample is the most valuable tool available to the speech/language pathologist for assessing a child's language and subsequently developing goals for intervention. As such, it is desirable to obtain a sample which represents a child's optimal linguistic performance. Past research has focused on the influence of numerous variables on length and complexity indices obtained from a language sample. Factors studied have included who elicits the sample (e.g., Cazden, 1970; Longhurst & Grubb, 1974; Olswang & Carpenter, 1978), the interaction style of the elicitor (e.g., Musselwhite & Barrie-Blackley, 1980), where the sample is elicited (e.g., Scott & Taylor, 1978; Kramer, James, & Saxman, 1979), the materials used to elicit the sample (e.g., James & Button, 1978; Longhurst & File, 1977; Longhurst & Grubb, 1974), and the length of the sample (e.g., Johnson & Tomblin, 1975). One variable possibly influencing the sample which has not been investigated is the position of the collection of the language sample in the evaluation, i.e., whether it is elicited before or after more formal testing procedures. The present study investigated the effect of the placement of the language sample collection in a language evaluation on the length, complexity, and spontaneity of the language produced by a young child, and the amount of adult talk produced in obtaining that sample.

The following diagnostic procedures exemplify those routinely given as part of an initial evaluation: (1) hearing screening, (2) receptive language test, (3) oral mechanism test, (4) articulation test and (5) spontaneous language sample. While there is general agreement regarding the content of such an evaluation, the order in which the procedures are carried out varies considerably. The collection of the spontaneous language sample may occur at any point during the evaluation, depending upon the preference of the clinician. Yet, the actual order of these diagnostic procedures may have an effect on the results of the spontaneous language sample obtained.

Of the procedures listed above, only the spontaneous language sample and the articulation test require any expressive language from the child. It is possible that the language sample collected later in an evaluation could be affected by the response patterns set up by such preceding tests. That is, procedures requiring little or no expressive output from the child that occur first in an evaluation may in effect be training the child to give nonverbal or minimal verbal responses. So having tests such as a receptive language test or an articulation test occur first in an evaluation may effect the length, complexity and spontaneity of a language sample recorded next in the assessment, and on the efficiency of collecting (i.e., how many adult utterances were produced). On the other hand, it is possible that beginning with procedures requiring little or no expressive language allows the child to become accustomed to the situation and thus results in a more optimal language sample. A third possibility is that the placement does not effect these parameters in a language sample. Although the sampling situation and variables affecting it have been heavily researched, no investigation of this particular question has been reported in the literature. Hence, the present study sought to determine what, if any, effect the placement of a language sample in a language evaluation (i.e., before or after formal testing) has on the length, complexity, and spontaneity of the language produced by a child.

METHOD

Subjects

The subjects for this study included 27 children aged 3-0 to 3-11 who attended several different preschools in the Austin area. All children were native speakers of English. In order to increase the generalizability of the results, no effort was made to control for either SES or race. The subjects had never been suspected of having a language delay by either family or day care personnel. Since the experimenters wished to test children who

had never before participated in a speech and language evaluation, these impressions of normal language development were corroborated post hoc by the scores the children received on the language testing done as part of the experiment (the *Test for Auditory Comprehension of Language,* Carrow, 1973). These scores revealed that all children were functioning within normal limits (± one standard deviation) for their chronological age.

Instrumentation and Procedures

The research question required that each subject be administered part of a regular speech/language assessment battery. It was also important that the experimental situation be the subject's first experience in speech/language testing so that results would be more comparable to a real assessment situation. Therefore, the subjects were not pre-evaluated to determine their eligibility as normal subjects for the study. Each subject was administered a standardized receptive language test and participated in the collection of a spontaneous language sample as part of the testing situation. Thus, the criteria for subject selection and experimental treatment were necessarily confounded. Since all the subjects performed within one standard deviation on the tests administered, no subjects were eliminated from the study.

