Social Speech and Social Interaction: Egocentrism Revisited

Catherine Garvey and Robert Hogan

Johns Hopkins University

GARVEY, CATHERINE, and HOGAN, ROBERT. *Social Speech and Social Interaction: Egocentrism Revisited.* CHILD DEVELOPMENT, 1973, **44**, 562–568. To study the development of social speech, 18 dyads of children (3½–5) were videotaped in 15-minute play sessions. Behavior was coded in terms of the time children spent in mutual interaction, and speech was coded in terms of the degree to which utterances were adapted to the verbal or nonverbal behavior of the partner. The results indicated a high level of mutual responsiveness in both speech and behavior. The paper suggests that children in this age range are capable of genuinely social behavior; it concludes that early forms of social speech entail a surprising level of interpersonal understanding, and that these speech forms are amenable to systematic study.

Studies of young children's speech typically emphasize its egocentric, parasocial, or private quality (Kohlberg, Yeager, & Hjertholm 1968; Piaget 1926; Vygotsky 1962). According to these studies, the progressive decline in overt forms of private speech is accompanied by a growth in collaborative activity (Piaget 1926); by the ability to recode or modify messages taking into account the informational needs of an interlocutor (Flavell, Botkin, Fry, Wright, & Jarvis 1968); or by the development of inner speech or thought (Vygotsky 1962). The attention paid to the form, function, and fate of private speech rests in part on the assumption that the thought and behavior of children is initially egocentric and that with the passage of time their actions become increasingly social as a result of cognitive development and social experience. Such a viewpoint has unfortunately led social scientists to neglect children's early social interaction. Recent research suggests, however, that it may be useful to regard children as "sociocentric" essentially from birth (e.g., Borke 1971; Ferguson 1971; Stayton, Hogan, & Ainsworth 1971). Even the first stages of language acquisition may depend on the child's ability to discriminate the intent of others (cf. Macnamara 1972); if so, then the view of the child as initially egocentric requires reexamination. A role-theoretical perspective (Sarbin & Allen 1968) further suggests that language, the origins of social behavior, and the development of interpersonal understanding (in the sense suggested by Weber's term "meaningful behavior" [cf. Winch 1958]) are all related. It may well be that a major function of early language use is social, in the sense of establishing and maintaining interpersonal contact. As such, children's talk could also serve as a vehicle for learning those concepts that underlie social intercourse (e.g., concepts of reciprocity, obligation, and complementarity).

Before the uses and forms of social speech can be understood, it is necessary to know how frequently and extensively such speech occurs and how it occurs in relation to other behaviors in a potential social interaction. Mueller (1972) found that among previously unacquainted children (ages 3½–5½) in a dyadic play situation, the majority of utterances (62%) received a definite response. Only 15% of the utterances failed to elicit a response, while 23% attracted the listener's attention. The present paper examines the distribution of activity throughout the play sessions to determine the extent to which children engage in spontaneous interaction. It

This work was supported by research grant #1 RO3 MH 21610-01 from the Public Health Service, National Institute of Mental Health. The authors wish to thank Edward Mueller for his comments on the manuscript and Richard Haier for his assistance in the collection of the data. Authors' address: Department of Psychology, The Johns Hopkins University, Baltimore, Maryland 21218.

[*Child Development*, 1973, **44**, 562–568.

further inquires how extensively the interactions created by talk are sustained. Finally, the paper describes, as an example of the social use of speech, how a state of mutual engagement is created.

Method

Subjects.—Eighteen white, middle-class children in the 3½–5-year age range served as subjects. These children were recruited from local nursery schools where the directors were asked to nominate children of English-speaking "professional" families. There were 11 girls and seven boys.

