Neuroscience and Biobehavioral Reviews 81 (2017) 150–157

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

journal homepage: www.elsevier.com/locate/neubiorev

Review article

Statistical evidence that a child can create a combinatorial linguistic system without external linguistic input: Implications for language evolution

Susan Goldin-Meadow[a,*], Charles Yang[b]

[a] University of Chicago, Departments of Psychology and Comparative Human Development, 5848 South University Avenue, Chicago, IL 60637, United States
[b] University of Pennsylvania, Departments of Linguistics and Computer Science, 619 Williams Hall, Philadelphia, PA 19081, United States

ARTICLE INFO

Article history:
Received 28 February 2016
Received in revised form 28 October 2016
Accepted 19 December 2016
Available online 29 December 2016

Keywords:
Language development
Computational linguistics
Linguistic input
Homesign
Sign language

ABSTRACT

Can a child who is not exposed to a model for language nevertheless construct a communication system characterized by combinatorial structure? We know that deaf children whose hearing losses prevent them from acquiring spoken language, and whose hearing parents have not exposed them to sign language, use gestures, called *homesigns*, to communicate. In this study, we call upon a new formal analysis that characterizes the statistical profile of grammatical rules and, when applied to child language data, finds that young children's language is consistent with a productive grammar rather than rote memorization of specific word combinations in caregiver speech. We apply this formal analysis to homesign, and find that homesign can also be characterized as having productive grammar. Our findings thus provide evidence that a child can create a combinatorial linguistic system without external linguistic input, and offer unique insight into how the capacity of language evolved as part of human biology.

Contents

---

* Corresponding author.
E-mail address: sgm@uchicago.edu (S. Goldin-Meadow).

http://dx.doi.org/10.1016/j.neubiorev.2016.12.016
0149-7634/© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

There is no doubt that language evolved as a biological capacity (Hauser et al., 2014). As a complex trait that emerged as recently as 100,000 years ago (Tattersall 2012), language must have been integrated within the broad human cognitive system, parts of which are shared with other species and lineages. But to really understand how language evolved in the extremely brief history of Homo Sapiens, we need to identify the defining characteristics of language (Hauser et al., 2002).

It is widely acknowledged that *a* (if not *the*) hallmark of language is the combinatorial use of a finite inventory of linguistic units—phonemes, morphemes, words, etc.—to form an infinite range of expressions (Chomsky 1965; Berwick and Chomsky 2016). How children acquire a combinatorial grammar has been viewed as “the most promising guide to what happened in language evolution” (Hurford 2012:590). In a recapitulationist turn, the development of child language is interpreted as retracing the steps of language evolution (Bickerton 1995; Studdert-Kennedy 1998). Young children’s language has, in fact, been likened to signing in non-human primates–both display limited combinatorial use of grammar, which is assumed to be nothing more than rote learning in children (Tomasello 2000,2003) and chimps (Terrace et al., 1979). But this characterization of child language is controversial and has been challenged on theoretical and empirical grounds (e.g., Fisher 2002; Lidz et al., 2003; Valian et al., 2009).

In this paper, we probe the question of language evolution by studying a particular type of language development–homesign, a gestural system created by deaf children whose hearing losses prevent them from learning spoken language, and whose hearing parents have not exposed them to sign language (Goldin-Meadow and Feldman, 1977; Feldman et al., 1978; Goldin-Meadow, 2003). These children do not have an adult model for language and thus the combinations they produce cannot be the product of rote learning. The study of homesign thus affords a unique opportunity to investigate the emergence of a combinatorial linguistic system de novo. Here we apply a statistical benchmark that characterizes the profile of grammatical rules (Yang 2013) to homesign. Using this rigorous procedure, we find that homesign has productive grammar and we discuss the evolution of language in light of these findings.

