Constituency and Bonobo Comprehension

Robert Truswell, Center for Cognitive Studies, Tufts University

Thanks to Ray Jackendoff, Tim O’Donnell, Sarah O’Neill and Moira Yip, and to audiences at Tufts, Groningen, York and Essex, for comments on earlier versions of this material.

The author is now at the University of Edinburgh. Correspondence concerning this article should be addressed to: Robert Truswell, Linguistics and English Language, School of Philosophy, Psychology, and Language Sciences, The University of Edinburgh, Dugald Stewart Building, 3 Charles St, Edinburgh EH8 9AD, UK. Email: rtruswel@staffmail.ed.ac.uk.

Constituency and Bonobo Comprehension

Abstract

This paper presents new evidence concerning the level of syntactic competence of Kanzi, a language-trained bonobo. Although his responses suggest that he can make use of information concerning linear order in interpreting spoken English commands directed at him, specific deficits in his ability to interpret function words and structures involving coordinated noun phrases suggest that Kanzi has a model of English syntax that does not include hierarchical constituent structure. This evidence provides a new source of support for the conclusions of Fitch & Hauser (2004) concerning limitations on pattern recognition abilities in nonhuman primates.

Keywords: Phrase structure, ape language experiments, Kanzi, NP-coordination, language evolution.

1 Introduction

An almost-undisputed finding from the earliest, foundational work on generative grammar concerns the indispensability of *constituency*. Humans appear to organise their grammars in such a way that smaller units can be recursively built into larger phrase-structural units, which can then be manipulated without worrying about their internal complexity. One classic example of this, from Chomsky (1975), is that the auxiliary which is fronted in English yes/no question formation is not the first auxiliary in the string, or the second, or last, or any other position defined in purely linear terms, but rather the first auxiliary *after the subject noun phrase*, regardless of how many auxiliaries that noun phrase contains. Of course, such a generalization cannot be formulated without the notion of *noun phrase*, which, in turn, requires reference to a hierarchically organized phrase structure.

(1) a. *Is [the man who __ tall] is in the room?

b. Is [the man who is tall] __ in the room? (Chomsky, 1975, p. 31)

An interesting question, recently posed by Fitch & Hauser (2004), concerns the extent to which the same is true of the cognitive abilities of other primate species. However, finding an unequivocal answer to that question has not been straightforward. This paper provides a new source of evidence for Fitch & Hauser's claim that any syntactic competence of nonhuman primates, unlike that of humans, is not based on constituency. The evidence concerns a deficit in the linguistic comprehension of Kanzi, an otherwise extremely linguistically gifted bonobo. A corpus of 660 utterances directed at Kanzi, together with the actions that he performed in response, was presented in Savage-Rumbaugh et al. (1993). Moreover, a parallel set of utterances was directed at Alia, a human infant, and her responses were also recorded. This provides a very useful resource for comparisons between the linguistic ability of an ape and a child.

Overall, Kanzi is very impressive, and marginally outperforms Alia in many respects. However, in a couple of areas, and most clearly with regard to one set of stimuli, Kanzi's performance plummets. These stimuli require Kanzi to perform some action on multiple objects described by a coordinate noun phrase. In these cases, Kanzi's syntactic comprehension was arguably at chance.

So why do we find the deficit specifically here? I will argue that, although Kanzi is exposed to a wide range of stimuli, and shows clear evidence of a basic type of syntactic competence, namely sensitivity to word order, Kanzi's performance dips whenever interpreting the input correctly requires sensitivity to constituency. In other words, Kanzi's syntactic competence, at least on the evidence of this corpus, is accurately described by the hypothesis that he can learn patterns related to word order, but falters with more complex constructions which can only be interpreted correctly with reference to constituent structure.

The paper is structured as follows: we start by giving some background on Kanzi and his general syntactic abilities. Three main types of sentence will be discussed: those that Kanzi responds accurately to, and which provide clear evidence for the presence of some syntactic ability (section 2); those that Kanzi responds accurately to, but plausibly without making use of syntactic information (section 3); and finally, those that Kanzi surprisingly fails to respond accurately to, providing evidence for the *absence* of some syntactic ability (section 4). Taking all these together, we see that the evidence for sensitivity to word order is strong, but most putative evidence for sensitivity to phrase structure is circumstantial and inconclusive. Moreover, the only clear evidence concerning phrase structure is the NP-coordination data, which points to an inability to handle this aspect of syntax. Finally, section 5 concludes, and sketches the implications for a research program aimed at situating this finding within a broader evolutionary and developmental cognitive approach.

2 Kanzi's Linguistic Achievements

By now, the story of Kanzi is well known (see in particular Savage-Rumbaugh et al. 1993; Savage-Rumbaugh & Lewin 1994; and Savage-Rumbaugh, Shanker & Taylor 1998). He attended sessions designed to teach an English-like communication system to his adoptive mother, Matata, but picked up the system himself, without explicit instruction. He quickly (for a bonobo) acquired a lexicon of several hundred items, and showed a high degree of comprehension of English utterances, either spoken or encoded on a special keyboard. Moreover, anecdotal evidence suggests that he shows more intuitively human-like patterns of language use, for example “talking” to himself by signing on the keyboard when there are no humans around, as opposed to the more reward-driven patterns of communication evident in most other English-based primate communication systems.

Nevertheless, the response of theoretical linguists to Kanzi’s story has generally been one of skepticism. Although Kanzi is streets ahead of most other apes trained in this area, he still clearly underperforms relative to a human of equivalent or younger age, despite the greater focus on the development of Kanzi’s linguistic abilities than those of a typical child. Moreover, his development did not continue as we might have hoped: there is no evidence for the development of a fully-fledged syntax and semantics of the sort that develops almost automatically during a typical human childhood. That is sufficient, in many eyes, to dismiss the Kanzi project as at best tangential to the core business of understanding the human language faculty.

I believe that this is a mistake. Kanzi clearly understands (and, in a sense, speaks) some English, in the same way that I speak and understand some German. And like my understanding of German, Kanzi’s understanding of English is far from perfect. However, we will see below that Kanzi fails to understand English in ways that are quite unlike adult human patterns of error. They are quite obscure at first, and this makes it tempting to dismiss them as blips, as rules that Kanzi failed to spot. But the simplicity of the rules in question (from an adult human perspective), combined with the apparent complexity of some of the rules that Kanzi does master, makes this analysis quite implausible.

To appreciate this, we need to know what Kanzi *is* capable of, linguistically. This is where the corpus presented in Savage-Rumbaugh et al. (1993) is such a valuable resource. This corpus contains 660 novel, unpredictable English instructions given to Kanzi, together with his responses. The basic format is as in (2): we start with an item number (I reproduce this in lieu of a full citation — all items can be found in Savage-Rumbaugh et al. 1993, pp. 111–210). Next, we have a code for the correctness of Kanzi's response: C and C1–5 are correct, but with increasing amounts of hesitation or human intervention; PC, OE and other codes are incorrect in various ways. Next, in italics, is the sentence uttered. Finally, in parentheses, we have a description of Kanzi's actions (which I will sometimes truncate to remove unnecessary details) and an explanation of the code assigned by Savage-Rumbaugh et al., if necessary.

(2) a. 329. (C) *Put the rubber band on your ball.* (Kanzi does so.)

b. 444. (C) *Put the rock in the water.* (Kanzi does so.)

c. 287. (C) *Kanzi, take the tomato to the colony room.* (Kanzi makes a sound like "orange"; he then takes both the tomato and the orange to the colony room.) [C is scored because it is assumed that Kanzi is announcing that he wants to take an orange and have it to eat.]

Novel though they may be, these activity descriptions are mainly built up of fairly predictable parts, namely simple actions such as *giving*, *putting*, *showing*, *taking* and *getting*, and a large class of nominal expressions. Kanzi is judged as having correctly performed an action if he either takes the correct object to the named location, as in (2c), or arranges the two objects such that they instantiate a fairly predictable relation. Objects are generally given or shown *to* animate beings, placed *in* receptacles, bodies of liquid, and so on; and *on* objects that they couldn't be placed *in*. If an object is present in Kanzi's immediate vicinity, it can be taken *to* a location. If not, it can be retrieved *from* a location.

