FIRST LANGUAGE

Article

First Language
31(4) 442–460

DOI: 10.1177/0142723711406882
fla.sagepub.com
SAGE

The role of dialogue in the ontogeny and phylogeny of early symbol combinations: A cross-species comparison of bonobo, chimpanzee, and human learners

Kristen Gillespie-Lynch
University of California, USA

Patricia M. Greenfield
University of California, USA

Heidi Lyn
University of Southern Mississippi, USA

Sue Savage-Rumbaugh
Great Ape Trust, USA

Abstract
What are the evolutionary and developmental origins of linguistic creativity? Cross-species comparison of the clade consisting of bonobo, chimpanzee, and human suggests that creative word combinations arise from conversation. Analysis of conversational data shows that novel symbol combinations are initially dependent upon conversational input – through the processes of deferred imitation and joint construction – for a bonobo and a chimpanzee exposed to a humanly devised symbol system, as well as for a human child. In all three species, reliance on conversational input for novel symbol combinations fades with development, as novel symbol combinations come to be constructed more independently. These findings resolve the controversy between the

Corresponding author:
Kristen Gillespie-Lynch, Department of Psychology and FPR-UCLA Center for Culture, Brain, and Development, 1285 Franz Hall, Box 951563, Los Angeles, CA 90095-1563, USA.
Email: proserpinae@ucla.edu

claim that ape language is *limited* to imitation and the claim that apes are *not capable* of imitation. Imitation, like conversational co-construction, does not differentiate between bonobo, chimpanzee, and human; instead, imitation and co-construction differentiate stages of learning and development across all three species.

Keywords
ape language, bonobo, chimpanzee, language-competent apes, word combinations

Almost a decade after Herbert Terrace concluded in an article in *Science* that the attempts of his chimpanzee, Nim Chimpsky, to combine linguistic signs were only imitative and lacked the spontaneity of human language (Terrace, Petitto, Sanders, & Bever, 1979), Tomasello and colleagues made the opposite claim: that non-human primates *lack* the human ability to imitate (Tomasello, Davis-Dasilva, Camak, & Bard, 1987). Why this contradiction? Adding to the inconsistency, imitation is a potential mechanism by which human children learn to combine words before extracting the patterns necessary to combine them creatively (Tomasello, 2000). If that is the case, then Nim Chimpsky was doing something that was very human indeed. The present research resolves these theoretical contradictions by demonstrating that various dialogic mechanisms, including but not limited to imitation, are developmental precursors to the independent combination of symbols for a human child, bonobo, and chimpanzee.

While the relationship between language input and human linguistic development is hotly contested in theory (Chomsky, 1959; Skinner, 1957), the empirical findings are quite clear: input is important for phonological development (Saffran, Aslin, & Newport, 1996), lexical development (Andersen, Dunlea, & Kekelis, 1993; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Schwartz & Terrell, 1983; Weizman & Snow, 2001), morphology (Aksu-Koc, 1998), early word combinations (Brown, 1973; Corrigan, 1980; Greenfield & Smith, 1976; Herr-Israel, 2006; Ninio, 2001; Veneziano, Sinclair, & Berthoud, 1990), and syntax (Abbot-Smith & Tomasello, 2010; Huttenlocher, Vasilyeva, Cymerman, & Levine, 2002; Lieven, Pine, & Baldwin, 1997; Matthews, Lieven, Theakston, & Tomasello, 2005; Tomasello, 2000; Tomasello, Brooks, & Stern, 1998; van Veen, Evers-Vermeul, Sanders, & van den Bergh, 2009). Input which is matched to a novice's current stage of development provides a form of social scaffolding, whereby a new skill is drawn into being through interaction with a more experienced other and, only later, emerges independently (Vygotsky, 1962).

Both human children and the language-competent apes described in this study initially encountered linguistic input in the context of recurring routines (Savage-Rumbaugh, 1990; Snow, Perlmann, & Nathan, 1987). Hearing word combinations during everyday routines may allow language learners to reproduce combinations through deferred imitation, or 'reproduction of a model which has been absent for some considerable time' (Piaget, 1962), before they are able to generate equally complex constructions on their own (Brown, 1973; Clark, 1974; Snow & Goldfield, 1983).

Mixed evidence that imitation supports or does not support language development may be due to a failure to distinguish between immediate and deferred imitation. *Immediate* imitation may (Veneziano, 2005) or may not enhance language development (Leonard, Schwartz, Folger, Newhoff, & Wilcox, 1979). However, *deferred* imitation does support the emergence of new vocabulary and early word combinations (Bloom, Hood, & Lightbown, 1974; Kuczaj, 1987; Snow & Goldfield, 1983). Because deferred imitation, unlike immediate imitation, requires an internal representation of the modeled behavior, it may subserve language development (Kuczaj, 1987; Piaget, 1962). Indeed, deferred imitation early in development is concurrently correlated with language comprehension and predictive of gestural communication (Heimann et al., 2006).

Whether deferred imitation of word combinations develops independently of creative constructions (Brown, 1973) or provides the ingredients from which creative combinations evolve (Clark, 1974) has been debated. However, deferred imitation is helpful when new constructions are not yet mastered or when in an unstable language situation, such as when learning a second language or a Creole (Elsen, 1996; Youssef, 1994).

Extensive contact with humans and their symbols may allow highly social species such as apes and dolphins to demonstrate deferred imitation in response to human cues (Kuczaj, Paulos, & Ramos, 2005). Like human infants, enculturated apes exhibit age-related increases in deferred imitation (Bjorklund, Bering, & Ragan, 2000; Jones & Herbert, 2006). When comparing the language-competent apes described in the current study to mother-reared apes and human toddlers, the language-competent apes were better at deferred but not immediate imitation than both the human toddlers and the mother-reared apes (Tomasello, Savage-Rumbaugh, & Kruger, 1993).

The maximum amount of time after which deferred imitation can occur is unknown. Deferred imitation has been observed in toddlers after four months, the longest duration between model and opportunity for deferred imitation yet subject to experimental test (Meltzoff, 1995). Similarly, deferred imitation of word combinations was observed by a toddler reading a picture book with his mother months after hearing his mother produce the combinations (Snow & Goldfield, 1983). Deferred imitation is frequently used as an index of declarative memory and as such, may last indefinitely (Meltzoff, 1995). Thus, in the current study, any prior experience with a word combination is assumed to represent an opportunity for deferred imitation.

