TALKING, UNDERSTANDING, AND THINKING

Lois Bloom

Departments of Psychology and Speech Pathology, Teachers College, Columbia University, New York, New York 10027

The relationship between understanding and speaking has barely been touched on in language development theory and research. Children’s early speech has clearly received the lion’s share of attention. In contrast, what children understand of what they hear has been virtually ignored, largely because of the difficulties involved in measuring comprehension, not because of a lack of interest. A major problem in evaluating comprehension is that children’s responses are multidetermined—what the child does depends on many things in addition to what he hears. In contrast, child speech can be written down or otherwise recorded and, if nothing else, simply reported or described. However, there has also been an unfortunate tendency to take comprehension for granted, in the light of what little anecdotal and experimental evidence has been reported. The result is that the prevailing view of the developmental relationship between the two is, quite simply, that comprehension necessarily precedes production at every step along the way. However, as awareness about the processes involved in speech development has increased, there has been a growing interest in comprehension development, and deeper questions than whether the one precedes the other have begun to emerge. In particular, one would like to know more about the factors that contribute to understanding messages, how such factors relate to producing messages, and the relation of both to linguistic and cognitive development.

Understanding and speaking do not develop separately, with children learning different “rules” for each. Inasmuch as communication depends upon the extent to which the semantic intention of the speaker matches the semantic interpretation of the listener, the knowledge that each has could not be independent. An intuitively likely hypothesis about the developmental relation between the two is that both speaking and understanding depend upon the same underlying information (or competence), but each manifests a

1 The preparation of this paper was supported in part by Research Grant HD 03828 from the National Institutes of Health. The opinions and ideas as presented here have benefited from discussions and collaborations with Janelleen Huttonlocher, whose studies of comprehension have been particularly enlightening.
different performance mode. The theory of generative transformational grammar has described the linguistic knowledge of adult speaker-hearers as a grammar that can be influenced by different kinds of performance variables. In this view, understanding and speaking would involve learning the same words and linguistic structures, with different performance capabilities emerging at different times. Inasmuch as it is necessary to process words and semantic-syntactic structures in order to learn them, the production of speech might appear to be dependent upon the prior development of comprehension. But, while it is most probably not the case that speaking and understanding are altogether separate developments, it is by no means clear that the emergence of speech and understanding shadow each other.

Another hypothesis that will be proposed here is that the two represent mutually dependent, but different underlying processes, with a resulting shifting of influence between them in the course of language development. In short, it will be suggested that the developmental gap between comprehension and speaking probably varies among different children and at different times and may often be more apparent than real.

Before attempting to deal with the questions of the nature of speech and comprehension and the relation between them in development, it will be helpful to review how their interaction has been described in diaries or case histories of development and in recent psycholinguistic studies of young children. As will be seen, there is most often a strong assumption of a necessary temporal priority in the relation between comprehension and speech development. For example, according to McNeill, "children probably add new information to their linguistic competence mainly by comprehending speech" (1970, p. 102). Observations of small children, by both experimenters and parents, have almost always produced the strong impression that children somehow respond to much more language than they can actually say themselves. However, it also happens that there has always been a certain amount of skepticism in the early diaries about the nature of children's comprehension and exactly what it is in speech events that children are responding to when they appear to understand what is said. More recently, there has been growing skepticism about linguistic comprehension in the psycholinguistic literature as well.

THE COMPREHENSION-PRODUCTION GAP

Several studies have recently demonstrated that infants as young as 2 months of age perceive acoustic differences between sounds (Eimas et al., 1971; and Moffit, 1971) and between different intonation contours at 8 months (Kaplan, 1970). This ability to hear the difference between two sounds (such as [b] and [p]) involves a different set of capacities than is involved in the ability to associate an acoustic event (a word) with some aspect of the environment. However, the one is embedded in the other inasmuch as the
child who recognizes a relationship between a word and an object must necessarily discriminate the word from among other acoustic events that he also hears. Thus, the child must have developed a primitive ability to segment acoustic units within the larger acoustic event of heard speech (which, in turn, has had to be differentiated from nonspeech acoustic events).

First Words

Certain repeated acoustic events begin to stand out in the stream of speech that the child hears towards the end of his 1st year. A number of investigators (for example, Guillaume, 1927; Leopold, 1939; and Lewis, 1951) have emphasized intonation and stress variations as major factors influencing this primitive segmentation. According to Lewis (1951) children respond to intonation before they respond to phonetic form and will respond similarly to adult utterances with different phonetic form if the intonation contour is the same. To explain the child's transition from response on the basis of intonation to response on the basis of phonetic form, Lewis suggested that the child responds affectively both to the intonational pattern of what he hears and to the situation in which he hears it. And at this very same time he hears a phonetic pattern, inextricably intertwined with the intonational pattern and—in many cases—linked expressively or onomatopoeically with the situation. Then his affective response fashions a new whole out of these experiences, this new whole including the intonational pattern, the situation, and the phonetic pattern. When at last the phonetic pattern acquires dominance so that irrespective of the intonational pattern it evokes the appropriate response from the child, we say that he has understood the conventional word (p. 122).

Thus, in the earlier entries [in Lewis' diary account] the child's response to a word is largely affective and conational—consisting of the initiation or inhibition of an act as the result of the affective state aroused by the word; while in the later entries his response is increasingly directed to and concerned with one particular element in his experience—an "object" (p. 146).

Lewis has described the beginning of comprehension in terms of the affective coalescence of intonation contour, phonetic form, and situation into "a new whole," which is, presumably, the primitive mental representation of semantic information linking acoustic linguistic events (intonation and phonetic patterns) with visually perceptual, nonlinguistic (situational) events.

With respect to the relation between comprehension and emerging speech, Lewis noted a 1-month lapse between understanding reference to objects and, not until 17 months, the clear use of words for objective reference, for example "ba" (bath), "ba" (button), "ha" (honey):

It is clear that the changes in the child's responses to words parallel the changes in behavior accompanying his own speech, the latter development following, stage for stage, some months after the former. Along both lines we see a progressive increase in the reference to specific objects within the situation, and ultimately reference even to objects which are not actually present (p. 147).
Lewis did not, however, relate the early influence of intonation for emerging comprehension to the use of intonation in early speech.