Procedures.—Three children from the same school were brought at one time to the laboratory by their teacher. The triad was composed of both sexes and was either younger (age 3½–4⅓) or older (4½–5). The children met the adult observers, saw the observation room, and then drew straws to see who would "play some games" and who would go as a dyad to the playroom. The child who drew the long straw went into the observation room with two adults where he was asked to choose objects which were "same" or "different." The dyad was left alone in the playroom where their activity was videotaped. The playroom contained toy telephones, a tool belt, a wooden car (big enough for two to sit on), dress-up clothes, a large stuffed snake, and fish, blocks, cars, trucks, an iron, and a broom. There were posters on the walls, a carpet, a couch, and curtains drawn back from the windows (one-way mirrors). The sessions lasted approximately 15 minutes and varied in length so that the children were not interrupted in the middle of a game. After each session the dyad was changed so that three dyads were formed from a triad. Thus, there were six younger dyads and 12 older dyads.

Methods of analysis.—Verbatim transcripts were made of the audio material. All sessions were automatically timed at 15-second intervals and the speech was further divided into utterance units (UU)—stretches of one person's speech separated by pauses greater than 1 second or by another person's speech. The UUs were numbered and counted for both members of the dyad, and the mean length of the UUs in each dyad was calculated for each speaker. Rate of utterance for the dyad was determined by dividing the number of seconds in the session by the total number of UUs.

Judgments of mutual engagement, or focused interaction (e.g., a state in which the actions of members of a dyad are interdependent [cf. Goffman 1963]), were made on the videotaped data of 12 dyads. Three independent judges indicated (anchoring their judgments to the nearest UU) when they thought the children moved into or out of a state of mutual engagement or focused interaction. The judges, following written instructions, could request replay of parts of the tape, and they indicated their decisions (in focus, out of focus, unsure) to the experimenter. Interjudge agreement was estimated using a random sample of 60 UU points from each dyad's session. Overall agreement was 82%. A major source of disagreement appeared to be the judges' inability to concur on the exact point of the beginning or ending of the two states of focused and unfocused interaction. The amount of time children were in focused interaction was calculated by dividing the time all three judges indicated was spent "in focus" by the total time of the session. This figure, then, was based on periods for which 100% agreement was obtained.

Social speech was defined as speech that is strictly adapted to the speech or behavior of the partner. To determine the extent to which the children's speech was social, each UU was coded according to its result. It was thus the interaction which was coded, rather than the speakers' intent, the nature of which, of course, the judges could only infer. Five categories of results, or consequences, of the UU were distinguished:

1. No apparent consequence (pause ensues, or speaker continues without response from listener, who may ignore him or may have failed to hear him)
2. Unrelated speech (listener speaks immediately after UU but speech is unrelated to UU)
3. Attending behavior (listener turns to speaker, watches or appears to listen)
4. Appropriate nonspeech behavior (listener performs action specified by UU)
5. Appropriate speech (listener replies to UU)

There obviously could be several immediate consequences of an UU, as in the case where a request, "Help me fasten this belt," was followed by compliance (appropriate nonspeech behavior) *and* another UU, "I'll try" (appropriate speech). To avoid multiple coding of an UU only the highest numbered relevant code was assigned, in this case, 5.

A sixth category (0) included those cases in which an action on the part of one child elicited a question or comment from the other. An UU coded 4 or 5 was said to form an exchange with its consequence, and an UU coded 0 an exchange with its antecedent. Thus, an exchange is formed of two component events, at least one of which must be an UU. An exchange conformed to a criterion of fixed order, that is, a change in order of the component events would destroy or change the meaning of the exchange. A series of UUs produced by one speaker was bracketed. Next, sequences of exchanges, which were formed from groupings of 4, 5, and 0 codes, were identified. A sequence could be broken by incidence of code 1 or 2 from the alternating speaker. However, if a code 1 was bracketed with a code 4 or 5, the sequence was not broken. Table 1 illustrates the coding categories, the bracketing, and the groupings of exchanges into exchange sequences.

A check of interjudge agreement in the use of the coding system showed that judges concurred in their codings of 89% of the total UUs. The coded UUs were tallied and the percentage of UUs which entered into exchanges was calculated. Sequences of exchanges were also tabulated.

