A·B·A·I
Association for Behavior Analysis International

Check for updates

Toward a Behavioral Interpretation of English Grammar

David C. Palmer[1]

Abstract

Behavior analysis is virtually alone among disciplines in assuming that the orderly arrangement of words in sentences, or grammar, arises from exposure to contingencies of reinforcement. In the face of the novelty, subtlety, complexity, and speed of acquisition of verbal behavior, this position will remain difficult to defend until the field can show that a representative range of grammatical phenomena is within reach of its interpretive tools. Using modern English as a case in point, this article points to the important role of automatic reinforcement in language acquisition and suggests that Skinner's concept of autoclitic frames (e.g., *X is taller than Y*) is central to a behavioral interpretation of grammatical phenomena. An enduring puzzle facing this interpretation is how stimulus control can shift from word to word in such frames as one speaks, for such permutations of verbal forms are often novel and rapidly emitted. A possible solution to the puzzle is offered by a consideration of contextual cues, prosodic cues, and the stimulus properties of the roles played by the content words that complete the frames. That these roles have discriminable stimulus properties is supported by considering that in Old English such roles directly controlled case inflections that correspond to positions in autoclitic frames. Continuing to develop behavioral interpretations of grammar is an important pursuit in its own right, whether or not it is sufficient to build bridges to other paradigms.

Keywords Autoclitic frames · Automatic reinforcement · Case inflections · English grammar · Old English · Parity

"Building Bridges," the seedling from which this special issue of *Perspectives on Behavior Science* is a shoot, is a noble attempt by an interdisciplinary group to foster exchanges between behavior analysts and scholars in other fields in the reasonable expectation that all parties would benefit from the cross-fertilization. In the

✉ David C. Palmer
  dcpalmer@smith.edu

1   Department of Psychology, Smith College, Northampton, MA, USA

domain of language, there has been little such cooperation in the past, a circumstance only partly explained by Chomsky's frosty reception of Skinner's analysis of verbal behavior (cf. Chomsky, 1959, 1971; Chomsky & Place, 2000; Virués-Ortega, 2006). Behavior analysts adopt the axiom that language is behavior and should be analyzed as such, but there are few linguists, psycholinguists, cognitive psychologists, or developmental psychologists who take this position. A conspicuous feature of language is that sequences of utterances are typically both orderly and novel. To put it in traditional terms, we often speak in grammatical sentences that we have never uttered before. Whereas behavior analysts seldom discuss grammar at all, most scholars in other fields regard it as a defining feature of language and its central puzzle. Thus, they tend to see a bridge to behavior analysis as a bridge to nowhere.

Grammar is not behavior; rather, it is an abstract model of behavior formulated by scholars. But it is a short step from making such a model to assuming that it somehow "governs" or explains the behavior from which the model was derived, and this seductive assumption is commonly accepted in traditional formulations. In formal writing, one may indeed deliberately consult a grammar or style manual for guidance, but people typically speak fluently in ways that conform to the model without such guidance. This fact suggests that some kind of grammatical system must be deeply rooted in the mind, the biology, or the brain of the speaker. That speech is often riddled with false starts, dysfluencies, clause fragments, and verbal contortions that can be found in no reference book is considered irrelevant. We can get away with such verbal goulash only in familiar contexts with familiar people, who, so to speak, already know what we are trying to say. When listeners demand it, our speech reflects the grammatical conventions of our verbal communities (which may differ from the grammar taught in classrooms), and it is no simple matter to explain how we do so without appealing to the rules themselves as controlling variables. If behavior analysis is to be taken seriously by traditional language scholars, it must address the topic of grammar.

It is natural to assume that parents teach their children to speak grammatically. However, psycholinguists and developmental psychologists have observed that to the extent that parents arrange explicit contingencies for their children's verbal behavior, they tend to care less about the grammar of a child's utterance than its truth. For example, in an influential early study by Brown and Hanlon (1970), parents crowed with delight when a child said, "Mommy not a boy, he a girl"; but they corrected the child for mistaking the day on which a television show would air. If such examples were typical of verbal contingencies for children, it would be hard to explain the acquisition of standard verbal conventions. But Brown and Hanlon considered only explicit consequences that are conventionally assumed to be reinforcing, such as, "Good!" or "That's right." Subsequent work has suggested that parent–child interactions are actually riddled with reinforcement contingencies, if one counts instances of repeating, rephrasing, and responding appropriately to what the child has said (e.g., Bohannon & Stanowicz, 1988; Chouinard & Clark, 2003; Farrar, 1992; Moerk, 1983). Nevertheless, because such surveys have not been sufficiently detailed to account for many of the subtle grammatical distinctions honored in adult speech, such as those discussed below, they have blunted but not turned back criticisms of the position that language is learned like any other complex behavior.

Because children typically learn to conform to grammatical conventions with little, if any, formal instruction, it is tempting to assume that they somehow abstract the grammar of their verbal communities as listeners and then use the grammar to organize their speech, rather as one might learn to use a bulldozer by observing a skilled operator, jotting down some rules, and consulting them later at the controls. But a point forcefully made by Chomsky (1980) is that extracting grammar from examples of speech is a formidable task that challenges the best efforts of erudite scholars; is it credible that nearly every child is able to do so in every culture, in every variety of nurturing or neglectful environment, in a few years? This enigma underlies Chomsky's insistence that the child must be genetically prepared for the task (the "nativist hypothesis"). But notice that this proposal merely transfers the puzzle to evolutionary biology, where it remains as intractable as before: If grammatical conventions are too abstract and arbitrary to be learned, it is difficult to see how they could play a role in differential contingencies of survival (Palmer, 1986/2000).