Spitz (1957) has described the early development of awareness of prohibitive "no" as the child's first semantic notion. His explanation of its emergence is based upon the exaggerated and often abrupt change in emphasis and affect in mothers' speech that accompanies the stern "no" and head shake as the child starts to crawl about and investigate. Spitz described the beginning of comprehension in the association between a word, "no," and a set of events or behaviors that have been defined for the child by his mother as "prohibited." In this case, the child needs to associate the acoustic event with events that transcend the perceptual properties of different objects. Thus, plants, electric wires, knives, and matches share common reference, even though they look quite different, because they have been associated with "no" and the accompanying prohibiting behavior from an adult. Although prohibitive "no" is often reported in the diary studies to be responded to by children in their 1st year, there have been no reports of prohibitive "no" in children's earliest speech. Indeed, the use of prohibitive "no" develops after the use of "no" to signal the other semantic notions of nonexistence, disappearance, rejection and denial (Bloom, 1970, 1973).

Leopold (1939) reported that the beginning of his daughter Hildegard's comprehension was at 8 months, and was, at first, limited to:

her own name . . . [which] usually induced her to turn her head expectantly toward the speaker. There was no doubt that she referred these sounds in some way to herself. . . . In the second half of the ninth month she took a decisive step forward: both speaking and understanding began [although it was speaking in a very rudimentary sense (p. 21).

Subsequently the word "Daddy" and prohibitive "no, no!" . . . "made her pay attention, stop crying, and look around." Leopold reported that understanding increased rapidly thereafter while progress in speaking was slight. However, the earliest words that were understood: her name, "Daddy," and "no, no!" were not among the first words in Hildegard's subsequent speech.

Some time toward the end of their 1st year, then, children may indicate recognition of an association between an acoustic event and an object by a shift of gaze toward the object, or an arrest of attention. Usually such objects and events are among the most familiar to the child. My daughter, Allison, first recognized the word "birds" in association with the mobile above her dressing table and then, shortly after, the word "music" in association with the record player in her room, before her 1st birthday. Unfortunately, the information that exists about emerging comprehension as children begin to associate sounds and referents is anecdotal in just this way.2 The diary studies (Lewis, 1951; Leopold, 1939; Guillaume, 1927; etc.) reported first instances

2 Janellen Huttenlocher is currently investigating children's ability to recognize and discriminate linguistic stimuli in the first 2 years.
of such behavior but did not report whether or not such responses persisted after they were first observed. For example, Allison's original responses to "birds" and "music" were arrest of attention and shift of gaze toward the objects, but it was not clear if she responded consistently to these words thereafter. She did not recognize the words "birds" and "music" in any situations other than her mobile and her record player for many months. Moreover, neither of the words "birds" or "music" were among the first words that she said, when she began to say words 3 months later. Although speech recognition preceded speech production by 3 months, there was not a one-to-one correspondence between early recognized words and later spoken words.

There have been reports of overinclusion of reference for the first words that children say, where a word is used in situations which seem to share a common element for the child, but not necessarily for the adult (Bloom, 1973; Clark, 1973). For example, Werner (1948) described a child's use of "afta" to designate a drinking glass, a pane of glass, a window, and also the contents of a glass. The child did not understand the word used in each situation before using it himself because there is little likelihood that he heard the word in the same situations, but that did not keep him from using it. It appears that, for comprehension, the child needs to experience a word in each instance in order to understand it. But in speech, as in saying "afta," for instance, the child may not previously have heard the word used in the same context. Thus, it is not the case that production depends on prior comprehension for each instance in which a word is used. Although the child needs to have heard the word in order to say it in the first place, he may well learn to understand the word by learning how to use it—that is, by generalizing or associating properties of the situation in which he first heard the word to new situations.

Even though the first words that the child says are not necessarily the same words that are first understood, there seem to be other kinds of similarities between early production and early comprehension. For one, children respond initially to names of particular objects or persons, and first words also make reference to particular objects or persons (such as "birds" in reference only to birds in a mobile and "music" in reference only to a particular record player). For another, among the earliest words that children say are words like "more" to specify another instance or recurrence of an object or event, or "gone" to comment on the disappearance or nonexistence of objects—that is, children also recognize behaviors which different objects can share. There are different objects that can go "up" or be "up," and different things can be prohibited, etc. The use of the words "afta," "more," "up," and "gone" are similar in reference to the child's early comprehension of prohibitive "no," because the same word is ascribed to referents with different figurative properties. Thus children seem to learn words in reference
to particular objects, such as "Mommy," "bottle," and "blanket," and in reference across perceptually different objects, such as "no," "more," and "gone" for both comprehension and production. More important than any temporal continuum in development between understanding and talking is the fact that both result from the interaction of mutually influential underlying processes.

From Single Words to Sentences

Once children are well into the use of single word utterances in the 2nd year, the question of related development in comprehension assumes a different form. The question in the 2nd year of development is whether or not children process the syntax of sentences that they hear in order to understand sentences, even though they themselves say only one word at a time. The prevailing assumption has been that children do understand a great deal more than they are actually able to say in this period of time, and such presumed understanding is often used as an argument for the claim that single word utterances are holophrastic or "one-word sentences."

As pointed out by Bloom (1973), children between 1 and 2 years of age do appear to understand a great deal that is said to them and often seem to respond appropriately to complex statements and directions. But the directions and statements they respond to so readily refer to their immediate environment more often than not. Utterances in the speech of mothers addressed to children are frequently redundant in relation to context and behavior, and language learning no doubt depends upon the relationship between the speech the child hears and what he sees and does. There are, as yet, no reports of studies that attempt to evaluate whether or not children using single word utterances understand sentences in nonredundant contexts as well as they understand what is said to them about what they can see or hear, or what they do, aside from often repeated and familiar phrases. For example, the statement, "Will is going down the slide," spoken to a child who has been playing with Will on the playground, would be likely to cause her to look for Will or look toward the slide. If children do not understand sentences that refer to relations among objects and events that are not immediately available, then the extent to which they analyze the semantic-syntactic structure for understanding is certainly questionable. Thus, the same statement "Will is going down the slide" or the question "Is Will going down the slide?" would most probably draw a blank if the child hears it at the dinner table or while taking a bath.