Longitudinal records of child language suggest that conversational precedents for subsequent word combinations include deferred imitation of a modeled phrase (Brown, 1973; Snow & Goldfield, 1983; Tomasello, Akhtar, Dodson, & Rekau, 1997) and joint constructions between conversational partners wherein each supplies a part of an utterance (Bloom; 1973; Caselli & Volterra, 1990; Greenfield, Reilly, Leaper, & Baker, 1985; Greenfield & Smith, 1976; Herr-Israel, 2006; Ochs, Schieffelin, & Plat, 1979; Scollon, 1976; Veneziano et al., 1990). The development of linguistic competence in humans involves movement away from joint construction and toward organization within a single utterance (Greenfield & Smith, 1976; Ochs et al., 1979), a phenomenon observed in different languages, including sign language (Caselli & Volterra, 2000; Herr-Israel, 2006; Veneziano et al., 1990).

Two forms of joint constructions are propositions across speakers (see Table 1) and expansions (see Table 2). In the former, different speakers co-construct a proposition by providing semantically related utterances that are distributed across discourse turns (Ochs et al., 1979). Propositions across speakers often precede vertical constructions (see Table 3), which are similar to word combinations except that the elements of the constructions are separated by a pause or intervening speech by another (Scollon, 1976). In an expansion, the caregiver expands the child's shorter utterance into a more complete phrase. Like propositions across speakers, expansions arise through discourse, but unlike propositions across speakers they yield a modeled phrase.

Both joint constructions and modeled phrases generated by a single speaker (see Table 4) provide potential conversational models for deferred imitation of novel word combinations. Early word combinations may arise from deferred imitation of previous discourse (Greenfield & Smith, 1976; Ochs et al., 1979; Scollon, 1976). Indeed, discourse-based interventions have been effective at increasing the number of word combinations produced by both language-impaired and typical children (Schwartz, Chapman, Terrell, Prelock, & Rowan, 1985).

Both deferred imitation and joint construction are forms of social scaffolding utilized by human children as they begin to combine symbols. While evidence exists that apes use imitation or repetition pragmatically as part of conversational discourse (Greenfield & Savage-Rumbaugh, 1993; Pedersen & Fields, 2009) (as well as in other situations, see Bjorklund, Yunger, Bering, & Ragan, 2002; Buttelmann, Carpenter, Call, & Tomasello, 2007), the role of social scaffolding in supporting the emergence of symbolic combinations of apes has never been assessed. Because social scaffolding has been shown to be important for the development of symbolic play across the clade (Lyn, Greenfield, & Savage-Rumbaugh, 2006), social scaffolding may be central for symbolic development, and hence for the emergence of symbol combinations, not only in humans, but also in the other two species comprising the clade bonobo and chimpanzee.

The present study takes on this issue through cross-species comparison, both qualitative and quantitative. Building on prior comparative research demonstrating that enculturated apes comprehend spoken language at least as well as two-year-old children (Savage-Rumbaugh et al., 1993) while also producing two-element combinations that construct the same semantic relations as human children do (Greenfield & Lyn, 2006), this study investigates whether active participation in conversation provides a foundation for symbolic development. We compare the early symbol production of two humanly encultured apes, Panbanisha (a bonobo), and Panpanzee (a chimpanzee), who communicate using lexigrams (abstract visual symbols that represent English words), to that of a human child, NT.

By comparing the symbolic development of one representative of each species, it is possible to identify relationships between conversational input and emergent abilities that are shared with our closest living relatives, as well as those that may be unique to humans. The discovery of a similar role for, and developmental departure from, social scaffolding in symbolic development regardless of species would suggest that the capacity to develop the ability to creatively combine communicative symbols is present across the clade. Note that with one representative of each species, we present an existence proof concerning species capabilities rather than a claim concerning frequency of particular processes.

Our first hypothesis was that conversational sources for symbol combinations would occur in all three species. We aimed to demonstrate that prior conversations provided material from which early word combinations could be generated. Each new word combination produced by language learners could have no precedent or could be derived from an observed precedent, from an unobserved precedent, from generalization across multiple precedents only one of which was observed, or could even be unrelated to observed conversational precedents. We were concerned only with whether or not early word combinations had conversational precedents, not whether the documented precedent was the origin of a given word combination. We saw our analysis as sampling potential precedents, not as always identifying the operative precedent.

Our second hypothesis was that the prevalence of potential discourse sources would decrease over developmental time in all three species, that there would be developmental movement from interdependent to independent symbol combination. We test the first hypothesis by discourse analysis of socially constructed antecedents to independent symbol combination. We test the second hypothesis by quantitative assessment of the developmental relationship between dialogic and independent symbol combinations. Through discourse analysis, we identified word combinations with or without an observed precedent. The corpus consisting of all symbol combinations with and without precedents was then subjected to quantitative analysis to test our second hypothesis. Hence, the examples given in the qualitative analysis are but a small sample of the corpus analyzed by quantitative means.

Method

Human participant

All communications between NT and those present were recorded during monthly formal observation sessions conducted when NT was 18–26 months old, described in detail in Greenfield and Smith (1976). Formal sessions consisted of an observer audio recording the child, his mother, and anyone else who was present in their home during the session. The observer also took written notes on verbal and nonverbal aspects of the situation. These sessions occurred approximately once a month for between one and four-and-a-half hours and were supplemented by diary entries by the child's mother wherein she recorded novel utterances and the context in which they occurred. However, only information from the formal observation sessions was used in the following analyses in order to make the data sources for the apes and human child as similar as possible.

Ape participants and their communication system

Panbanisha, a female bonobo, and Panpanzee, a female chimpanzee, were born about a month apart and raised together at the Language Research Center (LRC) for nearly the first four years of their lives. Data from the current study were collected when the apes were one to four years old. Although they had contact with Matata, Panbanisha's mother who also lived at the LRC, the young apes lived separately from her with human caregivers who communicated with them using lexigrams, gestures, and speech. As with human children, lexigram symbols were learned within the context of ongoing activities that were relevant to the apes (Brakke & Savage-Rumbaugh, 1996; Savage-Rumbaugh et al., 1993; Savage-Rumbaugh, Shanker, & Taylor, 1998).