Chomsky's nativism is still considered uncontroversial by contemporary structural linguists (e.g., Carnie, 2021), but it is not surprising that those who study language development tend to take a more expansive view of the role of experience (e.g., deVilliers & deVilliers, 1999; Goldberg, 2008; Goldberg & Ferreira, 2022; Langacker, 2016; Tomasello, 2003). However, such scholars draw short of endorsing an account rooted solely in basic behavioral processes. Tomasello, for example, argued that the success of theories of usage-based grammar arose from an appeal to social-cognitive skills "that Skinner and the Behaviorists (and Chomsky in his critiques) could never have envisaged" (p. 3), skills that lie in the domains of joint attention, theory of mind, intention-reading, and complex pattern-finding. That those skills might themselves emerge from basic behavioral processes was not explored. Thus the hypothesis that a behavioral view of grammar is possible is not widely considered, and surely the onus is on behavior analysts to show that it should be.

The present document will begin by briefly reviewing some of my previous attempts to provide a behavior analytic interpretation of grammar, in particular, discussions of automatic reinforcement, autoclitic frames, and the function of structure (Donahoe & Palmer, 1994/2004; Palmer, 1996, 1998, 2007, 2014). I will then turn to some suggestions about how to resolve one of the most formidable remaining mysteries, namely, the source of the fluctuations in stimulus control from moment to moment in time as we speak. All of my examples will be in English, but they are intended merely as cases in point. Presumably the analytical tools of the behavior analyst could be deployed across languages, although the details of the account would necessarily differ, perhaps substantially.

A Supporting Example for the Nativist Hypothesis

Evidence for the nativist hypothesis is the profusion of examples of subtle verbal conventions that fluent speakers respect but of which they are typically unaware and that they surely were never explicitly taught. Furthermore, it is sometimes difficult to think of natural contingencies that might have shaped them. Because

something else seems be responsible, the genetic endowment is invoked by default. However implausible or vacuous one might regard this inference, there is no doubt that many grammatical conventions are difficult to explain on the assumption that children are taught to speak. In a previous article (Palmer, 1998), I discussed a particularly compelling and empirically supported example:

> In a study by Gordon (1985), children were shown puppets of various monsters and invited, through modeling, to give them compound names. For example, the experimenter might say, "This monster eats mud; he's a mud-eater. This monster over here eats sand. What kind of monster is he?" whereupon the child, as young as 3 or 4 years, would respond, "A sand-eater." The dialogue might continue:
> This kind of monster eats mice. He's a—
> A mice-eater.
> This monster eats rats. He must be a—
> A rat-eater.
> Most of the children in this study respected this convention: mass nouns, such as mud, sand, or flour, and irregular plural nouns, such as mice, formed compounds unaltered. One might notice the relevance of echoic control here: *eats mud* evoked *mud-eater*; *eats mice* evoked *mice-eater*. However, when the experimenter modeled a regular plural noun, children dropped the *s* and formed compounds with the singular noun. *Eats rats* evoked *rat-eater*. (p. 8)

The force of this example arises from the fact that parents do not teach their children this rule; it is safe to say that nobody, apart from a few linguists, is even aware of it. The nativist alternative is to propose innate features of a mental architecture that process verbal behavior before it is emitted, a kind of workshop that hammers, chisels, and shapes meanings, intentions, and feelings, into verbal form:

> The theory of word structure explains the effect easily. Irregular plurals, because they are quirky, have to be stored in the mental dictionary as roots or stems; they cannot be generated by a rule. Because of this storage, they can be fed into the compounding rule that joins an existing stem to another existing stem to yield a new stem. But regular plurals are not stems stored in the mental dictionary; they are complex words that are assembled on the fly by inflectional rules whenever they are needed. They are put together too late in the root-to-stem-to-word assembly process to be available to the compounding rule, whose inputs can only come out of the dictionary. (Pinker, 1994, p. 146)

Notice that, in this account, the grammar is determined by the architecture of the mind, not by the subject matter, the speaker's history, or any communicative contingency. A feature of behavior analytic interpretations of complex behavior is that they are restricted to those principles that have emerged from the experimental analysis of behavior, so an account that invents such a collection of hypothetical structures, assembly lines, and processes would be dismissed as fanciful. But

what is the alternative? What is the history that enables such a performance, and what are the relevant controlling variables? What principles suggest such delicacy in emitting relevant verbal forms?

A Counterexample to the Nativist Hypothesis

At the time I read Gordon's article, I could not imagine an answer to these questions, but I had at my elbow a daughter of the appropriate age (3 years, 11 months) with whom I attempted a replication, but with some modifications (Palmer, 1998) In particular, I "reinforced" all responses consistent with the rule, in the sense that I said, "Good," or "That's right," immediately after the relevant response. Second, I repeated her response but with a slight modification that violated the rule. Finally, I reverted to conforming to the rule myself. Stripped to its essentials, the dialog went as follows (with her responses in italics):

> This is a monster that eats mud. He is a mud-eater. This monster eats mice. He must be a—
> *A mice-eater.*
> That's right. He's a mice-eater. Now this monster over here eats books. He's a
> *Book-eater.*
> Yes. That's right; he's a books-eater. This monster over here eats chipmunks. He must be—
> *A chipmunk-eater.*
> Right. He's a chipmunks-eater. This one eats marbles—
> *He's a marble-eater.*
> Yes. He's a marbles-eater. How about this one. He eats candles. He's a—
> *A candles-eater.*
> Right. He's a candle-eater. This one eats spiders. What's he?
> *A spiders-eater.*
> Good. He's a spider-eater. Now. . . .
> *You say "spider-eater," but I say "spiders-eater!"* (Palmer, 1998, p. 14)

This simple demonstration, one that I replicated several years later when my other daughter came of age, confirmed the results of Gordon (1985): both children, without any guidance or prompting in that context, dropped the final *s* of regular plurals when forming a novel compound noun. I was surprised that children of that age should have had sufficient experience with compound nouns to guide their behavior when coining a novel exemplar. Moreover, putative reinforcers like "Good!" or "That's right," although undoubtedly effective in shaping behavior in many other contexts, served no function in determining the form of the construction under investigation. This is consistent with the finding of Brown and Hanlon (1970) that, in their corpus of parent–child interactions, explicit expressions of approval or disapproval for grammatical form were rare enough to be dismissed as important controlling variables for subtle grammatical constructions. Like Pinker (1994) one is

tempted to take refuge in a nativist theory about the structure of the mind, however circular or vacuous it might be.