Leopold (1939) pointed out that comprehension during the period when speech is limited to single words depends primarily on the child being able to recognize the highly stressed and salient words in an utterance. Adults speaking to children help them understand by repeating key words and
exaggerating stress. Guillame (1927) vividly described the elaborations through gesture, emphasis, and repetition that are used to help very young children understand what is said to them. For example, “Give Grandma a kiss. Grandma. Give her a kiss. Give Grandma a kiss” with pointing and pursing the lips, and maybe even turning the child’s head toward Grandma would most probably produce the desired result. While such a description may seem somewhat extreme, such cues as repetition, exaggeration, pointing, and gesture come to be automatic for the adult, and children no doubt come to look for such cues and use them for getting meaning from the stream of speech that they hear. Several recent studies of mothers’ speech to children have demonstrated that their sentences are shorter, simpler, and more redundant than speech to adults (Broen, 1972; Snow, 1972).

The important issue in relating development in comprehension to development in speaking is the relation between the child’s mental schemas for processing such linguistic and nonlinguistic cues, on the one hand, and the mental processes which result in utterances, on the other hand. The cues of repetition, exaggeration, pointing, and gesture are also present in the child’s own behavior in the 2nd year, but it is not at all clear how such behaviors relate to the child’s perception of such cues produced by others. Children’s gestures have frequently aroused interest (for example, McCarthy, 1954), but there have been no studies of the development of kinesic movement in relation to the development of speech. An important issue that remains to be resolved is how children’s developing gestures relate to the systematic and conventional system of gestures and “body language” used by adults. But, further, one would like to know how the child’s movements augment his speech, and whether the interaction between expressive movements and speech relate to the child’s perception of combined gestural and auditory cues from adults.

One investigation that attempted to tap children’s understanding during the single word utterance period was reported by Shipley, Smith, and Gleitman (1969) using data from two groups of children: one group used only single word utterances, the second group used telegraphic two- and three-word sentences. Children were presented with commands that directed them to act on an object in their immediate presence. The commands were varied as 1) single words, e.g., “ball;” 2) telegraphic, e.g., “throw ball;” or 3) well formed “throw me the ball.” The first finding that was reported was that children who used only single word utterances themselves, responded most often to single word commands, providing no support for the idea that comprehension exceeds production in this period of development. Moreover, the tasks did not evaluate whether or not the children analyzed the structure

---

3 Bambi Schieffelen, at Teachers College, Columbia University, is currently involved in just such an investigation, using the video-taped material in Bloom (1973).
of the sentences, inasmuch as merely touching the ball or picking up the ball was accepted as a positive response.

The second finding reported by Shipley, Smith, and Gleitman (1969) was that the older children who were using two- and three-word utterances preferred to respond to well formed commands rather than to telegraphic or single word commands. However, this second result cannot be taken as evidence that comprehension exceeds production if the well formed commands manifested the same syntactic structure represented in the children's own telegraphic, that is, reduced, utterances. It has been pointed out that early two- and three-word utterances are often reductions of more complete underlying structure (Bloom, 1970). For example, the actual utterances "read book" and "Mommy book" would have the fuller underlying structure Mommy read book given 1) the relevant nonlinguistic state of affairs (Mommy reading, or about to, or supposed to read a book) and 2) evidence elsewhere in a large enough corpus of utterances that the child understands the linguistic relations between agent (of an action) and object (affected by the action).

There seems to be an asymmetry between the child's understanding words and understanding relations between words in the transition from using single word utterances to using longer, structured speech toward the end of the 2nd year. On the one hand, the child needs to understand something of the semantics of a word in order to respond to the word when he hears it spoken by someone else. For example, in order to recognize the association between the acoustic event "milk" and the substance milk that is available and offered to him with regularity in his daily routine, he must have perceived and mentally represented, however primitively, the acoustic configuration of the word. On the other hand, the child does not need to know or to understand the semantic-syntactic relations between words, that is, the structure of sentences, in order to respond when someone talks about the actual relations between Mommy and book, or between Baby and milk, or between ball and floor, when 1) he already understands the words separately, and 2) such objects and relations occur along with the utterances that make reference to them.

Knowledge of semantic constraints and knowledge of syntax are necessary for understanding linguistic messages that do not refer to the contexts in which they occur. In such utterances, the "meaning" is in the linguistic message alone. But when a sentence is redundant with respect to the context in which it occurs, then the amount of information which the child needs to get from the linguistic message is probably minimal. There is, then, another asymmetry between understanding and speaking multiword utterances in that children do not have to process syntax to understand reference to relations among immediate events, but children do need to learn something about the syntax of the language and semantic constraints in order to talk about such relations in any coherent way. Thus, knowing a word and knowing a
grammars, and understanding structured speech and using structured speech, apparently represent different mental capacities. It may be misleading to consider that such capacities develop in linear temporal relation. Such differences between the manifestations of underlying knowledge about aspects of language have not, as yet, been touched on in language development research. They will be explored further, after first considering how the interaction between speaking and understanding has been described in recent psycholinguistic experiments with somewhat older children.

PSYCHOLINGUISTIC TASKS TO COMPARE COMPREHENSION AND PRODUCTION

Several different experimental attempts have been made to investigate the relationship between understanding and speaking in the course of development. Three procedures that will be considered here are: 1) the Imitation-Comprehension-Production (ICP) tasks devised by Fraser, Bellugi, and Brown (1963); 2) training tasks; and 3) the use of elicited imitation.