These relatively simple guidelines account for 420 out of 660, or 64%, of the instruction–action pairs described in Savage-Rumbaugh et al. (1993). This could lead to the suspicion that Kanzi has acquired an impressive lexicon, and has a degree of pragmatic competence, in that he can construe instructions as instructions and figure out likely appropriate responses based on an array of default actions and the properties of the named objects and events, but that he is lacking any syntactic and/or combinatorial semantic ability. If this were true (and, given the clear evidence for Kanzi's lexical ability, it represents a sort of null hypothesis), it would constitute a version of what Anderson (2004) calls a "semantic soup" strategy: the meanings of the individual content words are pushed together into a coherent action in whatever way they fit best, without any attention to syntactic information in the signal.

However, there is evidence that Kanzi moves beyond semantic soup in responding to such instructions. Sometimes, particularly in the case of ditransitive sentences, the experimenter's instructions underdetermine the response that would be expected on the basis of the semantic soup hypothesis. Nonetheless, Kanzi is capable of understanding such instructions, even when they describe quite novel actions, and even when there are no contextual clues as to which argument should bear which thematic role. For example, in (3), either action is equally plausible on the basis of the semantics of the individual lexical items, but Kanzi still responds appropriately in each case.

(3) a. 525. (C) *Put the tomato in the oil.* (Kanzi does so.)

b. 528. (C) *Put some oil in the tomato.* (Kanzi picks up the liquid Baby Magic oil and pours it in a bowl with the tomato.)

There are 43 sentences presented in such alternations in the corpus (21 pairs, with one sentence repeated — see Savage-Rumbaugh et al. 1993, pp. 95–6 for details), and, as in (3), Kanzi generally performs at least as accurately in these cases as in cases like (2), where lexical semantics and familiarity with the experimental conditions alone would yield a good guess at the target response (specifically, he responds accurately to sentences like those in (3) in 33 out of 43 trials, or 77%, compared to an overall accuracy across the corpus, reported by Savage-Rumbaugh et al. 1993, p. 77, of 72%). This suggests that basic word order sensitivity is not a problem for Kanzi.

A slightly more involved argument supporting the same conclusion can be made on the basis of verb–particle constructions, as in (4).

(4) a. 317. (C) *Put your collar on.* (Kanzi does so.)

b. 430. (C1) *Turn the flashlight on.* (Kanzi puts a rock in his mouth, then turns on the light switch by the T-room door.) [Cl is scored because in the past the word *flashlight* has often been used to refer to any room light with Kanzi.]

The syntactic challenge here is that, on the basis of Kanzi's knowledge of the basic meanings of *put*, *your*, *collar*, and *on*, the semantic soup hypothesis might predict that (4a) asks Kanzi to put his collar on *something*. However, in (4a), such a literal response would be inappropriate: when putting a collar on, we don't put it on *anything*, but rather, round something's neck. Things get even worse in the case of (4b): the semantic soup hypothesis might predict that Kanzi would interpret this as an instruction to rotate the flashlight, while it was on top of some other object (indeed, Alia, the human control in this experiment, took this as an instruction to put the flashlight on the shelf, suggesting something akin to this confusion). Instead, Kanzi correctly (given his understanding of *flashlight*) flipped a light switch in response to this request. Although relevant data are quite sparse, Kanzi responded correctly in 10 of the 12 items (83%) involving particles or intransitive prepositions, again suggesting that these pose no particular problem for him.

Finally, note that there is no common factor between *put on* and *turn on* that can be attributed to the meaning of *on*. That is, the meaning of *V on* is not fully determined by the regular meaning of V and the regular, spatial meaning of *on*. The sensitivity to word order in verb–particle constructions therefore comes in two varieties: firstly, Kanzi knows to interpret the preposition *on* differently, depending on whether or not it is followed by a noun phrase; and secondly, he knows to interpret it differently, relative to the locally preceding verb. This gives further evidence that Kanzi displays a sensitivity to certain types of basic syntactic information.

One obvious hypothesis can therefore be dismissed: Kanzi's English comprehension is dependent on word order, and so, in one sense, on syntax. This rules out the possibility that bonobo comprehension of English is based on a heuristic cobbling together of word meanings floating in an unstructured semantic soup. Of course, though, once we attribute *some* syntactic competence to Kanzi, the question becomes *how much* competence we have to credit him with. This will be the focus of the next two sections.

3 Some Non-evidence for Phrase Structure

Having established that Kanzi is able to make use of word order information in computing the semantic content of an utterance, I next want to discuss, and ultimately dismiss as irrelevant, Kanzi's performance on a range of modification and complementation constructions, including adjectival modification, locatives, and sentential embedding. In a human grammar of English, the mapping between syntax and semantics in these cases makes crucial use of constituency: the noun that an adjective or relative clause modifies is, roughly, the one that heads that modifier's sister constituent. Competence with such constructions would therefore appear, initially at least, to suggest some ability on Kanzi's part to deal with constituent structure.

Kanzi does respond appropriately to many examples involving such structures. (5) shows a pair of cases in which he successfully distinguishes *hot water* from *ice water*.[1]

(5) a. 563. (C) *Show me the hot water.* (Kanzi vocalizes.) *Can you show me the hot water?* (Kanzi makes a sound like "hot," then picks up a paint brush, [and] points to the hot water with it, [. . .])

b. 564. (C) *Can you pour the ice water in the potty? Pour the ice water in the potty.* (Kanzi picks up the bowl of ice water and heads toward the potty.) E [the experimenter] says, "That's right." (Kanzi pours the ice water carefully into the potty.)

Similarly, the following cases show his ability to deal with locative modifiers: Kanzi heads outdoors or to the bedroom as appropriate to the stimulus.

(6) a. 500. (C) *Get the lighter that's in the bedroom.* (Kanzi does so.)

b. 508. (C) *Go get the lighter that's outdoors*. (Kanzi goes to the play yard, picks up the lighter, and [. . . eventually] brings the lighter back.)

However, it must be noted that the evidence for constituency is somewhat weak in the case of locatives. Specifically, Kanzi only ever has to deal with one locative in each instruction: he is not asked to put the lighter that's outdoors on the rock in the bedroom, for example.[2] It is plausible, then, to suggest that the locative functions as a sentential modifier, rather than a noun phrase modifier, for Kanzi: Kanzi's interpretation of (6a) is something like "Get the lighter (the bedroom is involved)." Such a possibility is consistent with every locative relative clause in the Savage-Rumbaugh et al. corpus — note that more demanding types of relative clause, with a gap in object or adjunct position, were not tested.

In fact, a similar analytical move can be made even in the case of the attributive adjectives that Kanzi successfully interprets. This idea is less intuitive, because there is no sentence-level equivalent of *big*, or *hot*, in a human grammar of English. However, that is not sufficient to show that Kanzi also perceives *hot water* as a complex syntactic unit. The alternative works as follows: Kanzi interprets *hot* as "something involved in this action is hot." Then there is only one choice for the relevant locus of hotness in the action described by *Show me the hot water*: the showing can't be hot, and the speaker and agent are probably no hotter than anyone else, so the water is the only pragmatically reasonable candidate for hotness. The fact that *hot* is intended to modify *water* can therefore be inferred on pragmatic grounds, without any need to assume that Kanzi has posited a *syntactic* relationship between the two words.

Of course, it is entirely possible to distinguish empirically the two positions reported here. On a model with hierarchically structured constituency, cases with multiple modifiers will not be significantly harder to comprehend than cases with single modifiers, because the modification relations are syntactically encoded. On a constituency-free model with pragmatic interpretation of modifiers, however, a significant indeterminacy would be introduced. The crucial sentences, then, are structured along the following lines.

(7) a. Pour the hot water in the ice water.

b. Put the rock that's outdoors on the TV that's in the bedroom.

Pouring the hot water in the ice water is just as plausible an action as pouring the ice water in the hot water. If Kanzi has a model of English syntax including determination of modification relations through phrase structure, such sentences would be disambiguated. However, if Kanzi's English grammar does not make reference to constituency, instead attaching modifiers purely on the basis of heuristics related to world knowledge, we expect him to perceive an ambiguity in such sentences, and for this ambiguity to lead to an indeterminacy in his responses to such sentences.[3] Unfortunately, there are no relevant documented examples that I'm aware of, either in Savage-Rumbaugh et al. (1993) or elsewhere in Savage-Rumbaugh's writings.