However, the speed with which the children picked up a novel and inconsistent ad hoc grammatical rule, dashes our hopes for such a theory. What becomes of the "root-to-stem-to-word assembly process" and its relation to the "mental dictionary" and the "compounding rule?" Nativism assumes a genetically fixed mental architecture that solves the learning problem for the child, but if the child can learn a novel construction with exposure to a few examples, perhaps there is no learning problem.

The demonstration suggests at least two lines of inquiry for behavior analysts: automatic shaping of verbal forms and the acquisition of intraverbal "frames," or in Skinner's taxonomy, "autoclitic frames" (Skinner, 1957). The former addresses the speed with which children acquire complex verbal forms in the apparent absence of explicit shaping, and the latter suggests an interpretation of the orderliness to be found in extended strings of responses, or in traditional terms, phrases, clauses, and sentences.

Automatic Shaping

Newborn infants have no idea, so to speak, in what century or in what culture they are born, but they must quickly adapt to the particular menu of contingencies facing them. The surest and fastest way to do so is to follow the example of others. When the stimulus properties of their own behavior "match" the stimulus properties of the behavior of others, past or present, the behavior is relatively likely to be reinforced, and matching can become a conditioned reinforcer (Donahoe & Palmer, 1994/2004, 2022; Palmer, 1996; Sundberg et al., 1996; Skinner, 1957; Vaughan & Michael, 1982). Discriminating such a match need not depend on instruction or differential feedback from other people: If the patterns of behavior of others become discriminative stimuli for a child, then those same stimuli, arising from the behavior of the child, will likewise be evocative; the evocation of relevant discriminative responses will certify that the child has "matched." In common parlance, the child can recognize successful imitations, and in appropriate contexts, such imitations will be reinforcing. The term "automatic reinforcement" is meant to capture the fact that the reinforcement is both immediate and unmediated by others. The term "automatic shaping" is meant to capture the fact that behavior can proceed from one form to a quite different form through a series of steps, each automatically reinforced, in the absence of a teacher, sibling, or parent.

The terms "matching," "modeling," and "imitation" are somewhat misleading. Elsewhere I have suggested the term "achieving parity" to capture the fact that to be automatically reinforcing, one's behavior need not have the same form as another's, so long as it has the same, or comparable, behavioral effect (Palmer, 1996). A pianist who plays a passage modeled by a bassoonist does not imitate the bassoonist, either in form or in response product, but the effect on a listener (and on the pianist) is in some sense "the same." That is, the listener tacts them as "the same" despite dramatic differences in response form and tonal quality, because a common discriminative response is evoked in each case ("recognizing the tune"). The behavior of members of a jazz trio goes far beyond

matching, but the response products are nevertheless automatically reinforcing to the musicians, and a discordant response product would be jarring.

More to the point of the present topic, a person who has never before played a piano can eventually pick out the bassoonist's tune through automatic shaping, i.e., successive approximations to the target passage. Note that the naïve pianist must be able to "recognize the tune" in order for such automatic shaping to occur. If the bassoonist plays a long sequence of random notes, such shaping will not occur, for the correct note in a sequence will not evoke the discriminative responses that certify that the tune is proceeding correctly. Notice also that the response product need not be "pleasant" to be reinforcing. A discordant note will be reinforcing if it conforms to the model. (See Palmer, 1998, for further discussion of this topic.)

Extrapolating to verbal behavior, note that children become discriminating listeners before they become articulate speakers. In other words, "they know the tune" of many verbal expressions before they begin to speak. Furthermore, neglecting slight differences in latency and tonal quality, speech has the distinctive feature that it stimulates the speaker at the same time and with the same stimulus properties as the listener. Thus when "picking out the tune"   of a novel verbal response, automatic differential reinforcement or punishment of a speaker's behavior will be immediate and reliable.

The scope of automatic reinforcement in shaping new response forms is vastly greater than that of social reinforcement, because virtually every utterance is a "learning trial." Thus, it must be considered in any investigation of the acquisition of complex verbal behavior by children. (For an overview of the role of automatic shaping in the acquisition of the passive construction, see Palmer, 2014. For reviews of the general acquisition of conditioned reinforcers in young children, see Cló & Dounavi, 2020; Lepper & Petursdottir, 2017; Lepper et al., 2013; Petursdottir & Lepper, 2015; see also Greer & Ross, 2008; Greer & Singer-Dudek, 2008; and Vandbakk et al., 2019.)

Explaining Novelty and Regularity in Sequences of Verbal Responses

Automatic shaping helps explain how a child might acquire new words or phrases in the absence of explicit instruction, provided those words or phrases have become discriminative stimuli, but extended sequences of utterances are characteristically both novel and orderly. Novelty usually appears in the arrangement of verbal elements, whereas orderliness is revealed in the fact that only some arrangements are effective on an audience. Even the most banal remarks—*I told Larry you'd pick him up at five outside the Hadley Walmart*—are likely to be novel; even so, only a few arrangements of the various terms in the utterance would lead to an appropriate response by a listener. Scrambling the word order would commonly yield unintelligible gibberish.

The Sentence as a Unit of Analysis

In linguistic analyses, an utterance is said to be a *well-formed sentence* if the elements are arranged in a sequence that is compatible with a set of rules that have been extracted by scholars by examining the speech and writing of native

speakers in the relevant verbal community. As mentioned above, some linguists believe these rules to be more than convenient descriptions; rather, they are taken to be part of the fundamental nature of the human mind, and for that reason they are of primary interest. The purity of these fundamental influences on individual speakers is assumed to be contaminated by confusion, loss of attention, memory lapses, and distraction. Thus, verbal behavior as it is actually emitted is viewed as a kind of palimpsest: often one needs to peer beneath the surface to detect the underlying order. It is the difficulty of this detective work that makes definitive conclusions about grammar so elusive (Chomsky, 1965, 1975, 1980; Pinker, 1994). This view has the additional advantage of protecting the linguistic enterprise from the obvious criticism that actual speech often violates the proposed grammatical rules.

