Verbal Planning Functions in Children's Speech

Brian MacWhinney
University of Denver

Harry Osser
Queen's University

MACWHINNEY, BRIAN, and OSSER, HARRY. Verbal Planning Functions in Children's Speech. Child development, 1977, 48, 978-985. This study sought to examine the role of communicative explicitness, sex, and social class upon children's utilization of a wide variety of hesitation phenomena. Speech samples were gathered from 20 British 5-year-olds under 3 conditions of elicitation. Occurrences of 13 varieties of hesitation phenomena were tabulated for each corpus and subjected to a multivariate test for equality of mean vectors for the 2 levels of sex and social class. Although there was a significant main effect for sex, there was no significant main effect for social class and only a weak interaction of sex with social class. The verbal planning functions of the several hesitation phenomena were examined by principal components analysis and interpreted in the light of textual relations. It was found that hesitation served 3 major functions: preplanning of verbalization not yet produced, coplanning of verbalization currently being articulated, and avoidance of superfluous verbalization.

Much contemporary research on hesitation phenomena in speech was derived from the work of Goldman-Eisler (1968), who argues that the analysis of speech pauses provides an external window upon the internal constructive processes of speech selection and organization. In some of the earlier studies of hesitation phenomena (Goldman-Eisler 1958a, 1958b; Lounsbury 1954), it was claimed that, once a speaker had begun a sentence, pauses only reflected lexical selection processes. This analysis was in accord with the information-theory view of sentence processing as a Markov process.

The advent of transformational grammar (Chomsky 1957) focused attention on the relations between hesitations and phrasal structure. Maclay and Osgood (1959) and Boomer (1965) found that pauses were related not only to lexical selection points, but also to grammatical junctures corresponding to phrasal units. More recently, Butterworth (1975) has pointed to the existence of macrorhythms in pause location correlated with idea units as well as clausal structure. This emphasis on ideational, syntactic, and lexical factors in decision making is in accord with current models of procedural semantics (Halliday 1973; Winograd 1973), semantic network systems (Anderson & Bower 1974; Kintsch 1974; Linde & Labov 1975), and the pragmatic theory of reference (Cole & Morgan 1975; Rommetveit 1974).

According to the pragmatic theory of reference, communications differ systematically in terms of levels of explicitness. Communications based on large amounts of shared knowledge can be rich in anaphora, ellipsis, and deixis. However, when the speaker cannot be sure that the hearer shares certain information with him, he must encode his intentions explicitly. When a communication is about concrete material, it is relatively easy to obtain a satisfactory level of explicitness, par-
particularly when the perceptual display is available to both speaker and hearer (Glucksberg & Krauss 1967). However, when a communication is about abstract material, it becomes increasingly difficult to maintain a satisfactory level of explicitness. In terms of processing time, increased levels of abstraction and explicitness lead to increased demands on the real-time processing of the system of semantic options (Halliday 1973). Inexplicit utilization of the system is less time-consuming than explicit utilization, all other things being equal.

Goldman-Eisler (1968, pp. 50-59) has shown how pauses increase with tasks that demand high levels of explicitness. Thus, subjects spent over twice as much time pausing when required to explain the meaning of a picture as when simply required to describe the picture. Similar results were obtained by Levin, Silverman, and Ford (1967) with children for use of pauses and several other hesitation phenomena.

Bernstein (1962) extended Goldman-Eisler's (1968) finding on pauses and levels of abstraction to social class differences in explicitness. He found that, compared to middle-class adolescent boys, lower-class adolescent boys spent significantly less time pausing, used a significantly longer mean phrase length, and had a significantly shorter mean word length. Following Goldman-Eisler's interpretation of pauses as indicators of verbal planning time, Bernstein argued that the lower-class boys were spending less time planning, presumably because the speech sequences they were producing were heavily "preceded" and, therefore, highly accessible. In a related study, Hawkins (1973) analyzed silent pauses in the speech of lower- and middle-class boys and girls. He found that the total number of pauses was virtually the same for all groups. However, differences did show up when the length of each pause was classified. For example, the middle-class boys had significantly more pauses in the 2-3-sec category than the lower-class boys. Second, the lower-class girls showed significantly more short pauses—up to 1 sec—than the middle-class girls, who had significantly more intermediate pauses, from 1 sec on, than their lower-class counterparts.

Hawkins argues that the fact that the middle-class children not only used significantly more "intermediate" length pauses than the lower-class children but also used more of them for within-clause lexical planning indicates that the middle-class children spent more time on lexical decisions than the lower-class children.