Lexigrams are abstract visual images, each of which represents a word, designed to bear the same arbitrary relationship to its referent as most spoken words do (Savage-Rumbaugh et al., 1993). At any given time, approximately 256 lexigrams were portrayed in organized rows on the lexigram board. Because the lexigrams were selected based on the apes' interests, some lexigrams were removed and some added over the course of the study leading to slight variations in the number and type of available lexigrams. However, most lexigrams remained consistently available across the course of the study. Apes used gestures and vocalizations to enrich their symbolic communication (Greenfield & Savage-Rumbaugh, 1993, 1994). The human caregivers used English speech and, to a lesser extent, gestures to enrich their communication with lexigrams.

Although both apes eventually developed similar vocabularies, Panpanzee relied on gestures for almost a year after Panbanisha had become competent with lexigrams (Brakke & Savage-Rumbaugh, 1996). Panpanzee frequently used a technique whereby she guided a caregiver's hand to point to a lexigram or combined a gesture with leading a human by the hand.

Ape and human database

While the human data are not an attempt at a complete record of all utterances, the ape data represent an attempt to transcribe every communication that the apes made over the course of a number of years. NT's data consist of monthly observation sessions from 18 to 26 months of age. The ape data document all observed communications by the apes that occurred during eight hours each day for three years; beginning when they were approximately one year old and ending at approximately four years of age (except for missing information from three months of ape symbolic development (the 12th and 19th months of Panbanisha's life and the 18th month of Panpanzee's). The less complete nature of the human data will make it more difficult to find dialogic antecedents to symbolic combinations, even if they are there to the same extent as in the apes' symbolic development. Because communications directed at both apes and children were mainly recorded when it was considered relevant to understand their symbolic productions, we may have underestimated the frequency of deferred imitation across all species.

The system for creating a written record of ape conversation and productions was based on the system used for NT and MG by Greenfield and Smith (1976). Concordance between each ape communication and its context was determined and coded as each utterance was entered into the database, generally in the evening of the day the combination occurred. Concordance was defined, as it has been in many studies of human linguistic development, by whether symbol use matched the linguistic and nonlinguistic context, including whether the apes exhibited behavioral monitoring of conversational partners and correspondence between the symbol selected and subsequent behavior (Bloom, 1970; Brakke & Savage-Rumbaugh, 1996; Scollon, 1976). As in the cited studies of human language development, reliability coding was not conducted for concordance. Similarly, the linguistic and nonverbal contexts of human data were used to determine which human utterances were concordant with their contexts. Ape communications that were not concordant with context were given a numerical code indicating either error (incorrect use of a lexigram) or babbling (lexigram use which did not appear to serve a communicative function and had no relationship to the ongoing situational and action context). If it was not clear to the observer whether the communication was an error or not, it was coded with a number that indicated that its function was unclear.

All ape symbol combinations that were coded as errors or as having an unclear meaning were discarded from subsequent analysis. Similarly, we discarded human data if the real-time observer noted that the communication was not interpretable.

Combinations that appeared initially within strings that were coded at the time of data collection as babble were neither analyzed in the present study as novel meaningful combinations nor excluded from being considered as novel meaningful symbol combinations were they to recur later.

When five hours of real-time coding of the ape data were checked against video of the same time period, both were equally accurate in what they captured although the real-time coder missed nine communications that the video coder captured (Greenfield & Savage-Rumbaugh, 1994). Overlooked utterances make it more rather than less difficult to find evidence of antecedents for combinations in early dialogue.

Both the ape and human data have been examined in numerous studies by Greenfield, Savage-Rumbaugh, and many collaborators. However, this is the first time that the human and ape data have been integrated into a single database making quantitative comparisons possible. It is also the first time that the data have been used comparatively to investigate co-construction. The ape database is unique in terms of the symbolic competence of the apes, the co-rearing of a bonobo and a chimpanzee, and the quantity of data. The child data are especially comparable because the variables used to record the ape data were based on the methodology reported by Greenfield and Smith (1976).

Analysis

Because of the special role of verbs in the grammatical development of English-speaking children (Akhtar, 1999; Tomasello, 1992; Tomasello et al., 1997), our analysis focuses on combinations in which at least one element is an action word. In mature human language, action words are a subset of verbs and entity words are a subset of nouns. However, in early language, the categories of verb and noun do not include the abstractions of mature language and are thus labeled more concretely as representing actions or entities.

Despite differences in the media by which the child and the apes communicated, all three learners passed through a period of babbling – productions that were unrelated to the context in which they occurred – before producing symbolic elements and symbolic combinations that were concordant with the ongoing context. The apes began to use lexigrams communicatively at around the same age as human children begin to speak (around nine-and-a-half months for Panbanisha and one year for Panpanzee). Over the course of this study, the apes were observed to produce an average of 319 novel word combinations involving a verb while the human child was observed to produce 246 such combinations. Despite being observed approximately .02 times as frequently as the apes, the human child was observed to produce a similar number of novel word combinations. Given the apes' slower rate of symbolic development, data generated by the human child up to 26 months of age was compared to data generated by the apes up to four years of age.

For both children and apes, the first recorded instances of all novel (never before observed) and concordant symbol combinations that consisted of an action symbol (word or lexigram) paired with one or more other symbols were identified. This was accomplished by entering each action word (child's word or English gloss of lexigram) into the search box in the electronic database, sorting all instances of that action word by date, and then recording each combination as it occurred for the first time. Potential action words searched for were drawn from Tomasello (1992).