In contrast, a behavior analyst assumes that verbal behavior as it is actually emitted is the subject matter of interest and that it can be understood by invoking well-established principles of nonverbal behavior, as well as by conducting controlled research on verbal behavior itself. Because the concept of grammar is derived, not from individual words, but from patterns of words, the task of a behavior analyst includes accounting for such patterns. This task must accommodate the fact that, although a pattern might have been encountered before, the words that make up a particular instance might never have occurred in that pattern. For example, when we say, *Steve is bringing lasagna tonight,* the pattern is familiar, but the particulars are likely to be novel. A young child who has heard such an utterance might be able to subsequently emit closer and closer approximations to repeating it through automatic shaping, like picking out a familiar tune on a toy xylophone, but a much more complex account is required if the child says, for the first time, *Stacey is bringing ice cream tonight.* To put it loosely, the child must have learned the pattern as well as how to insert appropriate words into the pattern, presumably without having been given explicit instruction in how to do so.

The sentence is not a technical term in Skinner's taxonomy. He cited the many functions that sentences serve and the many fluctuating variables that influence their emission, and he concluded, "It is scarcely worth dignifying the result of all these activities with a special name which might be taken to imply a single process" (Skinner, 1957, p. 354). Nevertheless, the sentence is not merely a formal term with meaning only to grammarians: a long stream of verbal behavior can typically be divided into segments, each of which has a kind of unity and a characteristic effect on listeners. The words within such segments "hang together" to achieve that effect, whereas other divisions of the verbal stream would likely yield nonsense. Trying to read a text from which the punctuation and capitalization have been removed makes the point forcefully: with few exceptions, we would have to discover the segment boundaries in order to make any sense of the passage. Therefore, some extended segments of verbal behavior have "behavioral reality" and must be included in our analysis.

The effect of such strings on the listener is a plausible defining feature of a useful unit of analysis, and as with other behavioral units, the effect would have to be determined empirically rather than defined formally. Verbal behavior commonly evokes effective action in listeners, or it can condition the behavior of the

listener with respect to an object, arrangement, or state of affairs. For example, a simple declarative sentence, such as *Bill's pet snake is venomous,* conditions the behavior of the listener with respect to the subject of the sentence, an effect that might become evident only much later when the listener gives Bill's snake a wide berth. As a first approximation, I offer these effects as defining features of a "behavioral sentence."

By this definition, the subject matter of interest is any verbal response that is effective in controlling the behavior of listeners in a systematic way, a category that excludes many strings that a linguist would regard as well-formed but that are too long or too complex to affect listeners in a systematic way. For example, a survey of listeners would lead us to dismiss the following string, which is offered by Chomsky and Miller (1963) as "a perfectly well-formed sentence with a clear and unambiguous meaning:"

> Anyone who feels that if so many students whom we haven't actually admitted are sitting in on the course than ones we have that the room had to be changed, then probably auditors will have to be excluded, is likely to agree that the curriculum needs revision. (p. 286)

This grotesque string was obviously composed as an exercise, not to affect a listener in a systematic way, but to illustrate how a complex arrangement of terms might conform to a set of rules. A problem-solving assault on this passage might make some sense of it eventually, but the magnitude and nature of that task excludes it from the generalization that my proposed definition attempts to capture. The formal structure of a sequence of words is of no interest unless it evokes a characteristics response in one's verbal community.

On the other hand, my proposed functional definition embraces many communicative behaviors that lack the required features of the sentence as traditionally defined. Pointing, gestures, simple phrases ("On the shelf"), clauses with no anchor ("Not if I can help it") and muddled forms ("Is it still raining outside yet, or what?") are all to be found abundantly in everyday speech, and they are usually effective. Our task is to make sense of verbal behavior as it is actually emitted, not to validate formal models or classroom definitions.

Pointing, gestures, phrases, and irregular fragments are often sufficient when both speaker and listener are in the same place, have similar histories, share motivational variables, and are discussing something at hand. Because the variables that control the behavior of the listener include both the speaker's behavior and many contextual cues, highly elliptical remarks may be sufficient. To *listen* is to speak along with the speaker (Palmer, 2007; Schlinger, 2008a, b; Skinner, 1957), and in such circumstances, the listener is likely to already have some tendency to say what the speaker is saying. But a listener may be on a telephone, may know nothing of the speaker's circumstances, may share no relevant history, and may know little of the topic at hand. In such cases the verbal behavior itself must carry all the burden of evoking an effective response. Under such conditions, the speaker is likely to speak in the grammatically complete sentences of linguistic theory. Thus the challenge of explaining verbal behavior that is both novel and orderly remains.

Autoclitic Frames

Skinner (1957) coined the term "autoclitic frame" to capture the concept of a verbal pattern into which circumstantial variable terms are inserted. We can define such a frame as a skeleton of fixed verbal responses, controlled by some variable common to all instances, intermingled with responses that vary according to circumstances. Prepositional phrases, such as *on the X, inside the X,* and *behind the X* are elementary examples, and it is a relatively small step to simple declarative sentences with two variables: *X is on the Y; X is inside the Y; X is behind the Y.* Such cases can be smoothly accommodated into a behavior analytic account and provide a toehold into the domain of more complex examples.

The verbal behavior and verbal environments of young children abound in autoclitic frames, such as *Sing X again; Give me X, Where's the X? X is bigger than Y; X is going to Y; Put X in Y; X gave the Y to Z.* Each frame represents frequently encountered contingencies that vary only in particulars from one occasion to another. In standard grammatical terms, both verbs and prepositions relate one or more terms to another set of terms in a systematic way, and autoclitic frames accommodate these relationships. It's the recurring relationship that calls for a frame on the part of the speaker, and the frame establishes for the listener how the terms in the frame are to be interpreted.

What verbal element is part of a frame and what is a variable term can't be determined by form alone: we must analyze controlling variables. In frequently recurring contingencies, long utterances might be strung together under the control of a single variable into an intraverbal chain. The spiel of a telemarketer or streetcorner evangelist might consist of long and invariant passages like the memorized lines of an actor. The stimulus control of such intraverbal chains poses no special empirical or interpretive challenges. In contrast, autoclitic frames invariably entail rapid, oscillating shifts in stimulus control between frame and variable terms. Sorting out these shifts is a formidable task (Palmer, 2007, 2014).