A second possible test to check for phrase structure in Kanzi's grammar would be to investigate his performance on *un*grammatical sentences. If there is no particular syntactic link between an adjective and a noun, it is possible (though not necessary) that Kanzi would respond equally well to a case in which the adjective is separated from the noun, as in (8).

(8) a. *Hot show me the water.

b. *That's outdoors go get the rock.

To a human English speaker, such sentences are probably ultimately interpretable, but much less automatically, or reliably, than a fully grammatical sentence with the regular English word order. With Kanzi, however, bets are off: there is a possibility that Kanzi would respond equally accurately to such ungrammatical sentences. However, Kanzi is only systematically tested on grammatical sentences, which rules out another potential source of information concerning his model of English syntax. In the absence of the evidence needed to argue in favor of a grammar with hierarchical phrase structure, then, the scientifically prudent thing to do is assume the weaker hypothesis that Kanzi has a model of English syntax making reference to linear word order, but not to hierarchical constituent structure, and hope for more decisive evidence in the future.

Another relevant case in this light concerns Kanzi's ability to handle sentential coordination, as in (9).

(9) a. 326. (C) *Take the keys and open the play yard*. (Kanzi makes a sound like "whuh" and does so.)

b. 461. (C) *Get Rose by the hand* [. . .] *Get Rose's hand* [. . .] *and take her to the colony room*. [. . .] (Kanzi takes Rose's hand . . . , leads her to the colony room, and waits for her to open the door.)

From a human point of view, such cases look like canonical examples of hierarchical phrase structure: two sentences are symmetrically combined to form a single larger sentence. However, it must be noted that the word *and*, which performs this combination, makes no obvious semantic contribution in these cases. Kanzi would respond equally well if he treated the utterances in (9) as two paratactically related simplex sentences. In other words, when we know we are dealing with a grammar which makes use of phrase structure, such as a human grammar of English, it is natural to see sentential coordination as a case of embedding of a constituent within a larger constituent. However, when we are dealing with a grammar where the presence of hierarchical constituency is up for grabs, sentential coordination is unreliable as a source of evidence.[4]

A final type of nonevidence concerning constituency can be illustrated with cases of sentential complementation. Once again, Kanzi can handle a variety of such structures, including one case with more than one level of embedding, as in (10).

(10) a. 111. (C) *I think we need to give the balloon to Kelly.*

b. 645. (C) *I want Kanzi to grab Rose*. (Kanzi turns around and grabs Rose on the leg, then walks away.)

c. 100. (C) *See if you can make your doggie bite your ball*.

d. 581. (C) *Kanzi, tell Rose that you want to go outdoors*. (Kanzi turns, looks at Rose, and gestures toward the play-yard door.) [. . .]

In all these cases, however, Kanzi's responses would probably be identical if he ignored the upstairs clauses, and just responded to the most embedded clause. Kanzi's response to (10a), for example, could equally well be a response to *Give the balloon to Kelly* in isolation. Similarly, Kanzi's usual method of getting outdoors, as in (10d), is by enlisting the help of Rose or another assistant, so *Go outdoors* alone might well have produced an identical response from Kanzi. Finally, the case in (10c) is slightly more complex, as Kanzi can't simply respond to *Bite your ball* alone in that case. However, we know from elsewhere in the corpus that Kanzi sometimes treats declaratives as instructions to bring about the state of affairs described in the sentence, as in (11).

(11) a. 633. (C) *Rose is going to grab Kanzi.* (Kanzi makes a sound like "grab," then looks around briefly at Rose. He sits stiffly with his back to Rose and appears to be waiting to be grabbed. [. . .])

b. 635. (C) *Liz is going to tickle Kanzi.* (Kanzi looks toward Liz, holds his hand out to her, vocalizes "enngh," then approaches Liz, then goes over and sits down near her and holds his hand out to her. Liz stands up. Kanzi motions toward himself, then laughs, then signs *tickle*, then leans down to be tickled. Liz tickles him.)

In that case, in (10c), Kanzi is plausibly responding to *Your doggie bite your ball* alone, rather than the full biclausal structure.

Overall, then, we have seen clear evidence for Kanzi's sensitivity to word order, but we have also seen that some potential pieces of evidence for sensitivity to phrase structure have to be ruled out, because the decisive examples have not been documented. However, in two cases, such evidence is available. We will see in the next section that Kanzi conspicuously fails to respond appropriately in just these cases.

4 Where Kanzi's English Fails Him

What is the minimum required to demonstrate the necessity of hierarchical constituent structure in language production or comprehension? Chomsky (1956), and other early writings by Chomsky, demonstrate convincingly that an adequate characterization of the set of possible sentences of English (and, by implication, other human languages) needs to make use of such structures, and cannot proceed solely on the basis of Markovian relations among elements in a string. However, those demonstrations rely on our intuitive knowledge of the infinite set of possible (not actual) sentences of English or other human languages. For instance, Chomsky discusses the possibility of recursive application of the rules in (12), generating a discrete infinity of sentences along the lines of (13).[5]

(12) a. S → If S then S
b. S→ Either S or S

(13) a. If [he comes] then [I leave].
b. Either [he leaves] or [I leave].
c. If [either [he leaves] or [I leave]] then [either [Jane will get upset] or [Bill will start screaming]]
d. If [either [he leaves] or [I leave]] then [if [either [Jane gets upset] or [Bill starts screaming]] then [we're in trouble]]

However, we have no way to access bonobo intuitions about the shape of English, and any finite subset of such sentences can straightforwardly be described without making use of constituent structure, at the very least by enumeration. Similar arguments can be made for most classic demonstrations of the necessity of phrase structure in describing human language.

I propose instead that the surest demonstration of hierarchical structuring of linguistic input comes from Kanzi's interpretation of it. Regardless of the syntactic structure assigned to a string, many elements need to form *semantic* units for proper interpretation.

In principle, this could also be true of the modification constructions described in the previous section. However, as we saw, the crucial data required to determine whether Kanzi really does interpret a modifier and its modifiee as a semantic unit are missing in those cases. In two other cases, though, we have the data we need to make the call. These concern, firstly, determiners and other functional structure, and secondly, NP-coordination. I will now discuss each of these in turn.

4.1 Failure to Interpret Functional Structure

A standard theory of formal semantics tells us that a determiner functions to specify the relation between the denotation of its nominal sister and the denotation of the predicate that takes the noun phrase as an argument. So when we hear that *The water is cold*, we know that the unique maximal contextually salient body of water is included within the set of cold things; when we hear that *John wants an ice cream*, we know that the intersection of the set of ice creams and the set of things that John wants is nonempty, and so on.

A further important property of determiners (at least, in the English to which Kanzi is exposed) is this: they work exactly the same, regardless of where in the sentence they occur. They take a property-denoting nominal, with which they form a constituent, as an argument, and output something that functions as an argument to the verbal predicate. In other words, Det-N forms a unit, and the wider syntactic and semantic context cares about that unit, rather than the specific atoms that combine to form that unit. This, of course, is just the behavior that *constituency* refers to, and so, if Kanzi's model of English syntax does not make reference to constituency, we expect him to struggle to interpret determiners.

Consider the following example.

(14) Make the doggie drink some water.

If someone whose linguistic competence did not include phrase structure were exposed to enough tokens like (14), he may be able to figure out how to interpret *the* along the following rough lines.

(15) a. The word (or maybe the noun) following *the* describes some class of objects, of which only one member should be salient in the immediate context.

b. That unique object does something. What it does is described by the word (or maybe the verb) following the noun we found in (a).

c. Other words (or maybe other nouns) following the verb describe (if appropriate) which other objects the unique object in (a) performs the action described in (b) on.

This interpretation has a certain amount in common with a typical semantic analysis of *the*. Moreover, it makes no reference to constituent structure. Rather it sets up a string of expectations: after *the*, we're going to encounter, first a noun, and then a verb, and then maybe some other nouns. However, this interpretation fails to generalize to cases like (16).