After identifying each novel symbol combination for each participant, human and ape, the database was searched to determine if modeled or joint constructions containing the symbols in each novel combination did or did not occur at some point prior to each novel concordant symbol combination (i.e. whether or not each combination had a conversational precedent). The data were sorted by date after entering the symbols (words in the case of the child, English gloss of lexigrams in the case of the apes) from the novel combination into all possible combinations of the data columns into which communications and their context were organized in order to determine if each combination of symbols had co-occurred in conversation (been present simultaneously within one or across multiple data columns) before the novice used them together independently. Our focus was on whether language learners use conversation to learn how to link words together rather than on whether discourse may help one acquire syntax. Thus, neither the order of words in the potential conversational precedents nor the presence or absence of intervening words that were not part of the novel combinations were considered when searching for conversational precedents.

Our definition of potential sources for deferred imitation or co-construction included all recorded conversations prior to the production of each novel combination. Once a potential conversational precedent for a novel word combination was identified, that combination was identified as having a potential precedent and the search for precedents for that combination was terminated. Thus, the amount of time elapsing between a potential conversational precedent and the first production of a novel word combination was irrelevant in our analyses.

In order to determine if word combinations became less likely to have precedents with development, we classified each novel symbol combination by the age of the individual when he or she first produced it. Thus, novel word combinations emitted by the apes during the last year of the study could have conversational precedents almost three years prior. Similarly, novel word combinations produced by NT during his 26th month could have conversational precedents up to eight months prior. This means that novel combinations emerging later in development would have had a greater chance of having a conversational precedent (i.e., a larger window of time in which a precedent could have been observed) if conversational precedents were equally likely across development. However, we hypothesized that conversational precedents would be more likely early in development even though the window of opportunity for observing them was smaller earlier than it would be later.

A greater number of novel word combinations were discarded for Panbanisha (110 out of 375) and Panpanzee (141 out of 263) relative to NT (29 out of 246) for being coded as babbling, incorrect, or lacking a clear meaning. Consequently, we wished to verify that the hypothesized pattern of decreased reliance on environmental input was specific to meaningful novel symbol combinations for the apes. We therefore focused on combinations produced by the apes that were coded as babbling and examined whether conversational precedents decreased across time for babble. We expected that conversational precedents to babbling would not decrease across time, thus distinguishing themselves from the development of meaningful combinations.

Results

Hypothesis 1: Qualitative findings

As shown in Tables 1, 2, and 3, different types of joint constructions (propositions across speakers, expansions, and joint constructions followed by vertical constructions respectively) occur prior to identical, but independently constructed word combinations for all three species. The last example for each action combination represents the first time that the learner was observed to produce that particular symbol combination without scaffolding by self or other.

Table 1. Joint construction (proposition across speakers) as antecedent to symbol combinations in child, bonobo, and chimpanzee

Joint construction: Human

Co-constructed model

NT (~age 18 months) pointed at a book: *Mommy.*

Mom: *Do you want a book? What do you want me to do with it?*

(Around seven months passed.)

First independent construction

NT (~age 25 months): He handed the cereal to his mom: *Mommy. Mommy do it.*

Mom: *You want me to do it?*

Joint construction: Bonobo

Co-constructed model

Panbanisha (~age 32 months) was being reprimanded after returning from running off: MILK SLEEP.

The caregiver told her that she had just finished her nap but she would HUG her.

(Around three months passed.)

First independent construction

Panbanisha (~age 35 months) vocalized happily as the caregiver prepared her milk for her nap. SLEEP HUG MILK.

Joint construction: Chimpanzee

Co-constructed model

The caregiver told Panpanzeee that she was going to OPEN the BACKPACK.

Panpanzee (~age 24 months): SURPRISE.

They did indeed find a surprise (the word for an unexpected treat) inside.

(Six days passed.)

First independent construction

A caregiver enters the room.

Panpanzee (~age 24 months) points at the door wanting to go into the other room to look for a surprise: OPEN SURPRISE.

They did not find a surprise in the other room but the caregiver took one from her pocket.

Note: Words that are entirely capitalized represent lexigrams. Each independent construction represents the first time that the learner was observed to produce that combination without scaffolding.

Table 2. Joint construction (expansion) as antecedent to symbol combinations in child, bonobo, and chimpanzee

Joint construction: Human

Co-constructed model

NT (~age 24 months) reached for the lid: *Lid. Lid*.

Mom: *He wanted to put the lid on*.

(Around one month passed.)

First independent construction

NT (~age 25 months). He had put water into a cup: *Lid on*.

Mom: *Do you want the lid?*

Joint construction: Bonobo

Co-constructed model

Panbanisha (~age 24 months) watched a caregiver and Panpanzee play with some soap. She visually searched for the lexigram for several minutes before she found it: SOAP.

The caregiver told her that she could go over and PLAY SOAP.

(Around one month passed.)

First independent construction

Panbanisha (~age 25 months) was having a physical: PLAY SOAP.

The caregiver got the soap out of the bag and Panbanisha played with it.

Joint construction: Chimpanzee

Co-constructed model

Panpanzee (~age 24 months): PLAY.

Elizabeth: PLAY HIDE *with* PANBANISHA *and* KAREN.

(Around nine months passed.)

First independent construction

Panbanisha was walking on the other side of the room with a hat.

Panpanzee (~age 33 months) charged toward her after touching: PLAY HIDE.

Note: Words that are entirely capitalized represent lexigrams. Each independent construction represents the first time that the learner was observed to produce that combination without scaffolding.

Table 3. Joint construction followed by vertical construction as antecedents to symbol combinations in child, bonobo, and chimpanzee

Joint construction: Human

Co-constructed model

NT (~age 24 months) put his car on the slope and rolled it down it: *Slope*.

Mom: *It went down the slope*.

(~An hour passed).

Vertical construction

NT placed a letter on the slope: *Down. Down*.

Mom: *mmh?*

NT: *Slope*.

Mom: *Down a slope? Yes, it's slipping down.*

(~A month passed).

First independent construction

NT (~age 25 months) pushed the truck down the edge of the couch: *Down the slope*.

Mom: *Down the slope*.