Bernstein (1972) has placed these, and other, results within the context of his sociolinguistic theory of socialization. The cornerstone of this theory is the contrast between the "elaborated" code of the middle class and the "restricted" code of the lower class. Bernstein holds that middle-class speakers use language abstractly, analytically, and explicitly, whereas lower-class speakers use language to establish group solidarity through inexplicit, ritualized (preceded), concrete communications.

Two theses arise out of the last 2 decades of research on hesitation phenomena. First, it appears that the number and distribution of pauses are dependent upon the extent to which the communication must be explicit. Second, it appears that there are social class differences in England in the use of pause phenomena. However, it should be noted that Bernstein's hypothesis refers not only to the use of silent pauses, but to all hesitations involved in verbal planning, whereas Hawkins (1973) has shown that class differences in verbal planning are not picked up by all hesitation measures. It may be the case that the two classes differ in their choice of hesitation phenomena, while not differing in the total amount of verbal planning they utilize. A similar argument can be made for task differences. Some verbal tasks may involve more of one type of hesitation than another. However, the total amount of verbal planning might be constant across all tasks. Thus, it may well be the case that there are no overall verbal planning function differences between the two social classes.

The present study is designed to examine these objections to earlier demonstrations of a relation between social class and verbal planning functions. It examines 15 categories of hesitation phenomena in the speech of children who, according to Bernstein's theory, would be likely to take up either the elaborated or the restricted code as their primary mode of communication. The speech was elicited by three verbal tasks. The tasks were designed to require different levels of explicitness. Thus, the study attempts to evaluate Bernstein's hypothesis against a wider array of tasks and measures than has been examined in previous research.
Children Development

**Method**

*Subjects.*—The subjects were 20 children in the same class in an infant school in the suburbs of London. There were 10 lower-class children (whose fathers were generally semiskilled workers) and 10 middle-class children (whose fathers were generally professionals). There was an equal number of boys and girls in each group. The average age for the lower-class group was 57 months, with a range of 51-62 months, and for the middle-class group the average age was 55 months, with a range of 51-59 months.

**Procedures**

*IQ testing.*—Each child was given the Stanford-Binet test. The IQ scores of the four sex and social class combinations were closely matched.

*Speech tasks.*—The children were given three communicative tasks.

*Picture task:* Three paintings by the French primitive painter Troït were presented to the child, and he was asked to describe them. The task was designed to tap the child’s ability to describe objects and their spatial relations.

*Game task:* The child was told by the experimenter that he did not know how to play certain children’s games (e.g., hide-and-seek) and the child was asked to explain how the game is played. The task was designed to tap the child’s ability to explain the logical and hierarchical relationships between objects and events.

*Conversation task:* The child was asked to respond to a number of questions about his day-to-day experiences (e.g., Do you go shopping with your mother? What do you do together?). The task was designed to tap the child’s ability to narrate events in temporal sequence.

These tasks were selected to require different degrees of explicitness in encoding. The game task was assumed to tap more "interpretive" processes than the others. The picture task and the conversation task differed from each other in that the former provided the speaker with visual support for his verbal performance, whereas the latter was without this support. However, the conversation task was simpler in one respect in that in the picture task the child was being asked to encode rather novel material, as the content of the pictures was somewhat strange to the children. The experimenter (H. O.) spent a week in the school before beginning the experiment.

**Data Measurement Techniques**

*Pause measurement.*—The total corpus contained over 21,000 words. A polygraph printout gave pitch and amplitude traces for the whole corpus. Pauses could be recognized as low spots or valleys in the pitch and amplitude traces. It was possible to measure the pauses to an accuracy of ±30 msec. Pauses below the limit of 250 msec were not included in the analysis, since these short pauses are principally artifacts of certain articulatory movements.

*Hesitation types.*—Fifteen measures were taken for each of the 20 children on each of the three tasks. The first two of these measures were time measurements:

1. **Seconds of unfilled pausing:** The number of seconds spent in unfilled pausing. The unfilled pause is defined as a period of more than 250 msec of nonvoicing occurring within the flow of voicing.

2. **Seconds of initial pausing:** The number of seconds spent in initial unfilled pausing. Initial pauses occur between the end of the experimenter’s probe and the beginning of the child’s response. If the child pauses, but fails to orient to the experimenter’s question, the resulting pause is judged to be noninitial.

The next 11 measures simply record the number of times a phenomenon occurred.

3. **Unfilled pauses:** As defined under (1) above.

4. **Initial pauses:** As defined under (2) above.