(16) Make a doggie drink the water.

The difference here is that the noun phrase including *the* is now in object position, rather than subject position. This means that in (16), the interpretation procedure would find a following noun (*water*), as predicted by (15a), but then the verb predicted by (15b) to follow *water* just isn't there: we reach the end of the string in (16) before we reach the end of the interpretation procedure for *the* in (15). So if we were to interpret *the* without a notion of constituency, we would need different interpretation procedures for *the* in object position and *the* in subject position. Considering further positions in which *the* may be found just makes the problem worse. In other words, if we didn't have access to constituent structure, we would have to posit massive lexical ambiguity for determiners like *the*, vastly complicating the task of acquiring the meaning of such determiners.

In fact, similar arguments can be made for other function words. Kanzi is only exposed to one other class of function words which require interpretation for accurate responses to instructions, namely spatial prepositions (he is also exposed to function words that he can safely ignore, such as *can* and *do* in yes/no questions used as indirect instructions, and the semantically vacuous complementizer *that*). And again, such words function identically regardless of the wider syntactic context in which they are located. So *on* does the same job in *Show me the snake (that's) on TV* and *Put the ball on the potty*, despite the different syntactic environments (and correspondingly different expected interpretations) in which the preposition is found (more distinct structures in which the prepositional phrase occurs nonfinally in the string only occur very rarely in the test material).

In fact, it is difficult to know exactly how accurate Kanzi is with such function words. As mentioned above, there is a default behavior of putting an object in Kanzi's vicinity *in* containers, giving it *to* animates or taking it *to* locations, and putting it *on* anything else. Certainly, Kanzi has no problem with these actions, as (17) shows, but the fact that Kanzi manages to do something sensible with instructions that are so predictable in this way tells us relatively little about his understanding of the prepositions themselves.

(17) a. 1. (C) *Put the pine needles in the backpack.*

b. 10. (C) *Get the toothpaste and put it on the Fourtrax*.

c. 12. (C) *Get the rock. . . give it to Kelly*.

d. 47. (C) *Get some money. . . Take it to the play yard*.

Instead, we need to look at cases where the spatial relation specified by the preposition is different from that predicted by these crude heuristics. Certainly, Kanzi is not hopeless in these cases, producing responses coded as correct over half of the time (by my count, Kanzi is coded as responding correctly in 32 trials with unpredictable prepositions out of 53, or 60% of the time). However, the value of these codes is questionable, because Kanzi does not need to bring about the specific spatial relation encoded by the preposition to be scored as correct, as explicitly indicated in the following example.

(18) 298. (C) *Put the hat on Rose*. (Kanzi tosses the hat to Rose.) [C is scored because Kanzi gives the hat to Rose and it is not assumed that Kanzi understands the difference between *on* and *adjacent to*.]

Once again, then, the crucial data are missing, but here, I feel we are justified in saying that it is unlikely that PP is interpreted as a unit. Rather, the most likely interpretation of the comment in (18) is that Kanzi pays no attention to the preposition, and simply interprets the nouns in isolation, inserting a default spatial relation to decide what to do to the theme argument. In the case of (18), then, for example, Kanzi only interprets *Put hat Rose*, and fills in the semantic blanks with default relations among the arguments.

A similar story can be told for determiners. Kanzi is only reguarly exposed to three determiners in the test material: *the, a* and *some* (there are also very rare occurrences of *this* and *that* in the corpus). His response to *the* is similar to that of a human English speaker. However, he has a tendency to respond as if *a* or *some* also mean *the*.[6] That is, Kanzi ignores the implicature that if we ask him to get a tomato, we don't want him to get all the tomatoes. Two such examples are the following.[7]

(19) a. 464. (C5) *Go to the refrigerator and get a tomato*. (Kanzi picks up the tomato in the array.) E says, "Huh un, no, put it down." (Kanzi puts it in the bin of water in front of him.) E says, "Go to the refrigerator and get a tomato." (Kanzi vocalizes right after *refrigerator*, then heads off toward the refrigerator [. . .] Kanzi points to a bag of small tomatoes. Rose gives them to him, and he heads back to E [. . .])

b. 584. (C3) *Kanzi, go get a carrot for Rose, carrot*. (Kanzi goes to the large 50-pound bag of carrots, begins opening it, and starts taking out all the carrots.) [. . . After a while,] E says, "Can you give Rose a carrot?" (Kanzi makes a sound like "carrot" and pushes several of them toward Rose with his feet.)

Note that these two responses are, once again, scored as correct. In a sense, this is accurate: many semanticists would argue that Kanzi is in fact complying with the explicit instructions in these cases, but simply ignoring a scalar implicature to the effect that, if we ask for a carrot, we only want one carrot. However, this does mean that the evidence concerning Kanzi's interpretation of determiners is once again equivocal: he looks in many cases like he simply interprets *N* as meaning *the N(s)* in every case, and ignores the actual determiner he is presented with.[8]

Taken in conjunction with the above arguments concerning NP-modifiers and prepositions, this has the effect of removing any evidence that Kanzi deals with NPs and PPs as constituents in a hierarchically structured model of syntax: we have no clear evidence that Kanzi interprets determiners or prepositions, and no clear evidence that adnominal modifiers like adjectives or relative clauses form a semantic unit with the head noun for Kanzi. In other words, we have no clear evidence that the strings corresponding to NPs or PPs form a constituent for Kanzi in the same way that they do for human English speakers. Moreover, there is some evidence that Kanzi ignores just those function words which require constituency for correct interpretation. If this is accurate, the arguments of verbs, for Kanzi, are plausibly not noun *phrases* or prepositional *phrases*, but nouns alone.

However, this conclusion, as it stands, is somewhat equivocal: for one thing, most of the relevant data are not systematically coded in the Savage-Rumbaugh et al. corpus, so this interpretation requires a certain amount of reading between the lines, and remains provisional until more accurate data become available. Secondly, the meanings of prepositions and (particularly) determiners are far harder to learn than the names for concrete objects and events that form the bread and butter of Kanzi's vocabulary, particularly given the possibility that such function words may not have been available as part of his lexigram keyboard at the time, limiting Kanzi to comprehension rather than production of such words.[9] In that case, there are plenty of nonlinguistic reasons why Kanzi might struggle in just these areas. However, evidence from a second deficit militates against the possibility that these considerations are the whole story.

4.2 NP-coordination

Part of understanding the lexical semantics of a verb involves knowing the number of participants involved in the event it describes. We know that giving requires a giver, a receiver, and a thing given. Likewise, scaring requires a scarer and a scaree. And part of our knowledge of a language concerns how to map this semantic predicate–argument configuration onto the configurations of words and phrases which the syntax gives us. This is how we know who did the giving and who did the receiving when we hear a sentence like *John gave Bill a present*.

One major constraint on such mappings between syntax and semantics is that every semantic participant should correspond to a syntactic argument, and vice versa.[10] Every major syntactic theory has some way of ensuring this happens: Government and Binding theory (Chomsky, 1981) and its successors have the θ-criterion, Lexical-Functional Grammar (Bresnan, 1982) has conditions on coherence and completeness, and so on. These all take a form along roughly the following lines:

(20) a. Every argumental noun phrase must bear a thematic role.

b. Every thematic role must be assigned to an argumental noun phrase.

Note that these constraints are stated at the level of phrases, rather than words. By hypothesis, Kanzi doesn't have such a notion in his grammar, and so the principle governing the mapping between nominals and participants must take a slightly different form in his grammar. This is the natural candidate:

(21) **Word-level θ-criterion:**

a. Every argumental *noun* must bear a thematic role.

b. Every thematic role must be assigned to an argumental *noun*.

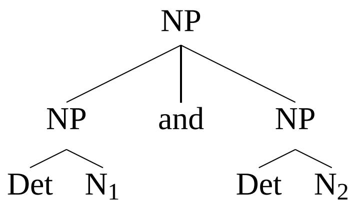
In many cases, the number of nouns and the number of noun phrases come apart. In fact, this is plausibly true already of the locative modifiers discussed in section 2: if Kanzi really pays no attention to determiners or prepositions, then a sentence like *Get the lighter that's in the bedroom* ((6a) above) reduces to simply *Get lighter bedroom*.