The experimental situation took place in the pre-school of the subject and consisted of two parts: (1) a standardized test requiring only a pointing response, and (2) the elicitation of a spontaneous language same of at least 50 child utterances. The language sample was collected by the same experimenter immediately after the standardized test was administered. The *Test for Auditory Comprehension of Language* (TACL) (Carrow, 1973), a test of receptive language which requires no verbal output from the child, was utilized to establish the pointing response pattern in the subjects since the TACL or an equivalent test is almost always a part of a standard speech/language evaluation. The elicitation of the spontaneous language sample consisted of the clinician conversing with the subject while playing with some toys.

The subjects were randomly assigned to two groups. Group I (containing 14 subjects) was administered the standardized TACL after the elicitation of the language sample. Group II (containing 13 subjects) had the TACL administered at the beginning of the session, before the elicitation of the spontaneous language sample. Sample sizes were unequal because it was not possible to obtain the home-collected baseline sample for all 30 children from whom school samples were collected. Hence, three subjects had to be eliminated for the final analyses, resulting in a total of 27 subjects. The sessions were tape recorded and supplemented by notes written by the clinician.

A baselining language sample of 50 child utterances was collected in each child's home. This was done to assure that one group did not have a significantly higher MLU or DSS score in the school-collected sample simply as a result of sampling error. That is, the comparison was performed to be certain that the random assignment of subjects resulted in groups of comparable levels of language skills. In this way, obtaining significant differences between the groups for the school-collected samples could be attributed only to the language sample placement variable. The baselining language samples were obtained after the children had participated in the school-collected sample so that the children's performance in the experimental situation was not influenced by a practice effect.

Five advanced graduate student clinicians in speech and language pathology volunteered to obtain the language samples—one collected the samples in the home situation, and four others collected the samples in the testing situation. These four clinicians tested approximately eight children each. None of the clinicians had been previously acquainted with the subjects. All followed the general guidelines suggested by Lee (1974) for collecting a language sample.

Stimulus materials for the school-collected language samples were constant across all subjects and consisted of a toy camping set including a camper trailer, pickup truck, motorcycle, boat and family of four. Stimulus materials for the home baselining samples taken at each subject's home varied slightly with each subject. The clinician obtaining this sample had the same camper outfit as the clinician in the testing situation, but the children were allowed to play with any toys of their own if they wanted.

Each language sample was tape recorded, transcribed and analyzed for utterance length, complexity and spontaneity, and the number of adult utterances produced in obtaining 50 child utterances. The transcriptions included utterances produced by both the child and the clinician, and some notes regarding non-verbal action. Mean length of utterance (MLU) was calculated according to the guidelines suggested by Brown (1973). Complexity was determined by the Developmental Sentence Scoring (DSS), a standardized procedure developed by Lee (Lee & Canter, 1971; Lee, 1974). All the language samples were also analyzed for an "index of spontaneity" (Morris, 1960), which is simply the percentage of child utterances in each sample which were not direct responses to adult questions nor a repetition of a clinician's utterance.

Reliability

All the tapes of the home and school-collected samples were transcribed by the first author. No reliability testing was done for the tape transcrip-

tions since interjudge agreement is usually very high (87%-98%) between experimenters for transcriptions of tapes (e.g., Prutting, Bagshaw, Goldstein, Juskowitz, & Umen, 1978; Scott & Taylor, 1978) since the TACL consists of a mere pointing response to one of three picture plates, it was not deemed necessary to obtain an independent reliability scoring on this measure. Ten of the 54 transcriptions were independently scored for DSS, MLU and index of spontaneity by the first author and by a student clinician, both of whom were advanced level masters students in speech and language pathology. The percentages of agreement were calculated by determining the total number of utterances each judge had scored or coded identically. The percentages for the three measures follow: DSS-89%; MLU-97%, and index of spontaneity-92%. Adult speech (as child speech) was segmented into utterances using the criteria of final intonation contour (either rising or falling) and pauses between preceding and succeeding remarks (see Johnson, Darley, & Spriestersbach, 1963, p. 167). Reliability was not obtained for the number of adult utterances per sample since this entailed a straightforward counting procedure.