Just as a completed jigsaw puzzle has a visual effect that is qualitatively different from that of the disconnected pieces, a completed autoclitic frame has an effect on a listener that is qualitatively different from that of the isolated variable terms or the same terms in a different arrangement. We respond one way to *The pig licked Sarah's foot* and in a quite different way to *Sarah licked the pig's foot.* Thus, even a small vocabulary of terms can yield innumerable examples and permutations of a given frame, each with a unique effect on a listener.

Listeners are typically sensitive to multiple concurrent contingencies and reinforce verbal constructions accordingly. Thus, frames can be, and often are, nested in other frames. *X* in the above examples could be *an X, that X, some X, Y's X, all Xs, the X inside the Y, the X that did such-and-such*, etc. Nouns are often nested in phrases with their associated adjectives and articles, such as *a bright blue X,* or *the small square X.* Whether a term is part of a nested frame or part of the larger frame is an empirical matter that cannot be determined by form alone. In the *Happy Birthday* song, for example, there is only one embracing frame, and only one variable, the celebrant's name.

For expository purposes, it is helpful to start with relatively simple examples. In what follows, I will take the first step of trying to analyze unnested declarative

frames in the present tense, and I will treat variables as single, unmodified terms. Explaining verbal accounts of past events is a more formidable problem, but the extra difficulty is unrelated to grammar.

As noted above, terms traditionally called verbs commonly entail autoclitic frames, for they typically specify relationships among terms. With support from a context, when subjects, objects, locations, etc. are implicit, a verb can stand alone (*Go!*) or can appear with a single additional term (*I am listening*), but in the absence of such support the frame becomes explicit: *X is going to Y,* and *X is listening to Y.*

More formidable interpretive problems are raised by more complex relationships in which there is little contextual support: *Send,* for example, relates a sender, a receiver, and the thing sent. *Pay* relates a payer, a receiver, and an amount or account paid. In modern English, such relationships are specified to the listener by a term's position in a relevant autoclitic frame. In *X is donating Y to Z,* X is the giver, whereas Y is the object given, and Z is the person or agency receiving the donation. In the passive construction, the frame might be *X is being donated to Y by Z,* in which case the roles of X, Y, and Z would be scrambled accordingly. As noted, our task is to explain how such strings can be emitted fluently and rapidly with novel permutations of variable terms. Linguists have formulated models of such strings, but their models serve only a descriptive function, not an explanatory one. A behavioral interpretation aims for more, but it requires a moment-to-moment account of the shifts in stimulus control as such strings are emitted.

A Stimulus Control Interpretation of Fluent Speech

To the extent that autoclitic frames are behavioral units, they are clearly relevant to the task of explaining grammatical regularities in speech, for the frames provide an invariant structure on which variable terms hang. But we must face the task of explaining how variable terms are interwoven with such frames. As Skinner put it, the filling in of an autoclitic frame "is not simply the emission of two responses separately acquired. The process [entails] multiple causation. The relational aspects of the situation strengthen a frame, and specific features of the situation strengthen the responses fitted into it" (Skinner, 1957, p. 336). To a first approximation, this sounds correct: in the presence of Tom passing a bill to Roy, the frame *X is passing Y to Z* would come to strength because the relational aspects of the situation have occurred before, and the frame has been emitted and reinforced. Then the variables, Tom, the bill, and Roy are "fitted into it," and our interpretive task is complete.

But is it? In what order should the variable terms be inserted? What controls our behavior, moment to moment in time as we speak, so that the variable terms are correctly interwoven with the frame? Why do we not say, "Is passing to the bill, Roy, Tom?" It is not enough to say that the latter makes no sense, or that it would mean something different if we said it differently. Nor can we appeal directly to a history of reinforcement for emitting the terms in that order: recall that it is putatively a novel event. What are the variables that evoke saying "Tom" at the moment we do, rather than earlier or later, and what variables then evoke jumping to the frame, rather than to the other names? In fluent verbal behavior, stimulus control shifts

rapidly from one term to the next. A complete account must specify these controlling variables and how they shift from moment to moment to evoke the next term in the sequence. I will next offer an inventory of such variables and suggest how they might come together.

To begin, the relational context of one man slipping a bill to another will, as Skinner says, tend to evoke the frame, *X is passing the Y to Z.* (We will assume that, for incidental reasons, this frame is more probable than the passive construction, as well as the innumerable other verbal responses that an observer might make to the scene.) The presence of a bill will tend to evoke *bill,* and the presence of Tom and Roy will tend to evoke their names. In all four cases, we appeal to tact control. Next, notice that the completed frame has prosodic features. In autoclitic frames, prosodic stress is placed on the variable terms, *X, Y,* and *Z,* whereas the terms making up the frame itself are typically prosodically recessive. The pattern is superimposed on any prosodic variation within words, giving the completed frame a characteristic rhythm, in this case: *DA da dada da DA da DA.* This rhythmic pattern will exert some intraverbal control over terms as they are emitted. Furthermore, the frames themselves consist of fixed terms, and because they have presumably participated in repeated contingencies of reinforcement, these elements will hang together partly under intraverbal control. Thus, once begun, there will be both rhythmic and phonemic intraverbal control between successive terms: *is* is partly controlled by an initial stressed term; *is* evokes *passing the; passing the,* followed by a stressed term, evokes *to; to* evokes a final stressed term.

Tact control and intraverbal control thus carry some of the explanatory burden of accounting for the emission of completed autoclitic frames, but at least one additional source of control is necessary: *Tom, bill* and *Roy* must take their places and not be mixed up. *Tom* is the actor, the agent, the "subject," and he must come first in this frame; the *bill* is the thing being transferred, hence the "object," and it must come second; *Roy* is the guy getting stuck with the bill, hence the "indirect object," and he must come third. If the roles played by the several terms could be adduced as physical controlling variables, we would have a complete account. Our completed autoclitic frame would be: *subject is passing the object to the indirect object,* with unique sources of control as each term is emitted.

