do this with such complex and abstract stimuli? Such a question is especially pertinent given that the subjects could not have been aware that their original performance was improved after naming was introduced, because testing of new relationships was not reinforced. An alternative suggestion may be that, as a result of the introduction of naming, certain contingencies became established so that the subjects applied them to the new set of stimuli without having to name these stimuli.

Research on this issue has not been developed sufficiently to provide unequivocal answers to these questions. Without fully exploring the matter, we should be wary of accepting the Horne and Lowe view *tout court* that naming is in itself the paradigmatic case of stimulus classifying behavior to the extent that it obviates the usefulness of a construct of stimulus equivalence.

LISTENER BEHAVIOR AND OSTENSIVE LEARNING

NATHAN STEMMER

BAR-ILAN UNIVERSITY

Three conclusions of Horne and Lowe's impressive article should be mentioned: It acknowledges the essential dependence of speaker behavior on previously acquired listener behavior, it shows that the transfer of listener to speaker behavior is primarily based on echoic behavior, and it argues that stimulus equivalence plays no role in the acquisition of verbal behavior. In my comment, I will concentrate on the first issue, because I largely agree with the last two points (Stemmer, 1995). My account uses some ideas of Catania and Cerutti (1986) and Cerutti (1989).

Because they view the learning of listener behavior as a crucial precursor to the development of linguistic behavior, Horne and Lowe describe it in some detail. They specify the following conditions: (a) The caregiver produces a vocal stimulus in the presence of an object and the child; (b) concurrently, the caregiver teaches the child how to perform conventional behavior in relation to the object; and (c) the caregiver's vocal stimulus in-

creasingly precedes and becomes discriminative for the child's performance of these object-related kinds of conventional behavior. This is the first stage of the process by which children learn object names.

However, a more precise analysis of listener behavior will show that Horne and Lowe's account has a serious shortcoming. It ignores the most important learning processes by which object names (and many other lexical items) are learned: ostensive processes. These processes do not satisfy the three conditions given by Horne and Lowe. On the other hand, the number of object words learned in processes that do satisfy these conditions is extremely low.

Notice first that in the earliest stages, children usually do not learn object names. Rather, they learn action names such as "come here," "drink," "more," "give me," "don't," "drop it," or "bye-bye" (e.g., Benedict, 1979; Dromi, 1987; Smith & Sachs, 1990). To be sure, sometimes the vocal stimuli have two (or more) elements, such as "listen to the music," "go to sleep," "go to the corner," "drink the milk," "walk slowly," "give shoe," or "drop sock," and some of these elements

Address correspondence to Nathan Stemmer, 9 Diskin Street, Jerusalem, Israel.

name objects. But it would be unwarranted to assume that in these very first stages the children learn two different names, one for the action and one for a property of the action or for the object. It is more plausible that the children's responses (e.g., going to a corner, walking slowly, giving a shoe, dropping a sock) are responses to undivided action names. The learning processes in this stage are operant conditioning processes. The responses to the action names are established by a process of differential reinforcement.

In a second stage, and out of the previous instances, emergent listener behavior usually develops, when children start to respond to novel combinations of elements of the undivided action names, as to "drink slowly," "walk to corner," or "drop shoe" (e.g., Goldstein, 1983; Striefel, Wetherby, & Karlan, 1976; for similar results in nonhumans, see, e.g., Herman, Richards, & Wolz, 1984; Savage-Rumbaugh, 1986). This emergent behavior suggests that some of the elements of the undivided action names now have acquired an independent status. For instance, the elements "drink *x*," "walk *x*," and "drop *x*" combine with the elements "slowly," "to corner," and "shoe" to evoke appropriate listener responses. Calling the former *action frames* and the latter *modifiers*, we can attribute the emergent behavior to this process: The modifiers now discriminatively control certain properties of the responses that are controlled by the action frames (cf. Catania & Cerutti, 1986; Cerutti, 1989). For example, the modifier "shoe," which occurs in "drop (the) shoe," determines that the response to "drop *x*" has the property of being the dropping of a shoe rather than something else. (The notion of action frame, which is a particular instance of the notion of *relational frame*, corresponds to Skinner's notion of *partial frame*, 1957, p. 336. The general notion of relational frame applies to expressions with varying numbers of arguments, e.g., "drop *x*," "*x* throws *y*," or "*x* gives *y* to *z*.")

Some of the response properties that are controlled by modifiers involve objects. These modifiers can indeed be called *object names*. But only very few object names are learned in this way. Notice also that at this stage the status of these object names is identical to that of other modifiers. They control prop-

erties of the responses that are controlled by the action frames.