This causes an instant problem with the word-level θ-criterion: Kanzi presumably realizes that he is the intended agent of this action, and also knows that getting involves a getter and a thing that is got. However, he has two nouns (*lighter* and *bedroom*) fighting for the single remaining thematic role, that of patient, or thing that is got.

In the case of locatives, a semantic soup strategy comes to the rescue. A lighter is something that Kanzi is capable of getting, and a bedroom is not, so he may reasonably assume that *lighter* is the patient of *get*. But *bedroom* is presumably not just noise, so it probably represents a location of the lighter that Kanzi is supposed to be getting. So a reasonable interpretation is arrived at, despite the apparent problem of having too many nouns floating around for the number of participants required.

Things aren't always so straightforward, however. An important construction in this respect is NP-coordination. In human grammars, this functions to combine two NPs into a single larger NP, and thereby allow them to share a single θ-role, as in (22).

(22)



However, this trick relies on the ability to group words into phrases, and group phrases recursively into larger phrases. If Kanzi lacks such an ability, such sentences should be a source of serious confusion, as there will always be more nouns around than thematic roles. The mapping from word order to semantics is disrupted by the presence of this surplus noun. Moreover, a semantic soup strategy is of no use here: Kanzi has no grounds for deciding that either N_1 or N_2 is *the* argument of V: when he hears *Show Sue the doggie and the rock*, he is expected to act on *both* the doggie and the rock. On the other hand, neither *doggie* nor *rock* is an obvious modifier in the way that *bedroom* was in (6a): to be sure, *Show Sue doggie rock* could mean *Show Sue the doggie next to the rock*, or *Show Sue the rock near the doggie*, and it might even mean *Show Sue the doggie and the rock*, but deciding between such options is inevitably going to be a seat-of-the-pants affair.

Here, then, we have a clear prediction: we predict chaos whenever NP-coordination is involved. We can also make a stab at representing what chance performance might look like in these respects: given the input $V\ [NP_1\ and\ NP_2]$, the three obvious possibilities are to ignore NP_1, ignore NP_2, or act on both (the fourth logical possibility, ignoring both NPs, just increases the θ-criterion worries, as Kanzi then not only has a surplus of noun phrases to deal with, but also an unassigned thematic role). As a null hypothesis, we might expect all three actions to be equally frequent. In that case, chance behavior would be to respond correctly $\frac{1}{3}$ of the time, and to ignore each coordinate NP $\frac{1}{3}$ of the time.

The number of relevant sentences in the Savage-Rumbaugh et al. corpus is small, but the patterns conform broadly to these predictions. I will now go through these results in detail.[11]

The Savage-Rumbaugh et al. corpus contains 26 sentences involving a coordinated NP object. Of these, one produced a clearly incorrect response (Kanzi was asked to show the lighter and the water, and instead put the lighter in the water). This trial will be ignored below, although Kanzi's response is incorrect in an interesting way, suggesting a default response when all else fails of ignoring the verb and substituting some basic action which relates the objects described by the nouns in the sentence. Moreover, seven trials produced responses where it was unclear whether Kanzi was genuinely responding to the coordinated NPs as a constituent or not. Three times, this was because Kanzi was already in contact with one of the two objects in question at the time of the instruction, and didn't perform any further action on that object as a result of the instruction, as in the following cases.

(23) a. 271. (C) *Show Sue the ball and the cereal*. (Kanzi picks the cereal up as E starts to talk. When he hears the sentence, he grabs the ball and tries to open the door to give E the ball. He keeps the cereal in his lap but appears to be attempting to show E the

ball.) [C is scored because Kanzi picks up both requested objects. Whenever either subject is asked to "show" something, they are given the benefit of the doubt if they act on it in some way that could serve to draw E's attention and if E can see this.]

b. 297. (C) *Show me the shot and the ball*. (Kanzi shows E the shot and looks at the ball that he is leaning on.) [C is scored because Kanzi has both items in front of E and can reasonably presume that E will see them [. . .]]

Four further such cases arose when Kanzi was asked to do something to two other apes (Panbanisha and Panzee, or Sherman and Austin), in such a way that it is impossible to tell whether he is really interacting with the two of them as a group, or whether he is really interacting with one, and the other one just happens to be present. For example, in (24a), it is inevitably unclear whether Kanzi intended to scare just Panbanisha, just Panzee, or the two of them, if they are occupying the same area. Similarly, in (24b), giving an orange to both Panbanisha and Panzee is impossible (unless he splits the orange in half): he has to give it to one or the other.

(24) a. 293. (C) *Get the monster mask and go scare Panbanisha and Panzee*. [. . . Kanzi spends a while looking for the mask] (Kanzi finds the mask and goes back to the tool room with it to scare Panbanisha and Panzee [. . .])

b. 311. (C4) *Take the orange to Panban and Panzee*. [. . . Kanzi initially tries to eat the orange] Rose [. . .] says, "Kanzi, stop." (Kanzi puts the orange back down on floor.) E says, "Pick it up and take it to Panban and Panzee. Carry it to Panban and Panzee. You take it to them." [. . . Kanzi gets E to open a door for him] (Kanzi goes to the middle room door and tries to shove the orange under; Rose comes and opens the door for him, and Kanzi takes the orange in to Panban and Panzee. [. . .])

Clearly, in all these cases, we can't learn anything about what Kanzi *intended* to do, even if the action that actually resulted happens to have been appropriate. These trials must also be discarded, then.[12]

This leaves 18 unproblematic NP-coordination trials. Of these, Kanzi ignored the first NP on 9 trials, as in (25a), ignored the second NP on 5 trials, (25b), and responded correctly to 4 trials (25c).

(25) a. 428. (PC) *Give the water and the doggie to Rose*. (Kanzi picks up the dog and hands it to Rose.)

b. 526. (PC) *Give the lighter and the shoe to Rose*. (Kanzi hands Rose the lighter, then points to some food in a bowl in the array that he would like to have to eat.)

c. 281. (C) *Give me the milk and the lighter*. (Kanzi does so.)

Kanzi's overall accuracy on this construction is therefore 22.2%. In comparison, Savage-Rumbaugh et al. (1993, p. 77) give his overall accuracy across the corpus as 72%. The same trials were presented to a human infant, Alia. Alia's accuracy across the corpus was slightly lower, at 66%, but her accuracy on the NP-coordination trials is around $\frac{13}{19}$, or 68.4%.[13] We have, then, a clear case of a species-specific, construction-specific deficit. Kanzi does not perform worse than a human infant across the whole corpus, but he performs much worse than both his usual standard (2-tailed Fisher exact test, $p = 0.00002$), and the human control ($p = 0.008$), on this one construction.

Of course, this raises the question of why this specific construction should be so problematic. The deficit does not go unnoticed in Savage-Rumbaugh et al. (1993), although it is treated as nonsignificant. They discuss it as follows:

> The simplicity of both the semantic and the syntactic components of type 2B sentences [those involving NP-coordination] suggests that Kanzi's difficulty was perhaps due more to short-term memory limitations on the overall amount of information than to processing limitations on the information that was available to him. Indeed, it seems possible that the semantic and syntactic structure in sentences such as "Feed the doggie some milk" permitted Kanzi to go beyond the typical constraints of his short-term memory system by enabling him to process or chunk the

> information in a meaningful manner. By contrast, sentences such as "Give the doggie and the milk" do not engage semantic chunking strategies but rather force reliance on short-term memory alone. (Savage-Rumbaugh et al., 1993, p. 85)

This is undoubtedly a partially accurate description of the problem. Moreover, as suggested to me independently by Moira Yip and Jim Hurford, relating the pattern to short-term memory problems raises the possibility that the noticeable but nonsignificant tendency for Kanzi to act on only the last of two coordinated NPs may be a recency effect. However, the question remains exactly why Kanzi would have problems with semantic chunking precisely here. If we assume that Kanzi has phrase structure, it's a mystery: many of the above sentence types look far more complicated on the surface, so why should he find, say, modification of NP by a locative relative clause so much easier than NP-coordination?