But is it plausible that "roles" have distinctive stimulus properties that can control word order? Surely not as physical objects: The generic roles of subject, object, and indirect object can be played by an indefinite number of people, animals, organizations, furniture, asteroids, etc. among which it would appear to be hopeless to find a distinctive stimulus property. But role-assignment in verbal behavior is controlled, not only by characteristic features of observed events, but also by characteristic responses of speakers themselves to those events. We interpret the world—that is, we respond differentially to events according to our respective histories—and our verbal behavior is multiply controlled by the world and our responses to it. A second observer, with a different history, might interpret the scene as Roy snatching the bill from Tom, in which case their grammatical roles would be reversed. A third party, who was present but "daydreaming" (i.e., looking at, but not responding discriminatively to the scene), might be wholly unable to assign roles to the transaction and would have nothing relevant to say. However difficult it may be to specify the

controlling variables in any given case, a consideration of the evolution of the English language suggests that roles as interpreted by speakers do indeed have discriminable properties sufficient to systematically control verbal behavior.

Modern English assigns roles to positions in autoclitic frames, and as a result word order in the relevant constructions is tightly constrained, but in Old English roles were directly tacted in the form of case inflections and case-specific definite articles. Grammatical subjects were inflected with nominative case markers, direct objects with accusative markers, indirect objects with dative markers, and possession with genitive markers. For example, consider the different roles played by the noun phrase *the king* in the following sentences. The corresponding phrase in Old English is shown on the right (Quirk & Wrenn, 1955):

The king is sending his emissary to the Pope.	*se cyning*
The Pope is sending the king to the Holy Land.	*þone cyning*
The Pope is sending his emissary to the king.	*þæm cyninge*

In modern English, the form of the phrase is unchanged, but its role is signaled to the listener by its position in the autoclitic frame. In Old English, the respective roles were signaled by case inflections. The position of a term, at least relative to modern English, would have been somewhat free to vary, as it formerly was in Latin and Ancient Greek, and as it currently is in Russian, Finnish, and many other languages. In its evolution from Old English to modern English, most case markers were dropped in favor of positions in autoclitic frames.

Because case markers are *tacts* of grammatical functions, those functions must have discriminable properties. Those properties are therefore available as controlling variables in the completion of autoclitic frames. That is, not only do the variables in a frame have position and prosody, they also have associated roles, or cases. To return to our restaurant scene, the frame previously expressed in terms of grammatical categories can be replaced with the discriminable properties of the events filling the variable terms: *X (as passer) is passing Y (as the thing passed) to Z (as receiver of the thing passed).* Listeners acquiring such a frame emit the relevant terms as echoics in the presence of the discriminable properties of the roles played by the terms, among other discriminative stimuli. When they speak at a later time those properties may exert control at the appropriate times for the completion of the frame.

To put it in concrete terms, Tom, the bill, and Roy are observed by the speaker playing their relevant roles, and the stimulus properties of those roles control their order of emission in the frame. The first term, "Tom," is a multiply controlled tact in response to Tom and his role as agent. The second term, "is passing," is emitted partly under intraverbal control of the prosodic properties (i.e., stress and duration) of "Tom" and partly under tact control of the relationship between Tom and other variables. The third term, "the bill," is partly under intraverbal prosodic control of "is passing," partly under tact control of the bill, and partly under tact control of its role as "object passed." The fourth term, "to," is part of the frame and is a multiply controlled intraverbal under phonemic control of "is passing" and prosodic control of a stressed response, "bill." The final term, "Roy," is partly under intraverbal prosodic control of "to" and multiple tact control by Roy himself as well as his role

as receiver. In short, there are sufficient controlling variables, shifting moment to moment in time, to account for the complete verbal response. With practice, there need be no deliberate composition or mediating behavior on the part of the speaker: one's behavior can be directly controlled by the relevant variables.

The fact that speakers of many languages, ancient and modern, can tact grammatical roles, as shown by their facility with case inflections and other grammatical devices, is sufficient evidence that such roles can serve as controlling variables in normal speech. Nevertheless, it is not obvious precisely what those controlling variables are. As noted above, the roles of "subject," "direct object," "indirect object," "agent," etc. appear to be able to be filled by an indefinitely large number of things, apparently with no stimulus properties in common.

But this way of putting it overstates the problem. Like other natural categories (*dog, flatfish, shrub*) a class of verbal responses need not have a necessary and sufficient defining feature. For example, the role of "subject" in a sentence frame can have one or more of several properties. It is often the topic of primary interest to the audience: To the homeowner, we say, *Your house is on fire*; to the fire chief we say, *The fire is on the corner of East and Main;* to the arson investigator we say, *Sam's ex-wife lit the fire.* Second, the role of subject can be the first term in a temporal sequence with the verb spread over the time interval: *Tom passed the bill to Roy; Sarah sent the book to her sister; Bob climbed to the summit on the Spur Trail.* Third, the subject can be the agent of an action: *Tom laughed at Roy's pun; Tom pushed Roy out of the path of the car; Tom sorted the files alphabetically.* A natural category need not exemplify every relevant feature.

Moreover, unlike the linguist, whose analysis must embrace every conceivable verbal string, even if successive words are drawn from a jar, the behavior analyst must explain only verbal behavior that actually occurs. People acquire new autoclitic frames one at a time, and it is likely that they generalize relevant variable terms along objective gradients only according to their histories. Having just learned to say, *Harry is turning the doorknob,* a speaker might on a later occasion say, *Michael is turning the doorknob* or *Harry is turning the dial*, for Harry and Michael, and knobs and dials, have features in common. In contrast the speaker would be highly unlikely to say, *Enthusiasm is turning the doorknob* or *Harry is turning the water vapor,* though a linguist might confirm that these are indeed grammatical sentences. In the frame *X is turning Y,* X might best be conceived of as "thing that can grasp and manipulate objects," rather than "subject," "agent," or "actor." With increasing experience, the class of terms that could play a role in the frame would increase accordingly, but lawfully, and not without constraints. For example, we would not necessarily expect the speaker to generalize to *Harry is turning the page* without further experience, for the respective actions are quite different in topography. The point is that speakers cannot just say anything; constrained by behavioral processes, they can only say what their histories and current conditions cause them to say, and that includes case inflections and terms in autoclitic frames. This does not relieve us of the responsibility of specifying the actual controlling variables in any instance, but we should not accept as an axiom that autoclitic frames can be completed by an indefinite number of exemplars with no features in common. To do so would be to confuse actual verbal behavior with a formal model of language. Except under

laboratory conditions, we may be unable to specify the controlling variables for novel strings of verbal behavior, but there is no reason to doubt that the conceptual tools of the behavior analyst are, at least in principle, adequate to the task.