RESULTS

Home-Collected Samples

For each subject, a DSS score and an MLU score was obtained from the language sample collected at the child's home. From this home-collected sample, these DSS and MLU scores for children who had participated in Group I (sample first) in the school-collected sample were compared to those scores of children who had participated in Group II (sample second) in the school-collected samples. Recall that this was done in order to assure that the children randomly assigned to the two experimental groups—sample collected first versus sample collected second—were in fact of comparable language levels. As such, this first analysis involved a comparison between these two groups even though the collection of the language sample at home was identical for all subjects. Two-tailed T tests comparing the home-collected MLU scores and the DSS scores of the two groups (sample first and sample second) showed that neither reached (or even approached) the .05 level of significance (for MLU t=.014; for the DSS t=.295). It was therefore concluded that the two experimental groups were comparable in their level of language skills.

School-collected samples

The DSS scores of the subjects in the school-collected language samples were then compared. No significance between group differences were

found in DSS (t=0.19, p>.05) or MLU (t=0.38, p>.05) scores.

The index of spontaneity scores obtained from the school-collected samples of Group I (sample first) and Group II (sample second) were compared using a two-tailed t-test of independent means to determine whether a significant difference in percentage of child-initiated utterances occurred between the two experimental situations. The index of spontaneity is defined as the percentage of child-initiated utterances in a spontaneous language sample, as opposed to those utterances which are directly elicited by the clinician. For this comparison, the t score of .095 did not reach a .05 level of significance.

Although spontaneity of the child utterances was not effected by language sample placement, it was informally noted that one of the four clinicians who collected the school samples seemed to elicit more spontaneous language, as measured by the "index of spontaneity" (Morris, 1960), than the other clinicians. This clinician seemed to have an interaction technique that included use of fewer direct questions, more comments on action-at-hand, and more expansions of the child's utterances. The other three clinicians seemed to have a different interaction technique involving the use of more exact repetitions of the child's utterances and greater persistence in asking questions to which the child was not responding. A t-test was used to determine whether a significant difference did indeed exist between these two intuitively-derived groups (i.e., clinician with different style compared to other three clinicians) in the percentage of child-initiated utterances each elicited (again, using Morris' 1960 criteria for spontaneous or child-initiated utterances). The t score of 1.764 did not reach the .05 level of significance but did reach a .10 level of significance. These results suggest that clinician style of interaction could have an effect on the spontaneity of the language sample obtained. Although statistical significance was not reached, these findings suggest that future research regarding language sample collection might fruitfully focus on variability in clinician interaction styles.

Again using a t-test for comparison, the number of adult utterances produced in obtaining 50 spontaneous child utterances (that met DSS requirements) was not effected by the placement variable in the school-collected samples. However, when the amount of adult talk produced in the home- versus school-collected samples was compared, a significant difference emerged—significantly fewer adult utterances were produced in the home-collected samples (t=2.90, p<.05).

DISCUSSION

The main question of this study was "What effect does placement of the elicitation of the language sample in the assessment battery have on

the language sample obtained?" The results of this study suggest that placement of the elicitation of the spontaneous language sample has no effect on the length, complexity, or spontaneity of the sample obtained, nor does it influence the amount of adult talk produced in obtaining that sample. In general, preceding the elicitation of the language sample with a test requiring only a pointing response does not seem to yield a less optimal language sample than placing the elicitation of the language sample first in the assessment battery. It should be remembered, however, that all the subjects were children with normal language development. It is still possible that children with language disorders might respond differently if they are not "locked into" a pointing response pattern before the elicitation of the language sample.