5. **Filled pauses:** Examples of this hesitation phenomenon include fillers such as /mm/, /Am/, and /A/, and the words "well" and "you know" when they are not used to convey a particular meaning.

6. **Drawls:** Drawls are marked prolongations of a speech sound without any pause in articulation, as in "gamesss."

7. **Initial segment phonological repetitions:** Here the repeated or stuttered material is less than a word in length and is usually only the first segment of a word, as in "b-b-boy." In the present data, any repetition of segments always includes the initial segment, except for one case of internal repetition: "ba-na-na-nas."

8. **Word-included phonological repetitions:** This phenomenon is like phenomenon (7) above in that there is a repeated string of segments shorter than a word, always including the initial consonant. However, the repetition also includes the word immediately preceding the initial segments, as in "I c- I can’t."

9. **Word repetitions:** This phenomenon involves repetition of a word without change and for
no grammatical purpose, as in "I want some
more."

10. Several-word repetitions: This is
identical to (9) except that several words are
repeated. Example: "I want some more some
more ice cream."

11. Sentence incompletion: Sentence
incompletions involve a string of words which
is intelligible, but which does not form a
complete phrase or sentence, due to a sudden
truncation before completion. Utterances
obeying the standard rules for ellipsis are
excluded from this category. Example:
"That's a... Where does that go?"

12. Retraced false starts: False starts
involve the correction of one or more words
through substitution, deletion, or addition of
material. Typically, the entire phenomenon
consists of three parts: (a) the material to be
corrected, (b) an editing signal which marks the
end of the material to be retraced, and (c) the
retracing correction which follows the editing
signal and continues with the train of speech.
This editing signal may come from at least one
of five distinct categories: unfilled pause, filled
pause, glottal stop, corrective word (e.g.,
"rather"), or raised pitch and volume.

13. Phonological corrections: This is the
shortest form of retraced false start and never
extends beyond the length of a word in these
corpora. Example: "musical... musical."

The final two measures were of the numbers
of words and sentences produced:

14. Words.

15. Sentences.

In total there were 15 measures taken for each
child on each task.

Results

No significant differences among the pattern
of results on the three verbal tasks were
identified. For this reason, the data were
merged across tasks.

The basic design involved a two-factor
analysis of multivariate with two levels of
the sex factor, two levels of the social class
factor, 15 linguistic variables, and the IQ
measure. Because one of the four sex x social
class cells was empty for initial segment
phonological repetition and for phonological
correction, it was necessary to merge these
variables with other functionally similar cor-
related variables. Since initial segment
phonological repetition correlated with word-
included phonological repetition at a level of
.45, these two measures were merged into a
general measure labeled phonological repeti-
tions. Similarly, phonological corrections

MacWhinney and Osser 981

were merged with retraced false starts into a
general measure labeled retraced false starts.
The 15 linguistic variables were thus reduced
to the 13 found in tables 1, 2, and 3.

A second correlation matrix for these 13
variables and the IQ measure was obtained
together with a set of principal components.
The raw scores were corrected for the overall
output factor by applying weights from the
first principle component. Using the corrected
values, a third correlation matrix was ob-
tained. Since the univariate effects of IQ
within each of the multivariate tests were not
significant, and since the design had closely
matched the IQs for each of the four sex and
social class combinations, further analysis was
confined to the 13 linguistic variables alone.
Finally, a fourth correlation matrix and set of
principal components was obtained for the 13
linguistic variables. The values for each of the
13 variables were again corrected by this new
first principal component.

Using these corrected values, the mul-
tivariate test of equality of mean vectors
yielded the following multivariate F ratios:
for the main effects of sex, F(13,4) = 7.72, p < .03;
for the main effects of social class, F(13,4) =
0.72, p < .71; and for the interaction of sex
and social class, F(13,4) = 4.30, p < .08. Although
the interaction of sex x social class is not
statistically significant at the p < .05 level, the
effect is strong enough to warrant further con-
sideration.

The variables whose production differs
significantly (p < .05) between sexes are sec-
onds of initial pausing, unfilled pauses, filled
pauses, word repetitions, and retraced false
starts. Table 1 lists the cell means of all the
variables after the correction for output. This
table shows that boys produced more unfilled
pauses, filled pauses, word repetitions, and
retraced false starts than girls. However, girls
produced more seconds of initial pausing than
boys.

In the case of the univariate F ratios
associated with the interaction, only filled
pauses and retraced false starts show values
significant at p < .05. The value for words atp
< .06 is just outside the acceptable range for
significance. As can be seen in the cell means
in table 1, the values on the measures of filled
pauses, retraced false starts, and words are
nearly identical for both sexes of the middle
class. However, in comparison to the middle
class, the lower-class girls produced far less
on these measures, while the lower-class boys produced far more.