Results

The analysis of rate of utterance and distribution of UUs by speaker indicated that the sessions contained a high level of verbal activity. The overall rate of utterance for 18 dyads was one UU every 4.6 seconds, and the average rates in the first and in the second half of the sessions were approximately the same. The rate of the younger dyads slightly exceeded that of the older. This does not mean that younger children spoke more rapidly; rather it means that their UUs were shorter. The overall rate was twice that reported by Mueller (1972) for his similar but longer observation sessions. The fact that the children in the present study were previously acquainted whereas Mueller's subjects were strangers may account for this higher rate of utterance. The overall mean and standard deviation for words per UU was 5.61 and 1.51; for younger dyads, 4.73 and 1.00; for older dyads, 6.04 and 1.65.

TABLE 1

SAMPLES OF CODING CATEGORIES, BRACKETING, AND GROUPING OF EXCHANGES

UU No.: Speaker A	UU No.: Speaker B	Code	No. of Exchanges
(Hums theme song from television show)		0	...
	127. We watch that.	5	...
128. Me, too.		5	...
	129. Isn't it funny?	5	...
130. I know, it sure is.			4
	176. My group got shot.	1	...
	177. Did your group ever get shot?	5	...
178. No.		5	...
	179. Well, mine did.	3	...
			2
180. Will you put this hammer in (tool belt) for me?		4	...
	(B puts hammer in)	0	...
			2
181. Thank you, I'll take it (tool belt) now.		1	...
182. Oh, the phone's ringing.	(B glances at A, then	2	...
183. Hello?	moves away, picks up snake)	3	...
	184. A snake can kill you.		
			0

The mean number of words and standard deviation per UU for younger girls was 5.05 and 1.28; for younger boys, 4.41 and 0.90. Among the older dyads, the mean number of words and standard deviation per UU for girls was 6.25 and 1.97; and for boys, 5.76 and 1.03.

The distribution of participation in the verbal activity was approximately equal for both members of each dyad. A dyad was said to be balanced in respect to relative participation if the percentage of UUs contributed by both members was equal or differed by 5% or less. Eleven of the 18 dyads were balanced. Of the remaining seven dyads, three showed differences of 20%; the remaining four showed differences of 12%, 16%, 16%, and 18%, respectively. Sex did not account for the unbalanced participation since in the seven unbalanced dyads (all but two of which were composed of both sexes) three boys were dominant and two girls were dominant. Nor were there consistent individual tendencies to domination or subordination in respect to production of UUs. Only three individuals consistently participated less in relation to *both* partners in their respective triads, and no individual dominated in both dyads in which he or she participated.

The dyads used for the focus judgments were considered to be "in focus," or mutually engaged, for an average of 66% of each session. Bar graphs drawn of the judgments over each continuous session showed an alternation of focused and unfocused periods. Periods judged "out of focus" were generally brief. Only one "out-of-focus" period exceeded 2 minutes, and this occurred at the beginning of a session. Many focused periods were also brief; however, each dyad showed sustained periods of mutual engagement. The basis for these sustained engagements included games and activity centered on toys or objects in the room, most usually with accompanying speech. More important, many periods of interaction included talk as the primary common focus of attention. In the following example the boy (A) and the girl (B) sat quietly, glancing occasionally at each other (this excerpt, composed of six exchanges, follows a sequence of exchanges concerning what A wanted to be when he grew up):

A. If I grow up my voice will change and when you grow up your voice will change. My mom told me. Did your mommy tell you?
B. No, your mommy's wrong. My voice, I don't want it to change. Oh, well.
A. Oh, well, we'll stay little, right?
B. What?
A. We'll stay little.
B. No, I don't want to. I *want* my voice to change. I don't care if it changes.
A. *I* care.

In summary, the children, though not continuously involved with one another, spent a considerable portion of their time in mutual engagement. The next question concerns the extent to which their speech was contingent on the speech or behavior of the partner.