In the Fraser, Bellugi, and Brown ICP tasks, 3-year-old children were presented with pairs of pictures that portrayed 10 different grammatical relationships such as between subject and direct object. Each pair of pictures presented two contrasting representations of a relation, for example, a girl pushing a boy, and a boy pushing a girl. The investigator presented each pair to the child, saying, “Here are two pictures, one of a boy pushing a girl, and the other of a girl pushing a boy.” In the imitation task the children were asked to repeat one or the other sentence “The boy is pushing the girl” or “The girl is pushing the boy.” In the comprehension task the children were asked to point to the picture that goes with one of the sentences: for example, “Show me the picture of the girl pushing the boy.” In the production task, the children were asked to say a sentence for one of the pictures. The results of the experiment were that imitation was easiest and production was most difficult, leading Fraser, Bellugi, and Brown to conclude that imitation precedes comprehension and comprehension precedes production in the course of language development. Subsequently, Lovell and Dixon (1967) repeated the ICP tasks with 2-year-old children, with the same results and apparent confirmation of the conclusion that children are able to imitate and then understand certain linguistic structures before they use them.

There are several problems in accepting the results of the ICP test. For one, in the production task, the children have already heard the sentence in the introduction to the instructions “Here are two pictures, one of a boy pushing a girl, and the other of a girl pushing a boy,” which gave some support to the child for making up his own sentence. Imitation was the easiest task, but it was also the case that the length of the sentences presented for imitation was always within the children’s auditory memory span. Children have difficulty repeating sentences that exceed memory span even when
they themselves are able to produce the same sentences spontaneously at a different time, as will be seen subsequently in the discussion of elicited imitation.

Fernald (1972) has recently challenged both the methodology and the conclusions of the ICP test. He pointed out that the response possibilities were not equated for the comprehension and production tasks and, in part, favored higher scores for comprehension. In pointing to a picture in the comprehension task, the child could be either right or wrong, depending on which picture he chose, and no other responses or behaviors from the child were considered. However, in the production task, Fraser, Bellugi, and Brown had counted irrelevant responses as errors. Fernald repeated the experiment but equated the response possibilities for both comprehension and production, and when looking at only the correct or incorrect responses in both tasks, comprehension and production were essentially the same. Baird (1972) also questioned the procedures of the ICP tasks. Other research reports have contradicted the evidence of the comprehension-production gap in language development.

Keeney and Wolfe (1972) tested the production, imitation, and comprehension of subject-verb agreement for number in English sentences by 46 3- and 4-year-old children. Production of subject-verb agreement was evaluated in the children's spontaneous speech; 94% of 906 utterances were correct for number agreement; and there were no errors in sentences with the copula "be." More than one-half of the sentences had pronouns rather than nouns as subjects. The authors concluded from this sample that the children reliably inflected verbs for number agreement with subject nouns in their spontaneous speech. In the imitation task, sentence stimuli were either grammatical or ungrammatical (for subject-verb agreement) and did not exceed memory span, for example "The bird is singing." Verbatim responses were scored; 83% of the grammatical sentences were repeated verbatim while only 51% of the ungrammatical sentences were repeated verbatim. However, in 93% of the ungrammatical sentences not repeated verbatim, the children corrected the noun and verb inflection, again demonstrating that they knew the rule for producing subject-verb agreement in sentences.

The results of three comprehension tests in the Keeney and Wolfe experiments were not so clear cut. In one of the comprehension tests, the verbal test, children were presented with either a spoken singular or plural verb and asked to say a subject-noun that agreed in number; results of this test indicated that the children were responding to the verb number inflection in that they most often produced subject-nouns in agreement. In the second and third tests, the children were asked to point to one of a pair of pictures that represented a bird or birds, acting out either a full spoken sentence, in one test, or only the spoken verb of a sentence, in the other (verbal) test. In responding to the full sentences, the children were most often right (appar-
ently responding to the noun number), but in responding to the verb alone they were not. Based on the results of this experiment, Keeney and Wolfe concluded that the children did not understand the relation between verb inflections for number and the meaning of singular or plural, even though they could produce subject-verb agreement for number in their speech: "in reference to the verb number inflection, then, production does indeed precede comprehension."

There are a number of questions that can be asked of the Keeney and Wolfe results—some of which were raised by the authors themselves. They pointed out, for example, that numerosity is a cognitive notion with respect to objects and is effectively signalled by the singular-plural inflection of nouns. The verb inflection is essentially redundant, and moreover, contradicts the noun inflection in form. Thus, a singular noun has no inflection, but plural nouns add [-s]; whereas the agreeing verb adds [-s] for singular nouns and nothing for the plural: "the bird sings" and "the birds sing." The result is that "sings" and "bird" are singular; "sing" and "birds" are plural.

One might indeed question whether the verb inflection actually means singular vs plural apart from the noun (or pronoun) number. Expecting children to recognize its quasi-semantic function may be a metalinguistic task that is beyond their capabilities at 3 and 4 years. The fact that the inflection was fully productive in their spontaneous speech was evidence that they had learned a syntactic rule which they could not identify. Such learning may reflect a certain integrity of the linguistic system as the child is learning it. That is, learning one aspect of the system necessarily entails knowing related aspects, so that knowing plural rules for nouns and speaking sentences will predict the use of verb agreement. In this case, as with lexical items, using a linguistic form may be a means of learning it.

In a recent study by Lahey (in press) there is support for the idea that children may incorporate surface features of the language into their speech before they fully understand the underlying meaning distinctions that such forms represent. Lahey compared 4- and 5-year-old children's comprehension of coordinate, center-embedded relative clause, and right-branching relative clause sentences under four conditions of presentation: 1) with both prosody and syntactic markers, 2) with markers but without prosody, 3) with prosody but without markers, and 4) with neither prosody nor markers. The children were asked to act out such sentences as "the cow hit the pig that chased the deer" with toy animals, and the number of semantic-syntactic relationships that were acted out was scored.