Table 3. (Continued)

Joint construction: Bonobo.
Co-constructed model
Panbanisha (~age 27 months): PLAY.
Human caregiver asked her to be more specific: TICKLE, BITE, HUG, GRAB?
Panbanisha: HUG.
(~Six months passed.)
Vertical construction
Panbanisha (~age 33 months) stood by the dog's gate: PLAY.
The human caregiver asked her to clarify her play request.
Panbanisha touched the fence: PLAY GRAB.
They entered to play with the dogs.
(Two days passed.)
First independent construction
The human caregiver is talking to Kanzi about tickling with Panbanisha.
Panbanisha: GRAB PLAY.
Panbanisha and Kanzi grab and tickle.
Joint construction: Chimpanzee
Co-constructed model
The caregiver offered Panbanisha some BANANA and GROOMING.
Panpanzee (~24 months): PLAY.
She was permitted to play for a few minutes.
(~Eight months passed.)
Vertical construction
The caregiver asked Panpanzee what she was thinking about.
Panpanzee (~32 months): PLAY
Caregiver: *No playing now.*
Panpanzee: BANANA
They ate a piece of banana together and got ready to move on.
(~Two months passed.)
Independent construction
Panpanzee (~34 months): BANANA PLAY.
It was near dinnertime. The caregiver suggested that they play first and then eat.

Note: Words that are entirely capitalized represent lexigrams. Each independent construction represents the first time that the learner was observed to produce that combination without scaffolding.

Table 4 illustrates that the ability to use deferred imitation of modeled phrases to construct symbol combinations independently at a later point in time is also present across the clade. In every example, the second time point for each example represents the first time that the learner was observed to produce each symbol combination. Note that the last example involves deferred imitation of a bonobo by a chimpanzee, whereas the prior child and bonobo examples involve deferred imitation of a human caregiver.

The existence of dialogic mechanisms for the construction of early symbol combinations across the clade confirms our first hypothesis.

Table 4. Modeled phrases as antecedents to symbol combinations in child, bonobo, and chimpanzee

Deferred imitation: Human

Model

Mom looked at NT in his high chair: *Do you want to get down?*

NT (~age 18 months): *Me-me*.

(Around five months passed.)

First independent construction

NT (~age 23 months) was about to climb down: *Get down*.

Mom: *Are you going to get down?*

Deferred imitation: Bonobo.

Model

Human caregiver: *You could* PLAY *with the* ORANGUTANs.

Panbanisha (~age 31 months) went over to the orangutans to play.

(Around five months passed.)

First independent construction

Panbanisha (~age 36 months): PLAY ORANGUTAN.

She was informed that she could play with orangutans later.

Deferred imitation: Chimpanzee

Model

Kanzi (a bonobo) touched symbols and then touched Panpanzee (~age 17 months): CHASE BITE GRAB.

They played together.

(Around 10 months passed)

First independent construction

Panpanzee (~age 27 months) was in the bedroom with her caregiver: CHASE BITE GRAB.

Her caregiver agreed and chased, bit, and grabbed her.

Note: Words that are entirely capitalized represent lexigrams. Each independent construction represents the first time that the learner was observed to produce that combination without scaffolding.

Hypothesis 2: Quantitative results

The qualitative examples described in Tables 1 through 4 are examples of a larger body of novel word combinations with conversational precedents. For Panbanisha, 92 novel symbol combinations were preceded by joint constructions, 81 by modeled phrases, and 202 had no observed precedent. For Panpanzee, 86 were preceded by joint constructions, 59 by modeled phrases, and 118 had no observed precedent. For NT, 32 novel symbol combinations were preceded by joint constructions, 44 by modeled phrases, and 170 had no observed precedent.

We now turn to testing the second hypothesis, that symbol combinations will become increasingly independent of dialogic origins with age in all three species. In order to do so, novel combinations were grouped by whether or not they had a conversational precedent (regardless of whether the precedent was a joint construction or a modeled phrase) and the age at which they were first produced independently.

The relevant samples of action combinations consisted of the first independent combination of every action symbol with every other symbol or sequence of symbols for child, chimpanzee, and bonobo. Quantitative analysis then tested the hypothesis that symbol combinations were more likely to be preceded by modeled or co-constructed phrases earlier in development than later. In order to test the prediction that combinations became more creative across developmental time in all three species, chi-square tests of independence were employed to compare the proportion of combinations with a conversational precedent across different time periods.

Prior research had demonstrated both the importance of conversational support for the transition to independent word combinations in children and the timing of this transition; between 17 and 23 months of age (Herr-Israel, 2006). Thus, for NT we compared combinations up to 23 months of age to those produced after 23 months of age. Combinations between 18 and 23 months of age were significantly more likely to have a precedent than those occurring between 24 and 26 months of age for NT (see Figure 1a; χ^2 (1) = 19.65, $p < .0001$).

Because novel word combinations increased more slowly in symbol-trained apes than in a human child, we compared combinations during the first three years of ape symbol production. For Panbanisha, the bonobo, and Panpanzee, the chimpanzee, combinations with a precedent were dominant between the first and second birthday; combinations without a precedent became dominant between the third and fourth birthday. This trend was statistically significant for Panbanisha (see Figure 1b; χ^2 (1) = 8.97, $p = .003$) and Panpanzee (see Figure 1c; χ^2 (1) 4.04, $p = .044$).

In order to verify that the observed pattern of increased creativity and decreased reliance on environmental input was specific to meaningful novel symbol combinations for the apes, we analyzed whether the number of potential precedents for word combinations coded as babbling decreased with age as they had for meaningful combinations. There were very few dialogic precedents for babble, and spontaneous babbling production decreased with age across the three years. This decrease in babbling contrasted

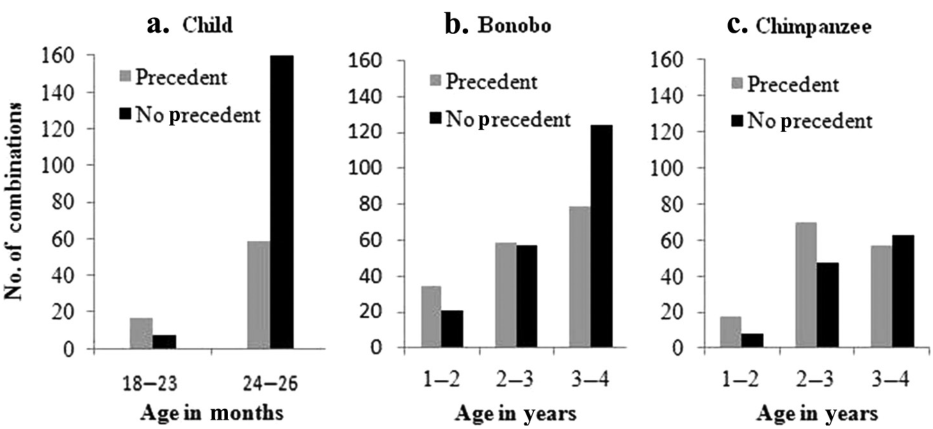

Figure 1. Number of novel symbol combinations with and without a source at different ages for human child, bonobo, and chimpanzee