The measure of the level of social speech was the percentage of UUs coded 4, 5, or 0; that is, those UUs that entered into a communicative exchange. The mean percentage of UUs which were coded 4, 5, or 0, that is, formed component events of exchanges, was 59%; SD = 13.2%. Among the older children the means ranged from 48% to 77%; among the younger children, 21%–64%. Single exchanges were by far the most numerous. All dyads, however, produced sequences of three exchanges, and the average number of such exchanges was nine per dyad. Sampling from the longer sequences, table 2 presents the number of dyads who produced sequences of 4, 6, 8, and 12 or more exchanges and the number of such sequences produced by those dyads. A larger proportion of older dyads produced longer sequences, for example, 11/12

TABLE 2

No. Dyads Producing Exchange Sequences of Varying Lengths and Numbers of These Sequences

Length of Sequences (Exchanges)	Younger Dyads (of 6) (*N*)	Total Sequences (*N*)	Older Dyads (of 12) (*N*)	Total Sequences (*N*)
4	5	15	11	42
6	3	7	11	15
8	2	4	7	8
12 or more	1	2	5	6

of the older dyads produced sequences of six exchanges, whereas only 3/6 of the younger dyads did so. Although sequences at the intervening lengths (e.g., 5, 7, 9, etc.) also occurred, the figures presented in table 2 demonstrate that the children were capable of sustaining mutually adapted speech well beyond the simple exchange and also beyond the three-component conversations (which would be comparable to our sequences of two exchanges) described by Piaget (1926, p. 11).

The preceding analyses show that the children spent considerable time in social interaction, much of which consisted of talk. Such talk can be regarded as social behavior and as a means by which children enter into mutual engagement—a state which seems to be intrinsically satisfying to both partners. This is not to deny, of course, that social speech may serve many other functions or that some of these can be realized by nonverbal means. It is the case, however, that children consistently use speech to achieve and maintain contact with each other. The function (of securing attention and engagement) can, of course, be realized by nonverbal means. If a child wishes to initiate interaction with his partner he may wave a toy, make a loud noise, or perform a diverting stunt. Although these acts may attract the attention of his partner, such tactics will not necessarily assure his partner's engagement. There is a more powerful means to secure the involvement of the partner, a technique that illustrates children's use of talk to achieve social contact.

The summons-answer routine as a conversational opener in adult speech was described by Schegloff (1968). This simple but virtually invariable routine follows the following form. Move 1, speaker A summons B. Move 2, B answers: A, "Hey, Fred." B, "Yeah?" Further, with the summons, A commits himself to having something to say, a "reason" for the summons. With the answer, B indicates his availability for interaction. The third move is up to A; he must produce the "reason" for the summons. The routine thus creates a state of mutual committedness and obligation. Correct use of this three-move routine by children would provide substantial evidence of communicative intent and of the ability to use a conventional gambit to secure the involvement of the partner.

The transcripts of all dyads were examined to determine the incidence of this routine. A total of 23 well-formed instances occurred. Almost half of these opened a longer interchange. An example of the well-formed routine is: A. "You know what?" B. "What?" A. "Sometime you can come to my house." Move 1 was typically either an opener, for example, "Guess what," a personal name, or an ascribed role title, for example, "Father." Move 2, which followed on move 1 almost instantaneously, was represented primarily by "What?" Move 3 was, of course, quite varied in form.

Seven aborted routines were identified. In two cases move 2 was missing (B failed to answer). In five cases move 3 was missing (A failed to provide a reason): A. "Mother?" B. "What?" A. (Silence—then A moves away). The number of aborted routines is probably greater than would be expected in adult interaction, for adult speakers would seek to repair or provide an excuse for an imperfect performance of the routine.

Excluded from well-formed routines were cases in which move 1 was an exclamation, which though perhaps an attention getter might be construed as an expression of delight or surprise, and cases in which move 1 contained content in addition to the summons, for example, content anticipating the reason.