Lahey reported that coordinate sentences were easiest to understand and right-branching sentences were the most difficult, which was in contrast to the report by Menyuk (1969) that nursery school children produced more right-branching sentences than center-embedded sentences. Further, scores for center-embedded sentences were significantly lower when prosody was
eliminated, but syntactic markers were present. Lahey concluded that word order was the major linguistic cue that the children used to process the sentences presented to them, even though they were at the age when children produce many syntactic markers in their spontaneous speech. Some of the children in the study spontaneously reproduced the sentences as they acted them out and either repeated the marker or supplied it when it was missing but, nevertheless, did not demonstrate the relationship that was marked by the relative pronoun.

Two studies have compared the ability of younger children to respond to word order variations in comprehension tasks with their ability to use word order in their own speech. Wetstone and Friedlander (1973) reported that children who used two- and three-word utterances with consistent word order responded equally as well to both normal word order and varied word order in questions and directions presented to them. Chapman and Miller (1973) used object manipulation to compare children’s comprehension and production of subject-object order in semantically reversible sentences with animate and inanimate subjects and objects. They reported that children used appropriate subject-object order in speaking significantly more often than in responding to subject-object cues in a comprehension task.

Of particular interest in the Chapman and Miller study was the result that the two younger groups of children (with mean length of utterance less than 2.5 morphemes) responded correctly at less than chance level to sentences with an inanimate subject and an animate object, indicating that they were responding to the animacy of the nouns more often than they responded to word order to determine the relationship of each noun to the verb. Although Chapman and Miller did not report the corresponding information for production, that is, the extent to which the children used animate or inanimate nouns as sentence subjects, their comprehension result is consistent with other reports that animate nouns predominate as sentence subjects in children’s early sentences (Bloom, 1970; Brown, Cazden, and Bellugi, 1969). It appears that semantic-syntactic knowledge of verbs influences comprehension and production similarly in early development. The semantics of verbs determines the selectional restrictions on nouns as subjects and objects, and the verbs that predominate in early grammar are those verbs that allow reference to people doing things and inanimate objects being acted upon. Eventually, children need to learn 1) that the semantic-syntactic restrictions with such verbs as “hit, chase, bump, push, pull, and carry” (the verbs used in the Chapman and Miller tasks) can be relaxed to specify inanimate agents and effect on animate objects and 2) that word order can override certain selectional restrictions of verbs as a stronger linguistic device. Again, as with lexical items and grammatical morphemes, using word order may be a means of learning more about its grammatical function, as the child actively applies his knowledge to new situations and thereby expands on it.
Training Tasks

In a study by Guess (1969) two severely retarded boys, each 13 years old and with reported mental age of 4.5 years, were trained for the acquisition of the plural morpheme. When trained to understand the forms, the subjects did not generalize to expressive use at the same time. Subsequently, after being trained to use the plural forms correctly, successful reversed receptive training did not generalize to reversed expressive use. Guess concluded that the subjects' ability to learn to understand the plural morpheme was independent of their ability to learn to use it.

In a follow-up study, Guess and Baer (1973) trained four retarded subjects concurrently in both comprehension and production of different allomorphs, either [-s] or [-z], of the plural morpheme. Subjects were trained to understand one plural form and to produce a different plural form. The findings indicated that even when trained concurrently to both understand and use the plural morpheme so that new, untrained stimuli were responded to correctly, three of the four retarded subjects exhibited partial or no generalization of performance from comprehension to production or vice versa with the same plural allomorph. As Guess and Baer pointed out, other studies have used a learning paradigm and have demonstrated the effectiveness of training in one modality (production or comprehension) for facilitating performance in the other, for example, studies of articulation (Winitz and Preisler, 1965; and Mann and Baer, 1971).

The most relevant issue with respect to the relative importance of understanding and using a speech form in the process of learning it has to do with the many variables that interact to affect each. On the one hand, one might well question how it is possible for an individual to use a linguistic structure or form without knowing something about it. However, the important fact is that both understanding and the use of a linguistic form—whether a particular word, a grammatical morpheme, or a syntactic structure—depend on a great many variables, both linguistic and nonlinguistic. The studies by Guess and Guess and Baer compared the effects of specific training for auditory discrimination and expression of a grammatical form, and their results indicated that there was little mutual influence between the two kinds of performance within the contingencies of the reinforcement paradigm. In the experimental situation, the contingencies of reward might be strong enough to preclude any behavior that is not directly reinforced. Even though subjects included new instances within the response category (when they supplied the learned morphemes to novel words), the variables that defined the response behavior were necessarily constrained. It is not clear what the subjects knew that led them to say or respond to the plural morpheme.

In a naturalistic situation, a child might respond to the form when he...
hears it, but what he understands of the form might be heavily dependent on
the situation in which he hears it or on the state of affairs to which it refers.
By the same token, a child's use of a particular form cannot be taken as
unequivocal evidence that he knows the form so that he can understand and
use it in unlimited reference and in any situation. As will be seen, this fact
becomes quite important when one considers children's performance in
elicited imitation tasks.

Elicited Imitation
An important limitation in studies of speech output has to do with the
constraints of topic and context as influences on the grammatical structures
or words that children use. That is, what a child says depends on what he is
talking about. One needs to obtain large samples of speech in order to be
assured of some degree of representation of the child's linguistic knowledge.
But the time involved in collecting and processing large speech samples has
limited the number of children that can be studied in this way, so that other
techniques for tapping children's knowledge are necessary. One such tech­
nique that has received some attention is having children repeat phrases and
sentences. It is possible to prepare sentences that present specific structures
one might want to test for in different groups of children. The underlying
rationale for presenting such sentences to children to imitate has been that
the child will process sentences that exceed auditory memory span according
to what he knows about grammar. As a result, it has been argued, one can tell
how much he understood of the original sentence by looking at how much of
the meaning of the original sentence is preserved in his own repetition of it,
and his repetition will provide evidence of his own coding abilities for
production. The underlying premise in elicited imitation tasks, then, is that if
a sentence is too long for the child to hold in memory he will process the
meaning of the sentence, and his imitation of it, although shorter and inexact,
will provide some evidence of 1) the extent to which he understood it and
2) what he knows about speaking such sentences.