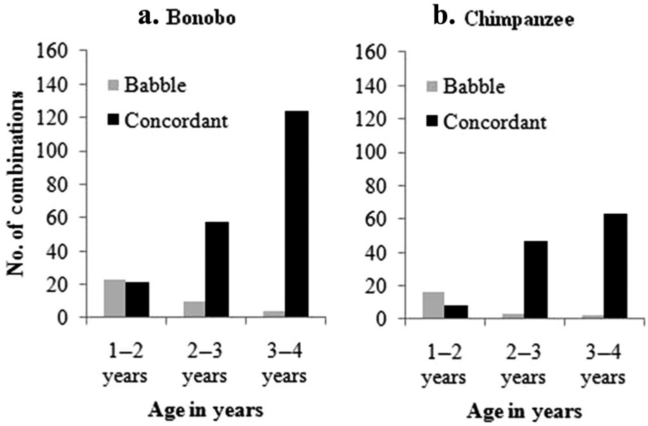

Figure 2. Changes with age in number of novel symbol combinations: babble vs. concordant with context (meaningful)

significantly with the increase in spontaneous concordant combinations over time for Panbanisha (see Figure 2a; χ^2 (2) = 60.47, $p < .0001$) and Panpanzee (see Figure 2b; χ^2 (2) = 60.314, $p < .0001$). The proportion of babbled combinations with a precedent relative to those without did not change significantly for either Panbanisha (χ^2 (1) = 0.173, $p = .68$) or Panpanzee (χ^2 (1) = 0.124, $p = .72$) across the three years, suggesting that the observed initial dependence on input followed by an increase in creativity was specific to combinations that were meaningful and communicative. These differential patterns suggest that the observed initial dependence on input followed by an increase in creativity was specific to combinations that were meaningful communications.

Discussion

Together, this pattern of qualitative and quantitative findings confirms our hypotheses. The qualitative evidence suggests that all three species use deferred imitation and other dialogic sources to generate early word combinations. The quantitative evidence suggests that over time representatives of all three species produce an increasing proportion of novel combinations without an observed dialogic source. This pattern occurs only for symbol combinations that are meaningful in context; it does not occur with meaningless babble that is unrelated to the situational context. This sequence of steps provides cladistic evidence for Vygotsky's (1962) developmental progression from inter-individual to intra-individual competence.

The holistic strategy of generating word combinations from deferred imitation of conversational precedents may provide information for an analytic strategy wherein key words commonly used in combinations in the learner's environment can be paired productively with multiple nouns (Ninio, 2001). The words 'yes' and 'no' are often used as key words by human children. We considered changes in the use of 'yes' and 'no' with development.

For safety reasons, the language-competent apes were more dependent on their caregivers to grant permission for them to engage in desired activities than many human children are. Thus, they frequently were told either 'yes' or 'no' when they requested an activity. Interestingly, none of Panbanisha's concordant word combinations over the course of the study included the word 'no' and only three of Panpanzee's did, all of which had prior conversational precedents. However, both of them began to use 'yes' productively over the course of the study. Panbanisha used 'yes' in a meaningful word combination once between two and three years of age (this first combinatorial use of 'yes' had an observed precedent) and 46 times between three and four years of age (40% of these later combinations involving 'yes' had an observed precedent). Panpanzee used 'yes' 19 times from one to three years of age (74% of these had an observed precedent), and 24 times from three to four years of age (50% of these later combinations had an observed precedent). NT was not observed to use 'yes' in a meaningful novel combination over the course of the study. He used 'no' four times (75% of which had a precedent) up to 23 months of age and two times (neither of which had a precedent) after 23 months of age. These findings suggest that the ability to pair key words with other words may initially be learned through deferred imitation and become more productive with development.

Another potential way that individuals may scaffold their own development, albeit with some conversational support, is through vertical constructions. Whenever a novel word combination was first produced via a vertical construction and only later via a modeled or joint construction, that construction was considered to not have an observed conversational precedent because the child initially supplied all of the elements. However, if a conversational precedent preceded a vertical construction which in turn preceded the independent production of a novel symbol combination, that combination was viewed as having an observed conversational precedent. This pattern can be seen in Table 3 and illustrates the general principle that children use many strategies besides or in addition to deferred imitation, such as vertical scaffolding and joint attention (Tomasello, 1988) to learn language. However, it is very much in line with the Vygotskian progression from inter-individual to intra-individual construction that our qualitative analysis indicates that social scaffolding precedes self-scaffolding.

The greater number of novel word combinations that were excluded as errors or babbling for the apes relative to the humans represents a potential limitation of our study. However, research examining the nature of the errors apes make during vocabulary tests suggests that their errors are not random (Lyn, 2007). Many errors are driven by proximity wherein the ape selects a symbol that is physically close to the symbol she meant to select. Thus, proximity may have influenced which word combinations were generated, particularly for combinations discarded as errors. However, the other most frequent type of error is categorical equivalence, or selecting a symbol from the same category as the desired symbol (i.e., bananas rather than blackberries). Auditory or physical similarity between the correct lexigram and the incorrect one also influences error rates. The multiple factors underlying the errors apes make imply that their erroneous combinations are not random but are influenced by a multiplicity of dimensions.

Another potential limitation of this study is that a visual lexigram board upon which all choices are immediately available may require fewer memory resources than the internal lexicon that a child accesses in order to speak. While the lexigram board may indeed reduce memory demands, it also requires that the apes use cognitive flexibility to select the best available option to express their desires from among limited options (Brakke & Savage-Rumbaugh, 1996). Additionally, they often had to keep the referent in mind while locating the desired lexigram, as documented in Table 2 when Panbanisha searched for the symbol for 'soap.'