2. The nature of homesign and child language

Children are able to learn any and all languages that span the globe. Even if exposed to a signed (as opposed to a spoken) language, children learn that language with equal ease (Lillo-Martin 1999; Newport and Meier 1985). The fact that children deal easily with such a broad span of linguistic inputs is often taken to suggest that they come to language learning with few expectations about what kind of system they are about to learn. This hypothesis is supported, at least circumstantially, by the fact that even at the earliest stages of language learning, children have successfully acquired many subtle features of the native language to which they are exposed. For example, English and Korean languages present children with very different ways of talking about joining objects. But children as young as 17 months have no trouble learning to talk about joining objects in terms of *containment* in English and in terms of *fit* in Korean (Choi and Bowerman 1991).

However, just because children catch on to the quirks of the language they are learning very early in development does not mean that they approach language learning without biases or predispositions. But it does mean that discovering those predispositions is going to be extremely difficult, precisely because the effect of a language model on language learning is both massive and early. Perhaps the best way to determine whether children bring their own biases to language learning is to observe children *before* they have been exposed to a language model. This is not easy to do.

It is almost impossible to find a child raised by human parents who has not been bathed in language. However, there are children who are unable to profit from the sea of language that surrounds them–congenitally deaf children whose hearing losses are so severe that they cannot profit from the surrounding speech, even with hearing aids and intensive instruction. If these deaf children are born to hearing parents who do not expose them to a conventional sign language until adolescence, the children are effectively deprived of a usable language model during the early language learning years. They consequently present us with the opportunity to uncover biases that children bring with them to language learning.

Goldin-Meadow and colleagues (Goldin-Meadow 2003; Goldin-Meadow and Feldman 1977; Feldman et al., 1978; Goldin-Meadow et al., 2015) have studied deaf individuals under these circumstances, and found that they are able to produce self-styled gestural communication systems–*homesigns*–that are similar to natural language in many respects. The structures that these individuals convey in their homesigns have not been influenced by a conventional language model, nor are they modeled after the co-speech gestures that hearing individuals use when interacting with them (Goldin-Meadow and Mylander 1983, 1984, 1998; Goldin-Meadow et al., 2007). These structures therefore come as close as we can currently envision to revealing the human child’s predispositions to communicate in a structured way.

In fact, many of the properties hypothesized to be central to human language, and thus good candidates for structures that children themselves bring to language, have been found in homesign. For example, lexical markers that modulate the meanings of sentences (negation and questions, Franklin et al., 2011), grammatical categories such as nouns and verbs (Goldin-Meadow et al., 1994) and subjects (Coppola and Newport 2005), and devices that refer to non-present events (i.e., that allow displacement, Butcher et al., 1991; Morford and Goldin-Meadow, 1997) are all properties of homesign. Importantly, the gestures in homesign are also composed of parts, akin to a morphological system (Goldin-Meadow et al., 2007), and those gestures combine to form structured sentences, akin to a syntactic system (Feldman et al., 1978). Homesign is thus characterized by levels of structure, and those levels are organized hierarchically. For example, homesigners use multi-gesture combinations–a demonstrative gesture plus a noun gesture–to serve the same semantic and syntactic functions as either the demonstrative gesture or the noun gesture used on its own. The larger unit can thus substitute for the smaller units and, in this way, functions as a complex nominal constituent embedded within a sentence (i.e., a sentence with hierarchical structure [[[that] [bird]] [pedals]], rather than flat structure [[that] [bird] [pedals]] (Hunsicker and Goldin-Meadow, 2012; Flaherty et al., 2017). The properties of language that crop up in homesign can be developed without input from a conventional language model and, in this sense, are good candidates for innate properties of language.

The strategy used to argue that homesign is structured like a language has been to compare homesign to the speech learned by hearing children and to the signs learned by deaf children who *are* exposed to a conventional language model. For example, early in development, homesigners produce gesture sentences that convey the same set of semantic relations that are found in the early communications of young hearing children (Goldin-Meadow and Mylander, 1984) and they structure those sentences just as young hearing children structure their early sentences (e.g., by following simple word order patterns, by selectively omitting and producing certain semantic elements, by conjoining propositions, Goldin-Meadow and Mylander 1984). Recently, however, the grammatical foundation of hearing children’s early productions has been called into question, which then calls into question the grammatical foundation of homesign.