Examiner variables might also have influenced the results of the study. Although the four testing clinicians had similar background and experience, it was informally noted that they did not elicit the samples in exactly the same manner. The clinicians varied in the number of direct questions they asked the subjects, in the amount of time they would spend pursuing an area of questioning which did not seem to be eliciting complete utterances from the subject, and in the number and length of pauses they would allow to occur in the conversation. Therefore, the difference in eliciting styles of the examining clinicians could have affected the results of the study. This might also explain why fewer utterances occurred in the home-collected samples. Recall that all of these samples were collected by one clinician. She may have exhibited a different interaction style from the other four clinicians who collected the experimental language sample.

From the current study, it can be concluded that the placement of the elicitation of a spontaneous language sample in an initial language evaluation does not influence the length, complexity, or spontaneity of the sample obtained. Further research is needed to explore the generalizability of these results to the language disordered population. Furthermore, the post-hoc analyses conducted point out another potentially fruitful area of investigation regarding language sample collection—that of the clinician's interaction style.

ACKNOWLEDGEMENTS

This study was conducted as a Master's thesis project at the University of Texas at Austin, and was partially funded by a grant from the Shell Foundation.

Anne van Kleeck
Department of Speech Communication
University of Texas at Austin
Austin, Texas 78712

REFERENCES

BROWN, R. (1973). *A first language: The early stages.* Cambridge, MA: Harvard University Press.

CARROW, E. (1973). *Test for auditory comprehension of language.* Austin, TX: Learning Concepts.

CAZDEN, C.B. (1970). The neglected situations in child language research and education. In F. Williams (Ed.), *Language and poverty.* (pp. 81-101). Chicago: Rand McNally College Publishing Co.

JAMES, S.L. & BUTTON, M. (1978). Choosing stimulus materials for eliciting language samples from children with language disorders. *Language, Speech and Hearing Services in the Schools, 10,* 91-97.

JOHNSON, M.R. & TOMBLIN, J.B. (1975). The reliability of developmental sentence scoring as a function of sample size. *Journal of Speech and Hearing Research, 18,* 372-380.

JOHNSON, W., DARLEY, F., & SPRIESTERBACH, D. (1963). *Diagnostic methods in speech pathology.* New York: Harper and Row.

KRAMER, C.A., JAMES, S.L., & SAXMAN, J.H. (1979). A comparison of language samples elicited at home and in the clinic. *Journal of Speech and Hearing Disorders, 44,* 321-330.

LEE, L. (1974). *Developmental sentence analysis.* Evanston, IL: The Northwestern University Press.

LEE, L. & CANTER, S. (1971). DSS: A clinical procedure for estimating syntactic development in children's spontaneous speech. *Journal of Speech and Hearing Disorders, 36,* 315-340.

LONGHURST, T.M. & FILE, J.J. (1977). A comparison of development sentence scores from head start children collected in four conditions. *Language, Speech and Hearing Services in the Schools, 8,* 54-64.

LONGHURST, T.M. & GRUBB, S. (1974). A comparison of language samples collected in four situations. *Language, Speech and Hearing Services in the Schools, 5,* 71-78.

MORRIS, H.L. (1960). Communication Skills of Children with Cleft Lips and Palates. Unpublished dissertation. State University of Iowa.

MUSSELWHITE, C.R. & BARRIE-BLACKLEY, S. (1980). Three variations of the imperative format of language sample elicitation. *Language, Speech and Hearing Services in the Schools, 11,* 56-67.

OLSWANG, L.B. & CARPENTER, R.L. (1978). Elicitor effects on the language obtained from young language-impaired children. *Journal of Speech and Hearing Disorders, 43,* 76-88.

PRUTTING, C.A., BAGSHAW, N., GOLDSTEIN, H., JUSKOWITZ, S., & UMEN, I. (1978). Clinician-child discourse: Some preliminary questions. *Journal of Speech and Hearing Disorders, 43,* 123-139.

SCOTT, C.M. & TAYLOR, A.E. (1978). A comparison of home and clinic gathered language samples. *Journal of Speech and Hearing Disorders, 43,* 482-495.