For present purposes, this completes our behavioral account of some elementary examples of the interweaving of variable terms and autoclitic frames in modern English. That account, in turn, suggests an approach to interpreting other grammatical regularities in behavioral terms. A complete account of the grammatical regularities in even a single language would be a Herculean task. The purpose of the present discussion is to identify relevant analytic tools. To the extent that these tools are adequate, we can dismiss one of the most compelling criticisms of the overall behavioral account of language. If they are not yet adequate, we have at least narrowed the gap in what remains to be done.

Conclusion

Science routinely attempts to make sense of everyday phenomena, under conditions in which experimental control is impossible. Why does thistledown float in the air? Why did one brother die from the flu, whereas his twin survived it? Where will the next hurricane make landfall? Science does so by extrapolating well established physical, chemical, meteorological, or biological principles to the available data. Suggested answers are not facts; they are possible explanations that are consistent with these principles, and as a result, they tend to allay curiosity and blunt the human impulse to invent mystical explanations for puzzling phenomena.

Language is behavior, and it is to the science of behavior we should look for possible explanations of its mysteries. Grammatical regularities in verbal behavior are particularly challenging, for it is apparent that they are not typically shaped by a process of successive approximations like exceptionally forceful lever-presses in a laboratory rat, nor chained together by a corresponding sequence of discriminative stimuli. The novelty and variability from case to case preclude such inferences. To date, behavior analysts have shown relatively little interest in studying grammar (but see Dal Ben & Goyos, 2019; Greer & Ross, 2008; Kohler & Malott, 2014; Østvik et al., 2012; Wright, 2006). This apparent neglect may be due to the difficulty of operationally defining the subject matter, or perhaps to the great variability in grammar from one verbal community to another: Any facts about grammatical behavior in one language might seem to be of limited generality and of narrow interest. However, grammar is special in that it seems to defy behavioral principles. Its importance as a subject matter lies, not so much in the details of a particular behavioral account in a particular language, but in testing whether such an account is possible with existing behavioral concepts and principles.

The present account is an incremental advance over previous speculations on the topic, for it adduces evidence from the evolution of English that the roles played by those things about which we speak plausibly serve as discriminative stimuli controlling the ordering of terms in autoclitic frames. Evidence from many other inflected languages could have served as well, but the history of English confirms the claim

that case inflections and autoclitic frames serve the same function, for the former were eventually dropped in favor of the latter. We can only speculate why the change occurred, but one possibility is that verbal operants necessarily occur in sequence; in inflected languages, order in the sequence is a wasted independent variable. Specifying relationships with autoclitic frames may therefore be simpler for the speaker, however troublesome they have been for the behavior analyst in search of the controlling variables in speech.

I am skeptical that any behavioral interpretation of grammatical regularities in verbal behavior will appeal to scholars from other disciplines. The view that language is behavior, and the assumption that it can be understood in the same terms as other behavior, is profoundly different from the that of other paradigms. But the goal of science is not to persuade others of one's point of view, right or wrong; it is to find order in nature.

Declarations

References

Bohannon, J. N., & Stanowicz, L. (1988). The issue of negative evidence: Adult responses to children's language errors. *Developmental Psychology, 24*(5), 684–689. https://doi.org/10.1037/0012-1649.24.5.684

Brown, R., & Hanlon, C. (1970). Derivational complexity and order of acquisition in child speech. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 11–53). Wiley.

Carnie, A. (2021). *Syntax: A generative introduction* (4th ed.). Wiley-Blackwell.

Chomsky, N. (1959). Review of *Verbal Behavior* by B. F. Skinner. *Language, 35*(1), 26–58. https://www.jstor.org/stable/i217081

Chomsky, N. (1965). *Aspects of a theory of syntax*. MIT Press.

Chomsky, N. (1971). The case against B. F. Skinner. *The New York Review of Books, 17*, 18–24.

Chomsky, N. (1975). *Reflections on language*. Pantheon.

Chomsky, N. (1980). *Rules and representations*. Columbia University Press.

Chomsky, N., & Miller, G. A. (1963). Introduction to the formal analysis of natural languages. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical philosophy* (Vol. II, pp. 269–321). Wiley.

Chomsky, N. & Place, U. T. (2000). The Chomsky-Place correspondence, 1993–1994. *Analysis of Verbal Behavior, 17*, 7–38. https://doi.org/10.1007/BF03392953

Chouinard, M. M., & Clark, E. V. (2003). Adult reformulation of child errors as negative evidence. *Journal of Child Language, 30*(3), 637–669. https://doi.org/10.1017/S0305000903005701

Cló, E., & Dounavi, K. (2020). A systematic review of behaviour analytic processes and procedures for conditioning reinforcers among individuals with autism, developmental or intellectual disability. *European Journal of Behavior Analysis, 21*(1), 292–327. https://doi.org/10.1080/15021149.2020.1847953

Dal Ben, R., & Goyos, C. (2019). Further evidence of automatic reinforcement effects on verbal form. *Analysis of Verbal Behavior, 35*(1), 74–84. https://doi.org/10.1007/s40616-018-0104-3

deVilliers, J. G., & deVilliers, P. A. (1999). Language development. In M. H. Bornstein & M. E. Lamb (Eds.), *Developmental psychology: An advanced textbook* (4th ed.). Lawrence Erlbaum Associates.