In one case a routine appeared to be headed for failure as A did not supply move 3. After a short pause, B repeated move 2. Then, to the surprise of the experimenter (and to the delighted astonishment of B), A played a verbal trick on B. After using the power of the routine to bring B into a position of social contact, A pulled the rug out from under her by supplying a joking insult as the "reason." That the routine was intended was evident from the pleasure that the two children displayed in the joke. Here is the sequence in its entirety:

[A, male, and B, female, are playing independently]:

A. Do you know what?
B. What? [Pause, B turns to A and moves toward him]
What? [Repetition is louder, with broader rising-falling intonation]
A. [A grins, then laughs before speaking] You're a nut.
B. What? What? What's a nut? What? [A and B laugh simultaneously, B dashes threateningly at A, shrieking the final "What?"]

[A and B drift apart after the laughter dies down]

The playful manipulation of the obligation contracted by the routine suggests considerable competence in the use of verbal means to achieve contact.

The successful routines described above have two features in common: first, they create a state of involvement; second, the "reason" always introduced a topic new to the interaction. If the routine is used competently, and if it has desirable consequences, then it seems likely that there would be a tendency to extend its use. A number of sequences, including the excluded cases, resembled well-formed routines but differed from these by occurring within an interchange, thus the element that corresponded to move 1 contained reference to a topic already introduced. In these sequences, move 1 was a request for response containing an indefinite pronoun or interrogative. Move 2 was a simple response in question form. Move 3 supplied the "reason" replacing the indefinite pronoun or interrogative of move 1. This sequence may be viewed as an extension of the summons-answer routine on the following grounds: (1) the form is similar, though move 1 is extended to include a topic; (2) the function is similar (move 1 requests B's attention to what A has to say, while A obligates himself to supply the reason underlying move 1). The major differences are that this variant occurs in the middle of an interaction rather than at the beginning and that the function changes from that of securing involvement to maintaining it. An example of this sequence which we can call the rhetorical gambit follows.

[A, male, B, female, playing together with tool belt]:

A. Hey, do you watch Horrible House?
B. No.
A. I watch Horrible House. Horrible House is silly.
B. O. K., gimme the flashlight. We gotta turn it on. [A hands flashlight to B.]

Move 1-A. Horrible House is funny. Do you know why?

Move 2-B. Why?

Move 3-A. 'Cause Branch always does silly stuff.

B. Unhuh.

[A goes on to describe what Branch does.]

The function of maintaining the partner's wavering attention is clear, for B is more interested in playing with the tools than discussing the television show.

Only five instances of the rhetorical gambit occurred, all of them among the older dyads. Given the small sample, it is improper to claim developmental significance for this finding. If, however, a routine is a series of reciprocal moves in fixed order with a specific function, then variation in the form and context of its use probably depends on learning the formally simpler variety in its basic function and simplest context (i.e., the discourse initial context rather than the embedded context). In adult speech, of course, the summons-answer routine usually serves as a first move in a more complex chain of sustained conversational interaction. Learning to converse may entail, at least in part, learning such short fixed routines, learning to extend or displace features or components of such sequences, and finally learning to combine the sequences and their variants.

Discussion

Our protocols, taken from children's free-play situations, included many instances of private (or egocentric) speech, for example, repetitions, monologuing, and collective monologuing (Piaget 1926), as well as muttering, self-answered questions, and task self-guidance (Kohlberg et al. 1968). The data suggested, however, that the children were mutually engaged the majority of the time and that most of their utterances were mutually responsive, that is, adapted to the speech or nonverbal behavior of their partner. Social speech, as defined here, appeared in abundance, and all dyads were able to sustain mutually responsive speech beyond simple exchanges. Furthermore, both younger and older dyads used a conventionalized series of verbal moves to create a state of mutual involvement. It should be emphasized that the present results do not contradict earlier estimates of the incidence of private or egocentric speech or accounts of its gradual decline. This paper stresses, however, that genuinely *social* behavior does in fact occur between children in the age range of 3½–5 years and that the spontaneous speech of this age may reflect the emergence of the social understandings that underlie such acts as invitations, requests, insults, and excuses. Further exploration of what is said and what is done in children's early social interaction is clearly warranted.