The sample correlation matrix is printed in table 2 and the rotated (VARIMAX) factor matrix is printed in table 3. The first three components in the rotated factor matrix of table 3 account for 97% of the total variance. The first of these components shows high absolute weights for seconds of unfilled pausing, unfilled pauses, filled pauses, sentence in-completions, retraced false starts, and words. The functional similarities between these variables indicate that this first component represents the contribution of speech planning and speech monitoring processes occurring during the articulation of an utterance. In the interests of brevity, this function will be referred to as coping planning.

The second component, with heaviest weights on seconds of initial pausing and initial pauses, appears to represent the contribution of speech planning processes occurring before the beginning of articulation. This second verbal planning function will be referred to as prepacing.

The third component has high weights for seconds of initial pausing and word repetitions, but these weights are in opposite directions. Looking at some of the variables with lesser weights, we find that several-word repetitions, filled pauses, and retraced false starts have moderate positive weights, while draws have a moderate negative weight. Thus, in this component, hesitation phenomena which gain time through silence or drawing are contrasted with those gaining time through repetitions of material or insertion of meaningless material. Thus, this third component may be characterized as involving the avoidance of superfluous verbalization in planning and monitoring occurring during production of an utterance. The last four components account for little of the total variance and will receive no further consideration.

The 151 recorded retraced false starts were quite clearly involved in attempts to increase explicitness. The total of 151 included 18 number corrections, 10 person corrections, 12 gender corrections, 5 count to mass corrections, 13 pronoun to noun corrections (i.e., "Mommy" for "I"), 8 additions of adjectives to nouns, 8 additions of conjunctions, and 3 deletions. In addition, there were 20 topicalization sequences such as "sand, we all play sand." The other 54 retracions involved 16 changes within categories (i.e., "sand" for "water") and 38 shifts within subcategories (i.e., "ice cream" for "jello"). The important point is that, without exception, these retracions increased the explicitness of the element being retraced.

Word repetitions, on the other hand, involve by their very nature no increase in explicitness. What is interesting is the fact that common words were most often repeated. Thus, of the 171 repetitions from all the children, 22% were of subject personal pronouns, 22% of the conjunction "and," 16% of articles, 8% of prepositions, and 7% of deictic pronouns. Together, these categories constituted only slightly over a dozen lexical items, but accounted for 75% of all word repetitions. Since the selection of such common words...
<table>
<thead>
<tr>
<th></th>
<th>Seconds of Unfilled Pausing</th>
<th>Seconds of Initial Pausing</th>
<th>Unfilled Pauses</th>
<th>Initial Pauses</th>
<th>Filled Pauses</th>
<th>Drawls</th>
<th>Phonological Repetitions</th>
<th>Word Repetitions</th>
<th>Several-Word Repetitions</th>
<th>Sentence Incompletion</th>
<th>Retraced False Starts</th>
<th>Words</th>
<th>Sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seconds of Unfilled Pausing</td>
<td>1.00</td>
<td>1.00 .36</td>
<td>1.00 .04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seconds of Initial Pausing</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unfilled Pauses</td>
<td>.62 .90</td>
<td>.58 .15</td>
<td>1.00</td>
<td>.38 -</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial Pauses</td>
<td>.39 .57</td>
<td>.75 .13 -</td>
<td>.24 .61</td>
<td>.16 .09</td>
<td>1.00 .25</td>
<td>1.00 .03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Filled Pauses</td>
<td>.45 .48</td>
<td>.36 -</td>
<td>.16 .61</td>
<td>-.10</td>
<td>.11 .25</td>
<td>-.21 -</td>
<td>1.00</td>
<td>1.00 .65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drawls</td>
<td>.45 .48</td>
<td>.36 -</td>
<td>.16 .61</td>
<td>-.10</td>
<td>.11 .25</td>
<td>-.21 -</td>
<td>1.00</td>
<td>1.00 .65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phonological repetitions</td>
<td>Word repetitions</td>
<td>.15 .23</td>
<td>.26 .50</td>
<td>.25 .09</td>
<td>.23 .42</td>
<td>.25 .34</td>
<td>1.00</td>
<td>1.00 .45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sentence incompletion</td>
<td>.80 .39</td>
<td>-.16</td>
<td>.91 .61</td>
<td>-.15</td>
<td>.31 .52</td>
<td>-.30 -</td>
<td>.68 .75</td>
<td>.20 .51</td>
<td>.51 .40</td>
<td>1.00 .63</td>
<td>1.00</td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>Retraced false starts</td>
<td>-.16</td>
<td>.91 .61</td>
<td>-.15</td>
<td>.31 .52</td>
<td>-.30 -</td>
<td>.68 .75</td>
<td>.20 .51</td>
<td>.51 .40</td>
<td>1.00 .63</td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>Words</td>
<td>.66 .60</td>
<td>.10 .42</td>
<td>.83 .63</td>
<td>.04</td>
<td>.31</td>
<td>.02 .19</td>
<td>.71 .54</td>
<td>.08 .08</td>
<td>.77 .58</td>
<td>.76 .35</td>
<td>.66 .60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sentences</td>
<td>.66 .60</td>
<td>.10 .42</td>
<td>.83 .63</td>
<td>.04</td>
<td>.31</td>
<td>.02 .19</td>
<td>.71 .54</td>
<td>.08 .08</td>
<td>.77 .58</td>
<td>.76 .35</td>
<td>.66 .60</td>
<td></td>
<td>1.00</td>
</tr>
</tbody>
</table>
should not constitute a major search task, it appears that words are repeated not because of uncertainties about their own production, but in order to gain time for other decisions. In the case of several-word repetitions, the situation is similar. Commonly repeated structures are: subject personal pronoun + verb; preposition + article; and conjunction + personal pronoun.