As noted earlier, a hallmark of human language is the combinatorial use of words to create an unbounded number of meaningful expressions. But young children display limited and formulaic combinatorial flexibility in their early productions (Tomasello 2003). Take, for example, determiner-noun combinations. In English, singular nouns can interchangeably follow the singular determiners "a" and "the": "a book", "the book", "a dog", "the dog", etc. If every noun that a child produces follows "a" and also follows "the", the diversity measure for that child is 100%; if, however, the child produces one set of nouns after "a", and a different set of nouns after "the", the child's diversity measure is 0%. The syntactic diversity of determiner-noun combinations in young English-learners is, in fact, quite low: only 20–40% of singular nouns in child speech appear with both determiners; the rest appear with one determiner exclusively (Pine and Lieven 1997). Similar observations of low diversity have also been made about young homesigners' early productions.

According to some researchers, young children do not have an abstract grammatical system (Tomasello 2003). Low syntactic diversity in child language is due to the memorization of specific word combinations in adult language. That is, if the caretaker has not had the social and pragmatic occasion to combine a specific determiner and a specific noun (e.g., "a zoo", where "the zoo" is much more common as in "go to the zoo"), the child would not do so either. However, there are two problems with this interpretation of low syntactic diversity in child language.

First, the syntactic diversity in adult language is, paradoxically, also very low. A recent study found no difference in determiner-noun combinations between individual children and their mothers (Valian et al., 2009). In the Brown Corpus (Kucera and Francis 1967), a collection of professional writing, the determiner-noun combination diversity is only 25%, lower than some two-year-olds (cf. Pine and Lieven 1997). Yet the grammatical ability of mothers and professional writers is not in doubt (but see Pine et al., 2013, and a response by Yang and Valian 2014).

Second, the study of homesign calls for an assessment of syntactic diversity in the absence of comparable adult languages. Homesign is a deaf child's own system of gestures, one which is *not* modeled after a conventional sign language (since the child does not have access to such a model), nor is it copied from the co-speech gestures that hearing adults produce when interacting with the deaf child (Goldin-Meadow and Mylander 1983, 1984, 1998; Goldin-Meadow et al., 2007; Hunsicker and Goldin-Meadow 2012, Flaherty et al., 2017).

In this study, we apply a recently developed statistical test for syntactic diversity (Yang 2013) to homesign. We ask whether the gesture sentences produced by a homesigner are limited to a closed repertoire, or whether they are consistent with a grammar that combines independent and interchangeable linguistic units.

3. A statistical model for assessing the combinatorial diversity of a language sample

To assess the linguistic system underling production data, it is important to develop a well-formed hypothesis that connects an abstract grammatical system to the numerical measure of syntactic diversity. This measure must be applicable to the production data, whether the data are generated with or without access to an adult language model. Yang (2013) has recently developed such a statistical test, one that can detect the presence or absence of grammatical rules within a linguistic corpus.

Consider a syntactic combination that has two units, C and O. In our study, C generally has a small number of values and is thus similar to a closed-class item, whereas O has a wide range of values thus resembling an open-class item. For instance, in the determiner-noun case mentioned earlier, C is either "a" or "the", a fixed inventory; O is a singular noun of which there are many instances ("car", "dog", "chair", etc.). As a second example, consider syntactic predicates, focusing on verbs. C can be either a transitive or an intransitive verb; O is an open-ended noun phrase serving as an argument of the verb. We are interested in whether the two units in a syntactic combination can be used interchangeably in the production corpus of language users––in these examples, whether "a" and "the" (C items) are both found in combination with a given noun (O items); or whether transitive and intransitive verbs (C items) are both found in combination with a given noun phrase (O items).

If the combination of C and O is statistically independent, we can compute their joint probabilities by multiplying their marginal probabilities. To do so, we must take into account a very general statistical property of language known as Zipf's (1949) Law, which has been widely recognized in the quantitative study of languages (Baroni 2009). A characteristic pattern of Zipf's Law is that many words in a linguistic corpus appear infrequently, often only once, as they reside on a long tail. That is, in a corpus of C—O combinations, many of the O's will appear only once and therefore cannot be combined with multiple members of C. This factor contributes to the low level of syntactic diversity observed in previous research.