On the other hand, if Kanzi lacks phrase structure, we understand why he fails here: NP-coordination is a very simple case which is, if not impossible, at the very least extremely hard,[14] to represent without recourse to hierarchical phrase structure.

I take this to be the most compelling argument in favor of the hypothesis that Kanzi has a constituency-free model of English syntax: not only do we repeatedly fail to find positive evidence for the presence of phrase structure when such evidence might in principle be available (as was the case with locatives or attributive adjectives in section 2, for example), but we also see that Kanzi struggles specifically where a constituency-free model of his competence predicts that he should fail — he cannot interpret determiners and prepositions, and he is baffled by NP-coordination. The predictions of a model without constituency therefore provide an excellent fit for both the areas where Kanzi clearly understands, and the areas where he surprisingly fails to.

5 Conclusion and Discussion

Close scrutiny of the Savage-Rumbaugh et al. corpus provides strong support for the notion that Kanzi comprehends English using a grammar based purely on linear order, without the capacity to represent hierarchical constituent structure. This contrasts with the grammatical knowledge that humans have of natural languages, in which hierarchical constituent structure is pervasive. This result can be taken further in a variety of ways. I will mention three of these briefly here, by way of a conclusion and a program for future research.

5.1 Language Evolution

It is very hard to look at Kanzi's phrase structure deficit without being tempted into speculation about the evolution of the language faculty. If this result holds, it means that one of the closest species to humans in evolutionary terms has the capacity for some embryonic form of syntax. It is therefore plausible that the last common ancestor of *Homo* and *Pan* also possessed this capacity (our other closest relative, the common chimpanzee, has not been tested as extensively as Kanzi, but it seems likely from the earlier ape language experiments that chimps also have at least some capacity to learn patterns requiring sensitivity to linear order). However, it is also clear that Kanzi's syntax, lacking phrase structure and long-distance dependencies, is far less expressive than human syntax. This suggests that creatures relatively early in our evolutionary lineage had at least the latent ability to make use of *some* basic syntactic information, but says nothing about the subsequent development of full human syntax — a position suggested by Bickerton (1990) and compatible with that reached on quite independent grounds by Jackendoff (2002, ch.8).

This also has implications for the debate instigated by Hauser, Chomsky & Fitch (2002) concerning the unique components of the language faculty. Hauser et al. make the strong claim that the only distinguishing features of the human language faculty are some form of recursion and mappings from syntactic structure to the sensorimotor and conceptual-intentional interfaces, a position apparently lent some indirect support by an experiment reported in Fitch & Hauser (2004) showing that cotton-top tamarins fail to learn to recognize strings of the form A^nB^n, but can recognize strings of the form $(AB)^n$, for $n = 2, 3$ (see Gentner, Fenn, Margoliash & Nusbaum 2005; Perruchet & Rey 2005; Rogers & Pullum 2007; Corballis 2007; and Hochmann, Azadpour

& Mehler 2008 for largely skeptical discussion). The interest of these two stringsets is that the latter can be recognized by a very simple regular grammar: roughly, start with an A, make sure that B follows A and A follows B, and end with a B. A^nB^n, on the other hand, requires something beyond a regular grammar: one popular way of representing this language involves a context-free phrase structure grammar using recursive center-embedding, as in (26).

(26) a. S→a S b

b. S→a b

Note that this is unlikely to be the way in which humans successfully recognize whether a given string is in A^nB^n or not: this sort of center-embedding is extremely hard for humans to parse after a couple of levels of embedding, as shown by the following example, based on Rogers & Pullum (2007, p. 2).

(27) a. [People left]

b. [[People [people left]] left]

c. [[People [[people [people left]] left]] left]

All of these sentences should be grammatical: they result from modification of a noun phrase by a relative clause, containing a noun phrase which may itself be modified by a relative clause, and so on. However, after only a couple of levels of embedding, our ability to perceive the structure underlying the strings of *people*s and *left*s disappears.

It seems unlikely, then, that humans use a grammar like that in (26) to recognize the string language A^nB^n: instead, we rely on other, easier, strategies, like simply counting the number of As and Bs and checking if they match.[15] Even if we did, though, the failure of cotton-top tamarins to recognize such a stringset would not tell us much about their ability to handle phrase structure, as argued extensively in Perruchet & Rey (2005) and Corballis (2007).

This is where the data from Kanzi can make a contribution. One difference between the data presented here and those in Fitch & Hauser (2004) is that Fitch & Hauser's tamarins had only to learn to recognize an input as a member of a class, while Kanzi has to recognize an input, interpret it and act appropriately. We saw above that Kanzi's model of grammar lacks hierarchical phrase structure. A grammar with hierarchical phrase structure contains at least the latent possibility of recursive embedding, in that it is possible for a larger phrase to contain a smaller phrase of the same type. Therefore, the data presented here provide a complementary source of evidence supporting Fitch & Hauser's claim that nonhuman primates lack recursion. However, these new data also sharpen the claim: Kanzi's linguistic competence (and so possibly that of other nonhuman primates) not only does not include recursion, but also does not include the more basic ability to recognize hierarchical phrase structure.

5.2 Language Acquisition

A second area of potential interest raised by this research concerns the development of humans as individuals, rather than as a species. Bickerton (1990, ch.5) articulated the possibility that ontogeny recapitulates phylogeny in the sense that human infants go through a developmental stage in which their linguistic abilities stretch to stringing words together in the correct order, but stop short of hierarchical phrase structure. According to Bickerton's hypothesis, this "protolanguage" phase would be equivalent in these respects to the abilities of some present day language-trained apes, and some common ancestor of ours.

If we look at data concerning early acquisition, we find some suggestive points which appear to support such a hypothesis. However, that support is still a fair way from being anything like conclusive. To be clear, if Bickerton's hypothesis of a "protolanguage" phase in human linguistic development is accurate, we expect to find a stage at which infants struggle with the same constructions that Kanzi struggles with.

The first such problem concerned Kanzi's failure to interpret functional structure, and determiners in particular. Here, the evidence concerning infants is inconclusive. We know from

Katz, Baker & Macnamara (1974) that girls, at least, as young as 17 months distinguish between common and proper nouns in English on the basis that only the former are preceded by a determiner. We also know from Shipley, Smith & Gleitman (1969) and Gerken, Landau & Remez (1990) that children, at least by the time they reach the two-word stage, respond more reliably to adultlike English forms including a determiner than to telegraphic forms in which the determiner is omitted, even though they are still producing telegraphic speech themselves. This shows us that young infants already have a clear knowledge of where in a string determiners belong — they precede common nouns and they don't precede proper nouns. What hasn't been tested, to my knowledge, is the semantic contribution that a determiner makes at this stage. When do infants start to reliably differentiate between the meaning of *the* and *a*, for example, rather than simply treating them as markers of a common noun to come? The possibility exists (and appears to be hinted at in Shipley et al. 1969) that infants work out the distribution of these words before they work out what they mean, but I am unaware of any evidence supporting this position at present.

There is also some evidence that, although children learn to interpret coordinated noun phrases in subject position correctly by 25 months (Naigles, 1990), there is a tendency for younger infants to misinterpret coordinate structures. In currently unpublished work, Gertner & Fisher (2005) show that 21-month-olds interpret an intransitive sentence with a novel verb like (28) as describing a causal action, with the first noun in the string describing the agent and the second noun describing the patient. One possible analysis of this is that these infants are calculating argument structure on the basis of linear order among nouns alone, rather than treating the coordinate noun phrase as a constituent.

(28) The girl and the boy are gorping. (Gertner & Fisher, 2005)

Moreover, infants still struggle to interpret the same coordinated noun phrases correctly in certain other environments (e.g. *Find Big Bird and Cookie Monster gorping!*) at 28 months (Hirsh-Pasek & Golinkoff, 1996; Fisher, 2002). All of this, though far from conclusive, is compatible with an account on which infants are initially sensitive to word order alone, rather than phrase structure, with phrase structure emerging gradually after 21 months. Whether such an account can be shown to hold more generally is a question far beyond the scope of this paper.