The technique has been used with different research questions in mind.
Menyuk (1963) had children imitate well formed (by adult standards) sen­
tences and utterances produced by children that did not conform to the
model language. Others have compared imitations of standard English and
black English sentences by black speakers (for example, Labov, 1968) or by
black and white speakers (for example, Baratz, 1969). Rodd and Braine
(1971) attempted to assess the ability of 2-year-old children to repeat a
variety of phrase structures.

In order to further explore the relation between spontaneous speaking
and elicited imitation, one of the children studied longitudinally by Bloom
and her associates, Peter, was asked to play a repetition game. Peter had been
identified as an imitator in natural speech inasmuch as an average of one-third
of his different utterances consisted of repetitions of something spoken by someone else (Bloom, Hood, and Lightbown, 1974). Thus, Peter imitated easily in naturalistic situations, and he played the game when presented with sentences to elicit imitation. The following exchanges took place between an adult investigator and Peter, age 32 months, 2 weeks, as they played “Simple Simon says.”

Simple Simon says:  
1. This is a big balloon.  
2. This is broken.  
3. This is broken.  
4. I’m trying to get this cow in there.  
5. I’m gonna get the cow to drink milk.  
6. You made him stand up over there.  

Peter:  
This a big balloon.  
What’s broken?  
That’s broken.  
Cow in here.  
Get the cow to drink milk.  
Stand up there.

These sentences were among a list of 14 sentences presented to Peter. He readily attempted to repeat all but two: “That is not bigger” and “What’s that on your leg?”

Looking at his responses to the sentences 1–6 above, one might conclude a) that Peter probably understands but cannot produce the copula “is,” as in examples 1–3; and b) Peter may or may not understand, but he cannot produce causal connectives such as “gonna,” “trying,” and “made” in examples 4–6. However, all of the 14 sentences presented to Peter for imitation had actually been produced by Peter spontaneously the preceding day, as follows:

Speech event:  
2 and 3. (Peter showing airplane with broken tail to investigator)  
4. (Peter trying to get a colt’s feet to fit into a barrel)  
5. (Peter, holding the cow, going to the toy bag to get barrels)  
(Peter returning with the barrels)  
6. (Peter trying to get investigator to spread an animal’s legs, so that it will stand in a spot he had cleared for it)

Peter:  
This is broken.  
I’m trying to get this cow in there.  
I’m gonna get the cow to drink some milk.  
I’m gonna get the cow to drink some milk.  
You made him stand up over there.

It appears, then, that given the support of contextual events and his own behavior, Peter’s spontaneous speaking ability exceeded his ability to imitate.

---

4 From data collected in collaboration with Patsy Lightbown and Lois Hood. The numbered utterances on the left were presented by an adult investigator for Peter to repeat.
sentences presented to him for imitation. When asked to reproduce sentences that did not relate to immediate context and behavior, he could not produce the very same utterances he himself had produced with such support. In pointing out that their subject also could not reproduce the same sentence she herself had said earlier, Slobin and Welsh (1973) suggested that intention, in addition to context, contributed to the support for the production of messages.

Intention to speak and contextual support from the actual existence of a referent for the utterance are two factors contributing to production performance that are missing in elicited imitation performance and in comprehension tasks. Thus, a child's saying a sentence originates with 1) an internal state of knowing, needing, or intending; plus 2) a referent in context and behavior; plus 3) linguistic knowledge about structure and semantic constraints. In responding to a sentence that he hears in the certain absence of 1), and the possible absence of 2), the child can depend only on 3). Talking and understanding are clearly different behaviors and seem to involve something more than simply a temporal relation in the course of their development.

BETWEEN UNDERSTANDING AND SPEAKING

A disparity between perception and production has been presumed for a long time in speech and psychology as well as in art, and various attempts have been made to understand and explain it (see, for example, Olson and Pagliuca, 1968). Among the most often quoted examples of the difference between perception and production is the fact that children are able to recognize geometric shapes long before they can reproduce them. Adults can recognize and appreciate a work of art but few, indeed, can reproduce one. Some small children can hear the difference between "right" and "white" and yet cannot say the words differently themselves, saying "white wight;" and individuals learning a second language often understand more than they can say.

Among the explanations of the disparity that have been offered, there is reference to memory factors affecting recognition and recall (Mandler, 1967); the knowledge of extra details or attributes that is needed for reproduction as opposed to recognition (Maccoby and Bee, 1965); and the intervening mental reconstruction of an image as the basis for reproduction (Piaget and Inhelder, 1970). Such explanations have not been explicitly extended to young children learning language, but they may contribute to an exploration of the developmental relationship among understanding, talking, and thinking. The memory load for saying a sentence is presumably greater than for understanding, inasmuch as the individual needs to recall the necessary words and their connections to say them, but these linguistic facts are immediately available to him when he hears them spoken by someone else. The child can experience a sentence as more or less independent of its parts, but saying sentences
involves bringing together the elements to form a whole. In recognizing a word or a sentence, a child relates what he hears to existing perceptual schemas, but saying the word or sentence involves reconstructing an intervening representation in the form of a "symbolic image" (Piaget and Inhelder, 1971). In Piaget and Inhelder's view, the latter is more difficult.

It is not clear how representational images relate to either the acoustic signals that the child hears or the speech that he himself produces. Both seem to involve the reconstruction of existing schemas, but, perhaps, at different levels of complexity. It is worth considering that the mental representation that gives rise to an utterance is cognitively less complex than the search for an "existing" perceptual schema that is triggered by hearing speech.

One might consider that in the course of development, both linguistic events (A) and nonlinguistic events (B) provide input that is perceived and organized at (C) in Figure 1. What the child knows of language—that is, that information that allows him to both understand and produce messages—is represented by (C) and depends upon the mental organization and integration of (A) and (B). Whether one speaks of images, concepts, or schemas, there is necessarily some form of mental representation of events. There is not a one-to-one relation between (A) and (B), that is, between linguistic facts, on the one hand, and items and events in the real world, on the other. Rather, there is a necessary internal mental representation of experience, and language is a means for coding or mapping experience.