Donahoe, J. W., & Palmer, D. C. (2004). *Learning and complex behavior*. Ledgetop Publishing. (Original work published in 1994).

Donahoe, J. W., & Palmer, D. C. (2022). Acquired reinforcement: Implications for autism. *American Psychologist, 77*(3), 439–452. https://doi.org/10.1037/amp0000970

Farrar, M. J. (1992). Negative evidence and grammatical morpheme acquisition. *Developmental Psychology, 28*(1), 90–98. https://doi.org/10.1037/0012-1649.28.1.90

Goldberg, A. E. (2008). Universal grammar? Or prerequisites for natural language? *Behavioral & Brain Sciences, 31*(5), 522–523. https://doi.org/10.1017/s0140525x0800513x

Goldberg, A. E., & Ferreira, F. (2022). Good-enough language production. *Trends in Cognitive Sciences, 26*(4), 300–311. https://doi.org/10.1016/j.tics.2022.01.005

Gordon, P. (1985). Level-ordering in lexical development. *Cognition, 21*(2), 73–93. https://doi.org/10.1016/0010-0277(85)90046-0

Greer, R. D., & Ross, D. E. (2008). *Verbal behavior analysis: Inducing and expanding complex communication in children severe language delays*. Allyn & Bacon.

Greer, R. D., & Singer-Dudek, J. (2008). The emergence of conditioned reinforcement from observation. *Journal of the Experimental Analysis of Behavior, 89*(1), 15–29. https://doi.org/10.1901/jeab.2008.89-15

Kohler, K. T., & Malott, R. W. (2014). Matrix training and verbal generativity in children with autism. *Analysis of Verbal Behavior, 30*(2), 170–177. https://doi.org/10.1007/s40616-014-0016-9

Langacker, R. W. (2016). Working toward a synthesis. *Cognitive Linguistics, 27*(4), 465–477. https://doi.org/10.1515/cog-2016-0004

Lepper, T. L., & Petursdottir, A. I. (2017). Effects of response-contingent stimulus pairing on vocalizations of nonverbal children with autism. *Journal of Applied Behavior Analysis, 50*(4), 756–774. https://doi.org/10.1002/jaba.415

Lepper, T. L., Petursdottir, A. I., & Esch, B. E. (2013). Effects of operant discrimination training on the vocalizations of nonverbal children with autism. *Journal of Applied Behavior Analysis, 46*(3), 656–661. https://doi.org/10.1002/jaba.55

Moerk, E. L. (1983). A behavioral analysis of controversial topics in first language acquisition: Reinforcements, corrections, modeling, input frequencies, and the three-term contingency. *Journal of Psycholinguistic Research, 12*(2), 129–155. https://doi.org/10.1007/BF01067408

Østvik, L., Eikeseth, S., & Klintwall, L. (2012). Grammatical constructions in typical developing children: Effects of explicit reinforcement, automatic reinforcement, and parity. *Analysis of Verbal Behavior, 28*(1), 73–82. https://doi.org/10.1007/BF03393108

Palmer, D. C. (1996). Achieving parity: The role of automatic reinforcement. *Journal of the Experimental Analysis of Behavior, 65*(1), 289–290. https://doi.org/10.1901/jeab.1996.65-289

Palmer, D. C. (1998). The speaker as listener: The interpretation of structural regularities in verbal behavior. *Analysis of Verbal Behavior, 15*, 3–16. https://doi.org/10.1007/BF03392920

Palmer, D. C. (2000). Chomsky's nativism: A critical review. *Analysis of Verbal Behavior*, *17*(1)*,* 39–50. https://doi.org/10.1007/BF03392954 (Original work published in 1986)

Palmer, D. C. (2007). Verbal behavior: What is the function of structure? *European Journal of Behavior Analysis, 8*(2), 161–175. https://doi.org/10.1080/15021149.2007.11434280

Palmer, D. C. (2014). Verbal behavior. In F. K. McSweeney & E. S. Murphy (Eds.), *Handbook of operant and classical conditioning* (pp. 369–391). Wiley.

Petursdottir, A. I., & Lepper, T. L. (2015). Inducing novel vocalizations by conditioning speech sounds as reinforcers. *Behavior Analysis in Practice, 8*(2), 223–232. https://doi.org/10.1007/s40617-015-0088-6

Pinker, S. (1994). *The language instinct*. William Morrow.

Quirk, R., & Wrenn, C. L. (1955). *An old English grammar*. Methuen.

Schlinger, H. (2008). Listening is behaving verbally. *The Behavior Analyst, 31*(1), 145–161. https://doi.org/10.1007/BF03392168

Schlinger, H. D. (2008). Conditioning the behavior of the listener. *International Journal of Psychology & Psychological Therapy, 8*(3), 309–322.

Skinner, B. F. (1957). *Verbal behavior*. Appleton-Century-Crofts.

Sundberg, M. L., Michael, J. L., Partington, J. W., & Sundberg, C. A. (1996). The role of automatic reinforcement in early language acquisition. *Analysis of Verbal Behavior, 13*(1), 21–37. https://doi.org/10.1007/BF03392904

Tomasello, M. (2003). *Constructing a language*. Harvard University Press.

Vandbakk, M., Olaff, H. S., & Holth, P. (2019). Conditioned reinforcement: The effectiveness of stimulus—Stimulus pairing and operant discrimination procedures. *The Psychological Record, 69*(1), 67–81. https://doi.org/10.1007/s40732-018-0318-8

Vaughan, M. E., & Michael, J. L. (1982). Automatic reinforcement: An important but ignored concept. *Behaviorism, 10*, 217–227.

Virués-Ortega, J. (2006). The case against B. F. Skinner 45 years later: An encounter with N. Chomsky. *The Behavior Analyst, 29*(2), 243–251. https://doi.org/10.1007/BF03392133

Wright, A. N. (2006). The role of modeling and automatic reinforcement in the construction of the passive voice. *Analysis of Verbal Behavior, 22*, 153–169. https://doi.org/10.1007/BF03393036