5.3 Domain-specificity

A final interesting question raised by Kanzi's constituency deficit concerns the degree of domain-specificity that this deficit shows. Ever since Lashley (1951), researchers into human cognition have seen hierarchical structure everywhere, from vision and motor coordination through navigation and social cognition to planning and problem solving, language, and music. The question is, do any of these make use of the same hierarchical structure which was shown to be lacking in Kanzi's comprehension of English?

At this stage, the best answer I can give is "maybe", but we have to proceed with caution. One common type of hierarchical structure discussed in the literature concerns relations among items of stored knowledge: stored knowledge about how to drive contains stored knowledge about how to use the brake as a proper subpart. While there is clearly some hierarchical structure in such mereological relations, this is of a different nature from the sort of hierarchical structure we see in language use. Specifically, the hierarchical structure in language use extends beyond the fixed, memorized phrases (where an English speaker may have the phrase *Pleased to meet you* stored in long-term memory alongside the word *pleased*, for example), and shows up also in novel, creative utterances, which obey the same phrase structure rules as stored expressions.

This is important, because we know that even evolutionarily distant animals can *learn* to produce fixed sequences of actions which look from a human perspective like they have some constituent structure (see, for example, Byrne & Russon 1998). The more pressing question is whether they can spontaneously produce novel behavior on the basis of some hierarchically structured generative system.

To take one example from Köhler's (1925) classic study of chimpanzee problem-solving (see also Steedman 2002 for somewhat different discussion of these experiments from a linguistic perspective), the chimp Sultan can learn to push a box under a banana suspended from a ceiling, climb on the box and reach the banana. Moreover, this is essentially one-shot learning: Köhler's descriptions indicate a huge disparity between the time and effort taken to solve the problem of how to get the banana on first exposure, and on subsequent exposures. From a human perspective, this looks like a chimpanzee having the initial subgoal of creating a platform from which the banana is reachable, and then grabbing the banana, and these subgoals can be seen as defining the problem-solving equivalents of constituents (as in Jackendoff 2007, ch.4).

We must be wary of overinterpretation, however. If Sultan really were approaching the problem in a humanlike way, with subgoals defining constituents leading to an ultimate goal, he should be able to treat the subgoal of finding a box or other suitable platform as a proxy for the ultimate goal of getting the banana: in the same way as an intransitive verb can be treated as a function from noun phrases to sentences, the banana dangling from the ceiling can be treated as a function from platforms in the right place to food in Sultan's belly. Adult humans, at least, have no problem reasoning in this way: if we need something, but it's too high for us to reach and there's nothing around to hand, we go and grab a chair or a box to stand on, and bring it to where it's needed.

According to Köhler's description (particularly pp.53–5), chimps fail to show this behavior. If they can push a nearby box under a banana, so that they are making more or less monotonic progress towards their ultimate goal, then they will learn to do so. If, however, they have to walk *away* from the banana to find a box in an adjacent corridor, the task is hugely complicated, and only successfully completed with some degree of human intervention. This is suggestive (but once again, hardly conclusive proof) of a chimpanzee problem-solving procedure which consists of chaining together possible actions, each one bringing the chimp physically closer to his goal, in contrast to a human problem-solving procedure which can set up a subgoal as a stepping-stone towards an ultimate goal. This allows us to work monotonically towards attaining a subgoal, even if this means taking us physically further away from any higher goals, by structuring our novel, creative plans hierarchically.

Perhaps the most compelling evidence for creative, hierarchically structured planning in a nonhuman primate comes from Russon & Galdikas (1993:156) and Byrne & Russon (1998:679–80). These authors discuss a case in which an orangutan, Supinah, gets into a half-sunken canoe, unties it and bales it out, pushes it along a dock to allow her to get past a guard, scares away a group of laundry workers, steals soap and laundry, and washes some clothes. Unlike the case of Sultan above, this plan involved avoiding the most direct route to the laundry, which was blocked by a guard employed to protect the laundry workers from intimidation by orangutans, and taking a more circuitous route instead. It is at present unclear to me just how to treat this one (undoubtedly very impressive) observation. It is undoubtedly a novel plan, and suggestive of hierarchical structure, as opposed to a simple chain of steps bringing Supinah closer to her ultimate goal of soap and laundry. At present, it stands as an isolated incident, but it is exactly the sort of evidence that could shed light on the degree of domain-specificity of Kanzi's constituency deficit.

As ever, then, the evidence for a constituency deficit in bonobo comprehension of English raises more questions than it answers. Where does this leave the debate over the evolution of language? To what extent does ontogeny mirror phylogeny in this respect? And how domain-general is the constituency deficit that Kanzi showed? I believe that the last two problems, at least, are tractable, although we're currently far away from any firm answers. Hopefully, future research will enable us to evaluate the status of Kanzi's linguistic constituency deficit in the light of these broader questions.

References

Anderson, S. (2004). *Doctor Doolittle's delusion*. New Haven, CT: Yale University Press.

Bickerton, D. (1990). *Language and species*. Chicago: University of Chicago Press.

Bresnan, J. (Ed.). (1982). *The Mental Representation of Grammatical Relations*. Cambridge, MA: MIT Press.

Byrne, R. & Russon, A. (1998). Learning by imitation: A hierarchical approach. *Behavioral and Brain Sciences*, *21*, 667–684.

Chomsky, N. (1956). Three models for the description of language. *IRE Transactions on Information Theory*, *2*, 113–124.

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

Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht: Foris.

Corballis, M. (2007). Recursion, language, and starlings. *Cognitive Science*, *31*, 697–704.

Fisher, C. (2002). The role of abstract syntactic knowledge in language acquisition: A reply to Tomasello (2000). *Cognition*, *82*, 259–278.

Fitch, W. T. & Hauser, M. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science*, *303*, 377–380.

Gentner, T., Fenn, K., Margoliash, D., & Nusbaum, H. (2005). Recursive syntactic pattern learning by songbirds. *Nature*, *440*, 1204–1207.

Gerken, L., Landau, B., & Remez, R. (1990). Function morphemes in young children's speech perception and production. *Developmental Psychology*, *26*, 204–216.

Gertner, Y. & Fisher, C. (2005). Predicted errors in early verb learning. Abstract, 27th Annual Conference of the Cognitive Science Society, Stresa, Italy. Also presented at the 31st Boston University Conference on Language Development, 2006.

Hauser, M., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, *298*, 1569–1579.

Hirsh-Pasek, K. & Golinkoff, R. (1996). *The origins of grammar*. Cambridge, MA: MIT Press.

Hochmann, J.-R., Azadpour, M., & Mehler, J. (2008). Do humans really learn A^nB^n artificial grammars from exemplars? *Cognitive Science*, *32*, 1021–1036.

Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. New York: Oxford University Press.

Jackendoff, R. (2007). *Language, consciousness, culture: Essays on mental structure*. Cambridge, MA: MIT Press.

Katz, N., Baker, E., & Macnamara, J. (1974). What's in a name? a study of how children learn common and proper names. *Child Development*, *45*, 469–473.

Köhler, W. (1925). *The mentality of apes*. London: Kegan Paul, Trench, Trubner, & Co. Translated by Ella Winter. Reprinted (1999) by Routledge.

Lashley, K. (1951). The problem of serial order in behavior. In L. Jeffress (Ed.), *Cerebral mechanisms in behavior: The Hixon symposium* (pp. 112–136). New York: Wiley.

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, *17*, 357–374.

Perruchet, P. & Rey, A. (2005). Does the mastery of center-embedded linguistic structures distinguish humans from nonhuman primates? *Psychonomic Bulletin & Review*, *12*, 307–313.

Rogers, J. & Pullum, G. (2007). Aural pattern recognition experiments and the subregular hierarchy. To appear in the proceedings of Mathematics of Language 10, UCLA Working Papers.

Russon, A. & Galdikas, B. (1993). Imitation in free-ranging rehabilitant orangutans (*Pongo pygmaeus)*. *Journal of Comparative Psychology*, *107*, 147–161.

Savage-Rumbaugh, E. S., Murphy, J., Sevcik, R., Brakke, K., Williams, S., Rumbaugh, D., & Bates, E. (1993). Language comprehension in ape and child. *Monographs of the Society for Research in Child Development*, *58*, 1–252.