What are some of the factors that may be involved in producing and understanding messages, in the course of language development, given the basic premise represented by (C) above: that language does not code reality,

![Figure 1](image-url)
but, rather, language codes the individual's intervening mental representation or his experience of reality? The question is an enormously complicated one, and there does not appear to be existing evidence as yet that will lead directly to an answer. At best, what is offered here is a highly tentative and partial account that might lead to the formulation of testable hypotheses for obtaining such evidence.

With respect to producing messages, it has been observed repeatedly (for example, Leopold, 1939; Brown and Bellugi, 1964; Bloom, 1970) that children typically talk about what they are doing or what they can see in the immediate "here and now" in which their utterances occur. Thus, a child might climb on a tricycle and announce "ride trike," or fit a peg into a hole and comment "this fit" or "no fit" (if it doesn't). Child utterances are redundant with respect to the context in which they occur, and to which they also refer, more often than not. Adults, in marked contrast, do not ordinarily talk about what they see or what they are doing when the listener is there to see for himself. In adult discourse, an utterance that bears no relation to the perceptible situation can elicit an appropriate response that also does not relate to what the speakers see or hear or do at the same time.

Egocentricity is often invoked to account for the fact that children seem not to take into account the information that is already available to the listener—or, indeed, to take the listener into account at all. A second point that can be made is that mother's speech to children is probably no less redundant, and indeed, language learning probably depends on the fact that mothers will not typically discuss finances or an argument with a friend when they talk to their children. Rather, when the young child comes into a room with a ball, the typical comment might be "You've got a ball" or "That's your red ball," or "Do you have a ball?" Thus, it might be argued, that children talk in the "here and now" because their parents talk in the "here and now" (to them). However, both the fact that the child is centered on his own actions and perceptions when he produces messages, that is, his egocentricity, and the fact that mothers' speech to children is also redundant with respect to context may well be related to another cause: that young children are strongly dependent on the support of nonlinguistic context for both producing and understanding messages.

It is important to emphasize that it is not the immediate context per se that supports the child's message or that enables him to decode messages, but, rather, it is the child's mental representation of the circumstances and events in nonlinguistic context. Thus, in Figure 1, the nature of the child's mental representation at (C) will determine the extent of his success at producing and understanding messages. Consider Figure 2, in which a linguistic signal (A) encodes an event (B). It is proposed that, in early development, the

---

5 The following diagrams are presented here with full awareness that they are incomplete. They hypothetically schematize some of the observations about the relation...
success of the child's understanding the message in (A) would be dependent upon the extent to which the child understands the relations in (B), with the immediate perceptual representation of the event as the source of the child's information at (C). The linguistic signal may function simply to call attention or focus the child on the event, and extent of understanding would be relatively independent of the complexity of the signal, with limited linguistic processing at LP.

Similarly, in Figure 3, the child is again presented with an event (B), but it is the child who produces the linguistic signal (A) relative to event (B), as output. Here, the success of the child's producing the message depends, again, on the immediate perceptual representation of the event (B) but also on his ability to use a linguistic processor for coding the information in (C). Given the states of affairs represented in Figures 2 and 3, comprehension may well exceed or at least be equal to production depending on the need for and use of a linguistic processing mechanism. In both, the same information about event (B) is available at (C), but linguistic processing appears to be necessary for producing messages about event relationships (multiword utterances), but not necessary for understanding such messages, when these relate to immediately perceived events. But if understanding is greater than talking, it is important to recognize that understanding can be largely nonlinguistic—which, at least at this early stage of development, would seem to close the comprehension-production gap.

This appears to be a plausible account of at least part of the relation between understanding and speaking in the period when children say only

---

between understanding and speaking in language development that have been discussed. They are offered, tentatively, for the purposes of discussion and, hopefully, further elaboration after relevant empirical study.
one word at a time—given the information referred to earlier and other accounts as well (Bloom, 1973). That is, children in the single word utterance period say only one word at a time because they do not yet know a linguistic code for representing information through semantic-syntactic relations between words. The fact that children are limited to single word utterances is evidence that they do not know enough about semantic-syntactic structure to say sentences; how much they know of semantic-syntactic structure for understanding sentences remains to be determined.

What of the relation between understanding and speaking when children are using multiword utterances? The conditions in Figures 2 and 3 continue to prevail; that is, there is still the perceptual representation of immediate events at (C) to support both encoding and decoding. It can now be assumed that the linguistic mechanism is more complex and functions to provide information from the linguistic signal that the child receives, in addition to coding the information that the child puts into messages. Here the relation between comprehension and production is less clear cut and needs to be tested by careful comparative research. The temporal lag at this point in development, if it exists, appears to assume that production itself is evidence of comprehension, but that assumption remains to be verified as well. The results reported by Chapman and Miller (1973) suggested that comprehension during early development of grammar depends upon lexical decoding, with semantic relations between words cued by the relations between objects in event contexts at the same time that children are using knowledge of syntax (word order) to say sentences.

It seems that the major task for the child in the course of language development is the ability to speak and understand messages that are independent of external situations or internal affectual and need states. Language emerges as truly creative in the sense described by Chomsky (1966) the
ability to speak and understand indefinitely many sentences, never spoken or heard before—only through a very gradual process that no doubt continues well into the school years. Thus, language development probably culminates when the child becomes capable of the logical operations described by Piaget (1956), and language becomes for the child truly a means of knowing, at about age 12. But while the achievement of the ability to use a linguistic code relatively independent of context may be the major task of language development in the school years, it certainly has its beginnings in the early preschool years and, further, may never be completely achieved in the adult years, as will be discussed subsequently.

With respect to understanding that is independent of the situation in which it occurs, consider Figure 4. Here, the linguistic signal must be processed at LP, the linguistic processor, first. As the signal is processed, there is no support in the immediate perceptual representation for representing the information in the message. The linguistic signal is the only external source of information about the message. Therefore, information from the LP must be compared with what the child already knows (INFO in cognitive memory) in order to obtain a mental representation of the message at (C) and thereby understand it.