An important difference between the species is that the rate at which each of our participants learned to construct novel combinations independently was different. Not only did the child produce more independent then dependent novel combinations earlier in life than either ape, but the chimpanzee appears to have continued to rely more on conversational sources than did the bonobo throughout the study period, although this difference is not statistically significant. Given the small sample size, it is unclear if this difference between apes in the rate of producing creative symbol combinations represents a difference between bonobos and chimpanzees more generally. Although both apes eventually acquired comparable lexicons, Panbanisha had a larger productive vocabulary than Panpanzee at approximately three-and-a-half years of age (Brakke & Savage-Rumbaugh, 1996). This is noteworthy because the emergence of multi-word utterances is positively related to lexicon size (Bates & Goodman, 1997). A larger sample could unravel this question of bonobo–chimpanzee species differences, but such a sample does not currently exist.

This study demonstrates the importance of dialogic mechanisms for the early emergence of symbolic combinations across the clade. The results imply that conversational capabilities were likely present in our common ancestor and may have played an important role in the evolution of language. We now see why earlier debates about which species imitates and which does not could never be resolved: imitation, like conversational co-construction, does not differentiate between bonobo, chimpanzee, and human; instead, imitation and co-construction differentiate stages of learning and development across all three species.

Acknowledgements and funding

Data analysis and article preparation were supported by a grant from the Leakey Foundation to Patricia Greenfield and by Summer Research Mentorships from UCLA Graduate Division to Kristen Gillespie-Lynch. Ape data collection took place at the Language Research Center, Georgia State University, and was supported by grants from NICHD to Duane Rumbaugh and Sue Savage-Rumbaugh. Child data collection was supported by the Harvard Center for Cognitive Studies. Goldie Salimkhan played an integral role in the current project by transferring all of the human data from written records into a shared electronic database with the ape data.

References

Abbot-Smith, K., & Tomasello, M. (2010). The influence of frequency and semantic similarity on how children learn grammar. *First Language, 30*, 79–101.

Akhtar, N. (1999). Aquiring basic word order: Evidence for data driven learning of syntactic structure. *Journal of Child Language, 26*, 339–356.

Aksu-Koc, A. (1998), The role of input vs. universal predispositions in the emergence of tense-aspect morphology: Evidence from Turkish. *First Language, 18*, 255–280.

Andersen, E. S., Dunlea, A., & Kekelis, L. (1993). The impact of input: Language acquisition in the visually impaired. *First Language, 13*, 23–49.

Bates, E., & Goodman, J. C. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia, and real time processing. *Language and Cognitive Processes, 12*, 507–584.

Bjorklund, D. F., Bering, J. M., & Ragan, P. (2000). A two-year longitudinal study of deferred imitation of object manipulation in a juvenile chimpanzee *(Pan troglodytes)* and Orangutan *(Pongo pygmaeus). Developmental Psychobiology, 37*, 229–237.

Bjorklund, D. F., Yunger, J. L., Bering, J. M., & Ragan, P. (2002). The generalization of deferred imitation in enculturated chimpanzees (*Pan troglodytes*). *Animal Cognition, 5*, 49–58.

Bloom, L. (1970). *Language development: Form and function in emerging grammars.* Cambridge, MA: MIT Press.

Bloom, L. (1973). *One word at a time: The use of single word utterances before syntax.* The Hague: Mouton.

Bloom, L., Hood, L., & Lightbown, P. (1974). Imitation in language development: If, when, and why. *Cognitive Psychology, 6*, 380–420.

Brakke, K. E., & Savage-Rumbaugh, S. (1996). The development of language skills in *Pan* II: Production. *Language and Communication, 16*, 361–380.

Brown, R. (1973). *A first language.* Cambridge, MA: Harvard University Press.

Buttelmann, D., Carpenter, M., Call, J., & Tomasello, M. (2007). Enculturated chimpanzees imitate rationally. *Developmental Science, 10*, 31–38.

Caselli, M. C., & Volterra, V. (1990). From communication to language in hearing and deaf children. In V. Volterra & C. J. Erting (Eds.), *From gesture to language in hearing and deaf children* (pp. 263–277). Berlin & Heidelberg: Springer-Verlag.

Chomsky, N. (1959). Verbal behavior. *Language, 35*, 26–58.

Clark, R. (1974). Performing without competence. *Journal of Child Language, 1*, 1–10.

Corrigan, R. (1980). Use of repetition to facilitate spontaneous language acquisition. *Journal of Psycholinguistic Research, 9*, 231–240.

Elsen, H. (1996). Two routes to language: Stylistic variation in one child. *First Language, 16*, 141–158.

Greenfield, P. M., & Lyn, H. (2006). Symbol combination in *Pan*: Language, action, and culture. In D. Washburn (Ed.), *Primate perspectives on behavior and cognition* (pp. 255–267). Washington, DC: American Psychological Association.

Greenfield, P. M., Reilly, J., Leaper, C., & Baker, N. (1985). The structural and functional status of single-word utterances and their relationship to early multi-word speech. In M. Barrett (Ed.), *Children's single word speech* (pp. 233–267). London: John Wiley.

Greenfield, P. M., & Savage-Rumbaugh, S. (1993). Comparing communicative competence in child and chimp: The pragmatics of repetition. *Journal of Child Language, 20*, 1–26.

Greenfield, P. M., & Savage-Rumbaugh, S. (1994). Grammatical combination in *Pan paniscus*: Processes of learning and invention in the evolution and development of language. In S. Parker & K. Gibson (Eds.), *Language and intelligence in monkeys and apes: Comparative developmental perspective* (pp. 540–578). Cambridge: Cambridge University Press.

Greenfield, P. M., & Smith, J. H. (1976). *The structure of communication in early language development.* New York: Academic Press.

Heimann, M., Strid, K., Smith, L., Tjus, T., Ulvund, S. E., & Meltzoff, A. N. (2006). Exploring the relation between memory, gestural communication, and the emergence of language in infancy: A longitudinal study. *Infant and Child Development, 15*, 233–249.