There are, however, other factors. Empirically, the probabilities with which the two forms of a C-type item combine with an O-type item are highly imbalanced. For instance, for the noun "bathroom", the determiner "the" is favored over "a", as there are more instances of "the bathroom" than "a bathroom" in English corpora; but for the noun "bath", the opposite pattern is true––there are more instances of "a bath" than "the bath". Likewise, empirically it turns out that the noun phrase "ice cream" is produced far more often in transitive predicates (e.g., "eat/lick/buy/want ice cream") than in intransitive predicates (e.g., "ice cream falls/melts/spoils"). If the combination of C and O is likened to tossing a coin repeatedly, where syntactic diversity is measured by the mixture of heads and tails, then the coin is heavily biased toward one of two sides. The imbalance of syntactic combinations, which can be empirically quantified, also contributes to the low diversity in language use.

Yang (2013) gives the expected value of syntactic diversity if the two items C and O can combine independently and interchangeably. The calculation makes use of Zipf's (1949) Law–the notion that the frequency and rank of words are inversely proportional. That is, if a sample consists of N distinct word types, then the probability of the r-th ranked word is:

$$p_r = \frac{1}{rH_N} \text{ where } H_n = \sum_{i=1}^{N} \frac{1}{i}$$

which allows us to approximate the probabilities of open class items in O. In a sample of S pairs of C—O combinations, the expected probability of the r-th ranked item in O, combined with both elements in C, is:

$$E_r = 1 - (1 - p_r)^S - \sum_{i=1}^{2} [(f_i\, p_r + 1 - p_r)^S - (1 - p_r)^S] \qquad [1]$$

where f_i is the probability of the i-th C-type item combining with the r-th ranked O-type item. Without loss of generality, we assume f_1 to be the probability of the more frequently paired C for O (e.g., the probability of *the* for the noun "bathroom", and the probability of a transitive verb for the noun phrase "ice cream", as discussed earlier). The average diversity across all of the items in O is:

$$\frac{1}{N}\sum_{r=1}^{N} E_r$$

This formulation allows us to calculate the expected value of syntactic diversity using three empirical values of a linguistic corpus: The total sample size of C—O combinations (S); the total number of different types of O items (N); and the bias (B), the average value of f_1 over all Os in the corpus, which empirically characterizes the degree of imbalance in the C—O combinations. Previous work has established the validity of this formulation in linguistic samples taken from both adults and children. For example, in the Brown Corpus (Kucera and Francis 1967), a collection of English print materials, the syntactic diversity measure for determiner-noun combinations is surprisingly low (25.2%), but not statistically different from the expected value (26.5%) (Yang 2013). Similarly, in language samples taken from children learning American English at the two-word stage, the syntactic diversity measure for determiner-noun combinations was also low (average: 30.6%) and again statistically indistinguishable from the expected value (average: 31.4%) (Yang 2013). This finding suggests that children do follow an abstract rule that combines words independently.

The statistical test for grammar needs to be robust to serve as an appropriate benchmark for the investigation of linguistic systems. It should not only identify the presence of grammar in a corpus where we might expect productive syntactic combinations, but it should also be able to identify the absence of grammar in cases where the combination of linguistic units is not independent and the units are not interchangeable. Yang (2013) applied the test to the sign combinations produced by one of the language-trained chimpanzees, Nim Chimpsky (Terrace et al., 1979). The test showed that Nim's combinatorial diversity is statistically significantly lower than the diversity expected of a rule-based grammar, corroborating results from a quantitative analysis of Nim's videos carried out by his trainers (Terrace et al., 1979).

Taken together, there is considerable evidence that the statistical test developed by Yang (2013) is suitable for detecting the presence or absence of a combinatorial system that underlies a linguistic corpus. We now apply the test to detect the degree of combinatorial diversity in homesign.