Savage-Rumbaugh, S. & Lewin, R. (1994). *Kanzi: The ape at the brink of the human mind*. New York: Wiley.

Savage-Rumbaugh, S., Shanker, S., & Taylor, T. (1998). *Apes, language, and the human mind*. New York: Oxford University Press.

Shipley, E., Smith, C., & Gleitman, L. (1969). A study in the acquisition of language: Free responses to commands. *Language*, *45*, 322–342.

Steedman, M. (2002). Plans, affordances, and combinatory grammar. *Linguistics and Philosophy*, *25*, 725–753.

Wagner, M. (2005). *Prosody and recursion*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA.

Notes

[1]In fact, Kanzi's competence with adjectives is questionable. He appears to systematically confuse *big* and *little*, presumably two of the most frequent adjectives in his input, as in the following examples.

(i) a. 454. (PC) *Give the big tomato to Liz.* (Kanzi picks up both the big and the little tomato and gives them to Liz.) [PC is scored because this is not a "multiple-object error" but rather a confusion between the words *big* and *little*. Kanzi consistently confuses these words throughout the test.]

b. 455. (PC) *Show me the little bitty tomato*. . . (Kanzi interrupts). . . *the little itty bitty tomato*. (Kanzi points to the large tomato.) E says, "That's the big tomato. Show me the little tomato. Do you see the little tomato?" (Kanzi picks up the big tomato.)

This lack of humanlike behavior where we might expect most opportunities for learning raises questions. Has Kanzi, perhaps, learned *hot water* and *ice water* as single lexical items, with the formal relation between *hot water* and *water* no different from that between *cat* and *catalog*? I simply don't have enough evidence to decide such matters, but we will see presently that this does not impact on the wider issue either way.

[2]There is, in fact, one case broadly along these lines (item 613 in the corpus). Kanzi is asked to *take the potato that's in the water outdoors*. He does so, but takes the other potato in his vicinity outdoors too (a response coded as incorrect). More systematic testing of such possibilities may help to shed light on the strategies that Kanzi uses to interpret examples containing locatives.

[3]A further possibility, suggested by Ray Jackendoff, is that Kanzi's syntax has no constituency, but he attaches modifiers to the closest appropriate head, where *closeness* is defined purely in terms of linear order among symbols, without reference to any hierarchical structure.

This would perhaps predict correct responses in cases like (7), without attributing to Kanzi the ability to handle constituent structure. Again, this hypothesis is testable in principle, but we currently lack any data to bring to bear on the matter.

[4]Sarah O'Neill brought to my attention a passage in Savage-Rumbaugh & Lewin (1994, p. 173), which claims that Kanzi had difficulty understanding sentential coordination structures, but had no corresponding difficulty in understanding a sentence containing a locative relative clause. If this were true, it would be extremely interesting, suggesting that Kanzi copes better with subordination than coordination in some cases, in direct contrast to the evidence reported in this paper. However, the data in Savage-Rumbaugh et al. (1993) do not bear out the claim in Savage-Rumbaugh & Lewin (1994): Kanzi responds appropriately in $\frac{2}{3}$ of sentential coordination tests, and there is no significant difference between his performance here and his performance on items containing relative clauses.

[5]In practice, as with many cases of center-embedding, multiple embeddings of these rules quickly lead to perceived degradation. However, people easily realize that such sentences can *in principle* be extended infinitely, even if working memory limitations get in the way in the real world.

[6]How strong this tendency is could not be determined, because of, firstly, the scarcity of examples in which including an indefinite or a definite in the test sentence should produce different responses, and secondly, doubts that the coding made allowance for this level of semantic subtlety, as exemplified by the discussion of (19).

[7]A video showing one instance of this behavior is available on Sue Savage-Rumbaugh's website, at http://www.greatapetrust.org/research/srumbaugh/rumbaugh.php. Kanzi is asked *Could you pour a little Coke in the water? Could you pour some Coke in the water?* and responds by pouring the whole can in.

[8]An alternative analysis of this pattern, suggested to me independently by Ian Roberts and

Maggie Tallerman, would claim that Kanzi's vocabulary does not include any function words, including determiners. If this were true, these data would have no implications for Kanzi's handling of constituency, because we would expect him to struggle here in any case. However, anecdotal evidence, particularly in Savage-Rumbaugh et al. (1998, ch.1), suggests that Kanzi can at least understand sentential negation, usually expressed by items of functional vocabulary like *don't*. This suggests that Kanzi's vocabulary contains at least some functional items, and so blocks this particular line of analysis.

[9]Sarah O'Neill has pointed out to me, on the basis of images on Sue Savage-Rumbaugh's website, that several such function words are clearly now part of Kanzi's lexigram keyboard. However, that keyboard has been growing steadily over the years, and it remains unclear to me whether the function words were on the keyboard in the early 1990s, when these experiments took place.

[10]Of course, various manipulations of argument structure such as passivisation, middle formation, perhaps certain varieties of pro-drop, and imperatives (the latter, in particular, being widely used in interactions with Kanzi) make this mapping less transparent than this statement suggests, but I will abstract away from such operations here.

[11]My conclusions about Kanzi's accuracy, on the basis of the descriptions given, differ substantially from the codes reported in Savage-Rumbaugh et al. (1993). I will expand on the discrepancies between my interpretations of Kanzi's actions and the original codes when it becomes relevant below, but it is worth noting that these differences in coding lead to a difference between a noticeable but nonsignificant difference between Kanzi's and Alia's performance in Savage-Rumbaugh et al.'s original work, and a highly significant effect here. To be explicit, Savage-Rumbaugh et al. (1993, p. 78) gives the following results: Kanzi: $\frac{7}{21}$ correct trials, Alia: $\frac{12}{21}$ correct trials. The difference between these two datasets is not significant ($p = 0.21$, 2-tailed Fisher exact test). However, we will see below that there are reasons to suspect that these codes do not accurately reflect the subjects' performances.

[12]The major discrepancy between my report, and Savage-Rumbaugh et al.'s report, of Kanzi's accuracy with respect to NP-coordination comes from the fact that most of these trials were scored as correct in the original monograph. While I can't rule out the possibility that this is accurate, it certainly can't be stated with any confidence, and so the danger of overinterpretation is clear.

[13]Savage-Rumbaugh et al. (1993, p. 78) Alia's accuracy on NP-coordination trials as 57%. I can only assume this is partially due to human error. Firstly, Savage-Rumbaugh et al. list only 21 NP-coordination trials in their results table, while 24 such trials directed to Alia are reported in the corpus. Secondly, the figure of 57% corresponds to 12 correct trials out of 21, while the corpus lists 15 trials scored as correct ($\frac{15}{24} = 62.5\%$). Moreover, one trial with Alia is unclear for reasons parallel to the cases in (23) (it is unclear whether Alia really acted on both items), and two trials are unclear parallel to the "Panbanisha and Panzee" cases in (24). Finally, it is unclear how to interpret two trials coded as "OE", where the subject retrieves more objects than requested. If a subject is asked to show the ball and the rubber band, and also tries to show a shot and a rock, is that incorrect? As with the case of Kanzi's failure to interpret scalar implicatures associated with indefinites in section 4.1 above, the answer is unclear, so the prudent thing to do is discard these two trials as well. If all these five trials are discarded, Alia's performance is $\frac{13}{19} = 68.4\%$ accurate.

[14]As investigated thoroughly by Wagner (2005), humans have the ability to apply NP-coordination recursively, such that two coordinate NPs can themselves be coordinated, giving rise to constituents like *[[The ball and the rock] or [the doggie and the milk]]*. It is impossible to interpret the full range of these more complex examples correctly without constituency. However, we have no evidence for such recursive application of NP-coordination in Kanzi-directed speech. In that case, it remains possible in principle to interpret non-recursively applied NP-coordination without constituency, but only at the expense of severe redundancy and/or ambiguity in the lexical semantics of *and*, similar to that discussed for *the* in section 4.1.

[15]Hochmann et al. (2008) suggest that we generally don't even do this, but rather check for a single transition from As to Bs, disregarding information about the precise quantity of each.