In his study of covert speech in children, Conrad suggested that (at ages 3–5 when overt speech is fluent) "children do not talk silently

to themselves because they have nothing biologically useful to say" (1971, p. 403). This further emphasizes the fundamentally social nature of early speech. From a role-theoretical perspective one can argue that children are sociocentric from birth but lack the skills and talents necessary to interact. As a consequence, interaction among children is initially centered on games concerned with manipulating the physical environment, in what appears to be parallel and egocentric play (cf. Piaget 1964). Language may then develop, as Bruner (1972) suggests, within this context of action and rule-governed play. We would argue, however, that early language serves, not only to coordinate the children's actions, but also to facilitate mutual engagement which has those actions as its focus. As children become able to sustain an interaction per se, play activity becomes less important as a vehicle for promoting these relationships; this development continues until children can interact solely by verbal means. We propose, then, that early activities which promote the acquisition and use of verbal forms of interpersonal contact are biologically useful, for they must precede those later derivations from basic dialogue which become monologue, inner speech or thought, writing, and, finally, adult dialogue.

Piaget found, in his study of children's ideas about causation, that in the period of precausality questions about psychological motives far outnumbered questions about causal explanation. This suggests that learning to detect and interpret verbal cues of intent would be a major skill developed during the period when children practice social communication. It should also be the case that a significant proportion of social communication at this stage is devoted to organizing and structuring the social situation itself. We hope to find further examples of behavior sequences which, with practice, will form the habitual repertoires on which more mature communication behavior depends.

References

Borke, H. Interpersonal perception of young children: egocentrism or empathy? *Developmental Psychology,* 1971, **5**(2), 263–269.

Bruner, J. S. Nature and uses of immaturity. *American Psychologist,* 1972, **27,** 687–708.

Conrad, R. The chronology of the development of covert speech in children. *Developmental Psychology,* 1971, **5**(3), 398–405.

Ferguson, L. R. Origins of social development in infancy. *Merrill-Palmer Quarterly,* 1971, **17** (2), 119–139.

Flavell, J. H.; Botkin, P. T.; Fry, C. L.; Wright, J. W.; & Jarvis, P. E. *The development of role-taking and communication skills in children.* New York: Wiley, 1968.

Goffman, E. *Behavior in public places.* New York: Free Press, 1963.

Kohlberg, L.; Yeager, J.; & Hjertholm, E. Private speech: four studies and a review of theories. *Child Development,* 1968, **39,** 691–736.

Macnamara, J. Cognitive basis of language learning in infants. *Psychological Review,* 1972, **79**(1), 1–13.

Mueller, E. The maintenance of verbal exchanges between young children. *Child Development,* 1972, **43,** 930–938.

Piaget, J. *The language and thought of the child.* New York: Harcourt, Brace, 1926.

Piaget, J. *The moral judgment of the child.* New York: Free Press, 1964.

Sarbin, T. R., & Allen, V. L. Role theory. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology.* Vol. **1.** Reading, Mass.: Addison-Wesley, 1968, 488–567.

Schegloff, E. A. Sequencing in conversational openings. *American Anthropologist,* 1968, **70,** 1075–1095.

Stayton, D.; Hogan, R.; & Ainsworth, M. D. S. Infant obedience and maternal behavior: the origins of socialization reconsidered. *Child Development,* 1971, **42,** 1057–1069.

Vygotsky, L. *Thought and language.* Cambridge, Mass.: M.I.T. Press, 1962.

Winch, P. *The idea of a social science.* London: Routledge & Kegan Paul, 1958.