4. Empirical tests for combinatorial diversity in homesign

In the present study, we evaluate the syntactic abilities of a deaf homesigner called "David" (Goldin-Meadow and Mylander 1984; Goldin-Meadow 1979). The statistical test just reviewed, which provides a benchmark for quantitative assessments of syntactic combinations, is particularly suitable for our purposes because there is no target adult system against which to compare homesign, either a conventional linguistic system or an idiosyncratic gestural system (as noted earlier, the co-speech gestures that the homesigners' hearing parents produce when interacting with their children do not form an adequate model for homesign, Goldin-Meadow and Mylander 1983, 1984, 1998, Flaherty et al., 2017, and thus cannot serve as a comparative baseline). We apply the statistical procedure to a sample of homesigns previously found to display syntactic structure using behavioral analyses (e.g., Feldman et al., 1978; Goldin-Meadow 1979; Goldin-Meadow and Mylander 1983, 1984; Goldin-Meadow et al., 1994; Hunsicker and Goldin-Meaodw 2012). We thus determine whether these syntactic descriptions can withstand a more stringent statistical test of combinatoriality.

4.1. Transcribing and coding gestures

4.1.1. Identifying, parsing, and categorizing gestures

David was videotaped in his home during interactions with his family members and the experimenters every two or three months between the ages of 2;10 and 5;02 (years;months); 11 sessions, each lasting approximately two hours, were analyzed. The experimenters brought the same set of books, toys, and puzzles to elicit communication to each session. Coders who had not been present at the session had access to these items in the lab and could use them to contextualize the child's gestures. In addition, when the experimenters were uncertain about the meaning of David's gesture, they asked his parents to clarify; those conversations were part of the videorecording and thus accessible to coders.

We used two criteria to identify a gesture: the hand or body movement had to be communicative in intent (i.e., produced when the child had another's attention), but was not a functional act on an object or person. For example, reaching to pick up a toy communicates the child's desire for a toy but it does so by directly acting on the world, and was therefore not considered a gesture. In contrast, an open palm held out flat (a GIVE gesture), produced while making eye contact with the person holding the toy, communicates a request for the toy indirectly and so was considered a gesture.

Once isolated, gestures were coded along the three dimensions used to describe signs in conventional sign language: shape of the hand, location of the hand with respect to the body, and movement of the hand. A change in any one of these dimensions during the stroke of the gesture was taken to signal the end of one gesture and the beginning of another. Motoric criteria were also used to determine the end of a string of gestures and thus sentence boundaries. Two gestures were considered separate sentences if the child paused or relaxed his hands between the gestures. Gestures that were not separated by pause or relaxation of the hands were considered part of the same sentence (see Goldin-Meadow and Mylander 1984 for additional details).

Homesigners produce three different types of gestures: deictic gestures, iconic gestures, and markers. Deictic gestures refer to objects by pointing to, or holding up, the intended referent and can be used to refer to any entity that is present (and, in some cases, entities that are not present, Butcher et al., 1991). Iconic gestures represent an aspect of an object or action through pantomime (e.g., moving two fists as though beating a drum, BEAT) or visual depiction (e.g., forming a circle with the thumb and index finger, ROUND). An iconic gesture can be used as a noun (e.g., when the BEAT gesture is used to identify a *drum*; when the ROUND gesture is used to identify a *penny*), a verb (e.g., when the BEAT gesture is used to refer to the act of beating the drum, *beat*), or an adjective (e.g., when the ROUND gesture is used to comment on the shape of the penny, *round*); see Goldin-Meadow et al. (1994) for criteria used to distinguish these uses. Markers are typically conventional gestures (e.g., flipping the palms from palm-down to palm-up to question, or shaking the head from side-to-side to negate); markers are used to modulate sentences and are not included here in our structural analyses of propositions (see Franklin et al., 2011, for an analysis of negative markers and question markers in homesign).