Most object names are learned in so-called *ostensive processes*. In these processes, which may occur concurrently with the previous processes, the child is exposed to the pairing of a vocal stimulus with another (salient) stimulus. Such events are frequently sufficient to establish correct listener responses to the vocal stimulus (e.g., Baldwin, 1991; Nelson & Bonvillian, 1973; Tomasello & Farrar, 1986). For example, the caregiver, while holding a ball, says "ball" (or "this is a ball"). This ostensive event will often establish listener responses to "ball," such as afterwards responding to an utterance of "ball" with looking at a ball or, if "ball" is combined with a known action frame (e.g., "give *x*"), the combination may evoke correct listener responses. In the latter case, "ball" is controlling a response property.

It is not easy to specify the behavioral effects of ostensibly established response-property control. But for typical instances, it can be specified in terms of the stimulus class corresponding to the nonvocal stimulus of the ostensive event (Stemmer, 1992). Thus in the present example, "give ball" will evoke giving an object that is a member of the stimulus class corresponding to the ball of the ostensive event.

Ostensive learning processes have an important feature; they are not based on immediate reinforcement. The first relatively correct response to the ostensibly learned vocal stimulus occurs before the response has been reinforced in the presence of the stimulus. Only after at least one relatively correct listener response to the vocal stimulus has been emitted can reinforcement begin. Consequently, these processes are not operant conditioning processes. Nor do they satisfy condition (b) above; the caregiver does not have to teach the child how to perform conventional behavior in relation to the object.

Ostensive learning plays a crucial role in the early learning of listener behavior. Not only are many, probably most, early object names learned in ostensive events, but other vocal stimuli, such as "thunder" (e.g., by an exposure to the pairing of "Did you hear the thunder?" with thunder), names of smells, names of tastes, names of relations such as the relational frames "*x* is bigger than *y*" or

"*x* throws *y*" (Naigles, 1990; Tomasello & Kruger, 1992) are also learned this way. Again, in most of these cases, the caregiver does not teach the child how to perform conventional behavior in relation to objects. (What could be conventional behavior in relation to, e.g., thunder?) The ostensive event is usually sufficient to afterwards evoke correct listener responses to the vocal stimuli.

Some of the vocal stimuli learned in ostensive events control highly subtle response properties. For instance, a child who has ostensively learned the frame "*x* before *y*" (e.g., by hearing "John will jump before Bill" while observing an appropriate event) will later be able to correctly respond to the instruction "Take the cake out of the oven before you start watching TV." The stimulus "before" controls part of the temporal dimension of the child's (delayed) response. (It has been argued that such delayed responses contradict discriminative interpretations of listener behavior, e.g., Schlinger, 1993. The present analysis shows that this is incorrect; the time intervals derive from the discriminative control of temporal expressions.)

There are thus at least two learning processes involved in the early stages that establish listener behavior: the operant conditioning processes by which children learn action names and modifiers, and ostensive learning processes. What is the nature of the latter processes? Skinner (1957, e.g., p. 357) suggests that it is a variety of Pavlovian conditioning, and Staats (1968) adopts a similar conclusion. Stemmer (e.g., 1973, 1983, 1992) leaves this issue open because of some differences between Pavlovian procedures and ostensive learning. (One difference: after the exposure to Pavlovian pairings, the first stimulus elicits reactions that are unconditioned reactions to the second stimulus, whereas after ostensive pairings, the vocal stimulus evokes responses that are typical relative to the second stimulus, such as looking at a ball or giving a ball.) Still, the striking similarities between the processes suggest the use of ex-

perimental results obtained in the investigation of Pavlovian procedures (e.g., on stimulus generalization) in the analysis of ostensively learned listener behavior.

The present discussion argues that the learning of most object names is not based on the three conditions enumerated by Horne and Lowe. The names are learned in ostensive events, which do not include teaching the child to perform conventional behavior in relation to the object. Moreover, many other lexical items are ostensively learned. Horne and Lowe must therefore introduce fundamental changes in their account of the learning of listener behavior. In particular, they must include an in-depth study of ostensive learning. To be sure, because ostensive learning is not an operant conditioning process, behavior analysts may be reluctant to engage in such a study. But there is no way to avoid it.

The study of ostensive learning processes is also important for what soon becomes the most important listener learning process, namely, intraverbal learning, such as the learning of "amphora" by hearing "An amphora is a Greek vase with two handles" (Skinner, 1957, p. 360). It is likely that intraverbal learning is based on second-order ostensive learning, probably a variety of second-order respondent conditioning (Staats, 1968; Stemmer, 1983, 1992).

Let me close with Horne and Lowe's explanation of the so-called verbal explosion. They suggest that this explosion is largely a consequence of the acquisition of an echoic repertoire. But this can only be a secondary cause. Because speaker behavior is essentially dependent on previously acquired listener behavior, there *must* first be an explosion in listener behavior (implicit or explicit). Ostensive learning accounts for this explosion. Because it does not require the emission of responses and their reinforcement, it allows for the learning of many lexical items in a very short time.