Speaking that is independent of the situation in which it occurs would be schematized as in Figure 5. Here, it is the mental representation at (C) that precipitates the linguistic signal, but the representation originates internally, from the information in the child’s cognitive memory (INFO). Thus, the utterance codes what the child already knows. The LP receives the information for linguistically representing the content of the message, in the form of a mental representation, image, or schema at (C).

Comparing speech in Figure 5, where there is no external event (B), with

![Fig. 4.](image-url)
speech in Figure 3, where there is an external event (B), one can see that where there is an external event to be talked about, message content is more immediately accessible at (C) to the LP for linguistic processing for output. Although there is necessary interaction between (C) and stored INFO in order for the child to know what he sees in (B), the mental representation for the content of the utterance is in the immediate perception. Similarly, one can compare understanding in Figure 4 where there is no event (B), with understanding in Figure 2, where there is an external event (B) that interacts with the message (A).

In this account, one can see how young children's speech occurs most often in relation to immediately perceived events, and children apparently understand speech in relation to immediately perceived events. The diagrams (Figs. 2 and 3) do not explain speech and understanding in the "here and now"; they offer no new information but simply schematize speech and understanding in the "here and now" so that it is possible to see how it is characteristic of early development.

Comparing speaking and understanding, it is not difficult to understand why comprehension in Figure 2, where there is an external event (B), would occur more readily than where there is not an external event, as in comprehension in Figure 4 and speaking in Figure 5. However, the comparison that may be more pertinent to the language development evidence that was discussed earlier is between understanding and speaking in the absence of an external event (B), the comparison between events in Figures 4 and 5. Here, it is possible that understanding may well be less than or equal to speaking, because, in understanding, the mental representation of the content of the message originates with the linguistic signal (A), and the information for mental representation available at (C) is limited, first, to information from
the linguistic processor. However, producing an utterance originates with the mental representation at (C) which provides information to the linguistic processor for semantic representation.

Piaget and Inhelder (1971) argued that the need for constructing the mental image or schema makes production difficult. It would certainly make production in the absence of an event (B) more difficult than production given an external event to talk about. But it also appears that creating the mental representation as input to linguistic encoding may be cognitively less complex than deriving a mental representation as a result of linguistic decoding. In this regard, one can compare children’s utterances that occur in discourse, in response to what someone else says, with utterances that originate spontaneously. In current research in progress (Bloom, Higgins, and Hood, forthcoming), data from four children consist of records of discourse between child and others (investigators and mothers) from the age of approximately 19 months to 3 years. The children produced far fewer utterances in response to what someone else said than they produced spontaneously. Communicative effectiveness was greater for utterances that were not responses to adult utterances, that is utterances that originated with what the child was thinking about, or did, or saw. Further, communicative effectiveness increased developmentally for utterances that were contingent on preceding utterances from an adult. The argument is, essentially, that what an individual knows about an event will determine his ability to both speak and understand messages that code that event and, further, either the event in actual context or the mentally represented event will always be a richer source of information about messages than a linguistic signal alone.

Several experimental studies have demonstrated that what a child knows about an event will influence how he interprets a message about the event. Given the same linguistic message, the child will have greater or lesser difficulty interpreting it depending upon what he already knows about the state of affairs encoded in the message. In a series of experiments, Huttenlocher (1968a, b; 1971) has demonstrated that messages such as “The red truck pushes the green truck” will be responded to with varying delay depending on whether one or the other or neither of the objects is already placed in a three-space track or ladder. Thus, the child’s ability to determine the relationship between two nouns in a sentence is determined in part by how the corresponding objects appear to him in the situation in which the sentence occurs.

Earlier it was said that the major achievement in language development is the ability to use a linguistic code independently of the situational contexts in which utterances occur. In attempting to explain the development away from dependence on context, it is necessary to recognize the continual development of increasingly more complex systems of mental representation of both linguistic and nonlinguistic information. Also, it is clear that even in
adults, context continues to be important. Recent research by Johnson, Bransford, and Solomon (in press) and Bransford and Johnson (1972) has demonstrated how the availability of context and prior knowledge can influence comprehension and recall of linguistic messages by adults.

Bransford and Johnson pointed out that understanding a sentence is "a joint function of input information and prior knowledge." They presented five groups of subjects with a prose passage read to them under the following conditions: 1) no context (simply hearing the passage); 2) context before (seeing an appropriate pictured context before hearing the passage); 3) context after (seeing the picture after hearing the passage); 4) partial context (seeing a pictured context with the relevant items rearranged); and 5) no context, but two presentations of the passage. Subjects then rated the passage on a 7-point scale of comprehension difficulty and then were asked to recall (in writing) the passage as accurately as possible, or at least write down as many ideas as they could. Subjects who received the pictured context before hearing the passage recalled twice as much or more than subjects in any of the other conditions. Their mean comprehension rating was 6.1 (where 7 was the highest rating), while other ratings ranged from 2.3 to 3.7. According to Bransford and Johnson's results, adults understand more from hearing a prose passage when provided with an overt context. Moreover, their subjects reported that they attempted to mentally create situational context when they only heard the prose passage.

It appears that the mental representation of linguistic features of meaning continues to be influenced through adulthood by the interaction of informational, contextual, and pragmatic constraints. The above schematic diagrams are obviously very gross in their representation of the enormously complex interactions that take place with speaking and understanding. The components themselves need to be described and explained, and such explanation should result in considerable refinement of the network of their interaction. However, they do make it possible to see some of the sources of variability in speaking and understanding performance that have been observed. The performance of a particular child will differ to the extent to which learning and maturation have influenced development in any of the component systems: the mechanism for linguistic processing, cognitive memory, and capacities for mental representation of experience.

In summary, the only conclusion that seems appropriate to what has been presented here is that there is not enough information at the present time to explain the relationship between speaking and understanding in language development. At best, the research that has been done is contradictory and inconclusive. However, it is important to emphasize that the relationship is probably never a static one but, rather, shifts and varies according to the experience of the individual child and his developing linguistic and cognitive capacities.
REFERENCES