**Discussion**

The first major result of this study has been the identification of three major planning functions: preplanning, coplanning, and avoidance of superfluous verbalization. These styles in verbal planning reflect basic differences in cognitive processing. Underlying all three planning functions, however, is one central commonality—verbal planning takes time. While the speaker is trying to figure out what to say and how to say it, the conversation moves on. Given this inevitable forward movement in time and his own problems in formulating his utterance, the speaker may do one of two things. He may attempt to fully formulate what he is going to say before he says it. Alternatively, he may start talking and hope to be able to figure out his utterance in medias res. Whether he pauses initially or attempts to patch together an ongoing sentence, he has a further option. He may either use superfluous verbalization to cover his pauses and errors or he may simply remain silent. The principal components analysis in this experiment indicates that the 13 hesitation phenomena examined in this study can be grouped into these three functional categories: coplanning, preplanning, and avoidance of superfluous verbalization. The association of retracted false starts with Co-planning and of word repetitions, several-word repetitions, and filled pauses with preplanning echoes the relations found by Maclay and Osgood (1959) between, on the one hand, false starts and "lexical" words and, on the other hand, between pauses and filled pauses and "function" words.

The second major result of this study has been the finding that, for 5-year-olds, differences in verbal planning functions are more related to sex than to social class. Boys were found to do more coplanning, while girls made greater use of preplanning. Moreover, boys showed more use of superfluous verbalizations than girls. These differences extended to both social classes. It is important to note that the interaction effect between social class and sex was strongest for filled pauses, retracted false starts, and words, which were the central variables in the coplanning component. Whereas both sexes of the middle class produced roughly equal amounts of filled pauses, retracted false starts and words, there were great differences for these measures within the lower class. In effect, the girls illustrated highly preplanned speech and the boys illustrated highly coplanned speech. If the interaction effect were to be ignored, a fairly clear picture of care and explicitness among girls as opposed to wordiness, carelessness, and lack of planning among boys would arise from these data. However, this basic picture might eventually require modification, if the interaction effect were found to be of significant proportions.
It appears then that sex has a greater effect than social class in determining differential use of hesitation phenomena in British 5-year-olds. It is important to recognize that Bernstein's theory is primarily a theory of class differences. Although the 1972 reformulation recognizes the existence of differences in explicitness between older boys and girls in the lower class, the theory still implies that class differences should be more important than sex differences in determining verbal planning functions. This relative importance of class versus sex is not confirmed in the present data. Moreover, Bernstein has suggested that sex differences between lower-class boys and girls in verbal planning would first emerge in adolescence. However, the present data, like those of Hawkins (1973), show these differences in 5-year-olds.

At the very least, these data indicate the need for studies of a large variety of hesitation phenomena as used in maximally diverse communicative situations by widely different children of different age groups. Until such studies have been conducted, all hypothesized relations between social class, explicitness, and hesitations remain quite tentative.

References


Goldman-Eisler, F. Speech production and the predictability of words in context. Quarterly Journal of Experimental Psychology, 1958, 10, 96-106. (a)

Goldman-Eisler, F. The predictability of words in context and the length of pauses in speech. Language and Speech, 1958, 1, 226-231. (b)