Herr-Israel, E. (2006). *Single words to combinations: Longitudinal analysis reveals constructive processes integrating lexical and pragmatic development within a conversational context.* Unpublished doctoral dissertation, Rutgers University.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology, 27*, 236–248.

Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology, 45*, 337–374.

Jones, E. J. H., & Herbert, J. S. (2006). Exploring memory in infancy: Deferred imitation and the development of declarative memory. *Infant and Child Development, 15*, 195–206.

Kuczaj, S. A. (1987). Deferred imitation and the acquisition of novel lexical items. *First Language, 7*, 177–182.

Kuczaj, S. A., Paulos, R. D., & Ramos, J. A. (2005). Imitation in apes, children, and dolphins: Implications for the ontogeny and phylogeny of symbolic representation. In L. I. Namy (Ed.), *Symbol use and symbolic representation.* Mahwah, NJ: Lawrence Erlbaum.

Leonard, L. B., Schwartz, R. G., Folger, M. K., Newhoff, M., & Wilcox, M. J. (1979). Children's imitations of lexical items. *Child Development, 50*, 19–27.

Lieven, E. V., Pine, J. M., & Baldwin, G. (1997). Lexically based learning and early grammatical development. *Journal of Child Language, 24*, 187–219.

Lyn, H. (2007). Mental representation of symbols as revealed by vocabulary errors in two bonobos *(Pan paniscus). Animal Cognition, 10*, 461–475.

Lyn, H., Greenfield, P. M., & Savage-Rumbaugh, S. (2006). The development of representational play in chimpanzees and bonobos: Evolutionary implications, pretense, and the role of interspecies communication. *Cognitive Development, 21*, 199–213.

Matthews, D., Lieven, E., Theakston, A., & Tomasello, M. (2005). The role of frequency in the acquisition of English word order. *Cognitive Development, 20*, 121–136.

Meltzoff, A. N. (1995). What infant memory tells us about infantile amnesia: Long-term recall and deferred imitation. *Journal of Experimental Child Psychology, 59*, 497–515.

Ninio, A. (2001). Pragmatic keywords and the first combining verbs in children's speech. *First Language, 21*, 433–460.

Ochs, E., Schieffelin, B. B., & Platt, M. L. (1979). Propositions across utterances and speakers. In E. Ochs & B. B. Schieffelin (Eds.), *Developmental pragmatics.* New York: Academic Press.

Pedersen, J. & Fields, W. M. (2009). Aspects of repetition in bonobo-human conversation. *Integrative Psychological and Behavioral Science, 43*, 22–41.

Piaget, J. (1962). *Play, dreams and imitation.* New York: Norton.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926–1928.

Savage-Rumbaugh, S. (1990). Language as a cause-effect communication system. *Philosophical Psychology, 3*, 55–76.

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

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

Schwartz, R. G., Chapman, K., Terrell, B. Y., Prelock, P., & Rowan, L. (1985). Facilitating word combination in language-impared children through discourse structure. *Journal of Speech and Hearing Disorders, 50*, 31–39.

Schwartz, R. G., & Terrell, B. Y. (1983). The role of input frequency in lexical acquisition. *Journal of Child Language, 10*, 57–64.

Scollon, R. (1976). *Conversations with a one year old.* Honolulu: The University of Hawai'i Press.

Skinner, B. F. (1957). *Verbal behavior*. Acton, MA: Copley Publishing Group.

Snow, C. E., & Goldfield, B. A. (1983). Turn the page please: Situation-specific language acquisition. *Journal of Child Language, 10*, 551–569.

Snow, C. E., Perlmann, R., & Nathan, D. (1987). Why routines are different: Toward a multiple factors model of the relation between input and language acquisition. In K. E. Nelson & A. van Kleeck (Eds.), *Children's language* (Vol. 6, pp. 65–97). Hillsdale, NJ: Lawrence Erlbaum.

Terrace, H. S., Petitto, L. A., Sanders, R. G., & Bever, T. G. (1979). Can an ape create a sentence? *Science, 206*, 891–902.

Tomasello, M. (1988). The role of joint attentional processes in early language development. *Language Sciences, 10*, 69–88.

Tomasello, M. (1992). *First verbs*. Cambridge, England: Cambridge University Press.

Tomasello, M. (2000). The item based nature of children's early syntactic development. *Trends in Cognitive Sciences, 4*, 156–163.

Tomasello, M., Akhtar, N., Dodson, K., & Rekau, L. (1997). Differential productivity in children's use of nouns and verbs. *Journal of Child Language, 24*, 373–387.

Tomasello, M., Brooks, P. J., & Stern, E. (1998). Learning to produce passive utterances through discourse. *First Language, 18,* 223–237.

Tomasello, M., Davis-Dasilva, M., Camak, L., & Bard, K. (1987) Observational learning of tool use by young chimpanzees. *Human Evolution, 2,* 175–183.

Tomasello, M., Savage-Rumbaugh, S., & Kruger, A.C. (1993). Imitative learning of actions on objects by children, chimpanzees, and enculturated chimpanzees. *Child Development, 64*, 1688–1705.

van Veen, R., Evers-Vermeul, J., Sanders, T., & van den Bergh, H. (2009). Parental input and connective acquisition: A growth curve analysis. *First Language, 29*, 266–288.

Veneziano, E. (2005). Effects of conversational functioning on early language acquisition: When both caregivers and children mattern. In B. Bokus (Ed.), *Studies in the psychology of child language* (pp. 47–69). Warsaw: Matrix.

Veneziano, E., Sinclair, H., & Berthoud, I. (1990). From one word to two words: repetition patterns on the way to structured speech. *Journal of Child Language, 17*, 633–650.

Vygotsky, L. (1962). *Thought and language*. Cambridge, MA.: MIT Press.

Weizman, Z. O., & Snow, C. E. (2001). Lexical input as related to children's vocabulary acquisition: Effects of sophisticated exposure and support for meaning. *Developmental Psychology, 37*(2), 265–279.

Youssef, V. (1994). 'To be or not to be': formulaic and frame-based acquisition of the copula in Trinidad. *First Language, 14*, 263–282.