4.1.2. Coding types of nominal constituents

David used two types of gestures to refer to entities. (1) ***Demonstrative*** gestures: gestures that make reference by indicating a particular entity (e.g., point at a bird used to refer to that particular bird, *that*); and (2) ***Noun*** gestures: gestures that make reference by indicating the class of an entity, either by pointing at one object to refer to another (category pointing gestures, e.g., point at a bird used to refer to some other bird, thereby indicating the referent's class, *bird*), or by displaying characteristics of an object in an iconic noun gesture, e.g., flapping hands at the shoulders, which highlights an attribute of the referent's class, *bird*).

David also had a third way of referring to entities, called (3) ***Noun Phrases*** in our analyses. David would, at times, use both a demonstrative gesture and a noun gesture to refer to the same entity (e.g., point at bird combined with an iconic noun gesture for *bird*, palms flapping at sides). Note that the demonstrative gesture in this type of combination indicates the particular entity under discussion, whereas the noun gesture provides information about its class. Hunsicker and Goldin-Meadow (2012) found that these types of combinations function like complex nominal constituents in David's homesign system, thus warranting the label *Noun Phrase*.

We used the following criteria to identify Noun Phrases: (i) the two gestures in a Noun Phrase must refer to the same entity; (ii) the gestures must be within the same sentence; (iii) the gestures must be contiguous; (iv) the gestures must be of two different types (e.g., two pointing gestures at the same bird were not considered a Noun Phrase, even if they occurred in the same sentence and were adjacent); (v) the gestures must serve the same semantic role. This last criterion rules out predicate nominal sentences. For example, David sometimes points at a picture of a bird and then produces the noun gesture BIRD to identify the picture as a bird; in this case, the noun gesture is functioning as a predicate nominal (e.g., *that's a bird*), rather than as part of a nominal constituent (e.g., [*that bird*] *pedals a bike*). Predicate nominals were not coded as noun phrases.

4.1.3. Coding types of propositions

In addition to assigning meanings to nominal constituents, we also assigned propositional meanings to sentences. Once the boundaries of a gesture sentence were established using the motoric criteria described earlier, we used both the form of the gestures and the context in which the gestures were produced to assign meanings to propositions (Goldin-Meadow and Mylander 1984). David produced two types of propositions: ***Action*** and ***Stative***. Action propositions were coded when the child referred to an ongoing action (including pictures of ongoing actions) or an action that had just taken place or was about to take place (e.g., a request for an action); stative propositions were coded when the child described a static characteristic of an entity (see Goldin-Meadow and Mylander 1984 for details).

David used four types of Action propositions. Two of the four types were caused motion (i.e., transitive) events. (1) ***Transitive Crossing-Space***: an actor moves a patient across space to an endpoint or recipient (*I move jar to table*, a 3-place proposition, e.g., point at jar–MOVE, glossed as *that move;* or point at jar–point at table, glossed as *jar there*). Note that the child did not have to produce gestures for the all of the arguments in order for a sentence to be classified as conveying a 3-place proposition. In addition, sentences were classified according to type independent of the order of the gestures; that is, if the point at jar is produced after the MOVE gesture, it too is classified as transitive crossing space. (2) ***Transitive In-Place***: an actor acts on a patient in place (*I open jar*, a 2-place proposition, e.g., point at jar–OPEN, glossed as *that open;* or point at jar–point at self, glossed as *that me*). The remaining two types were spontaneous motion (i.e., intransitive) events. (3) ***Intransitive Crossing-Space***: an actor moves on its own across space to an endpoint or recipient (*I go to table*, a 2-place proposition, e.g., point at self–GO, glossed as *me go*). (4) ***Intransitive In-Place***: an actor moves on its own in place (*I dance*, a 1-place proposition, e.g., point at self–DANCE, glossed as *me dance*).

David used five types of stative propositions: ***Naming*** (e.g., point at bird–BIRD, glossed as *that [is] bird*), ***Describing*** (e.g., point at jar–BIG, *that [is] big*), ***Locative*** (e.g., point at jar–point at shelf, glossed as *that [belongs] there*), ***Possessive*** (e.g., point at jar–point at self, glossed as *that [belongs] me*), and ***Similarity between an object and a picture*** (e.g., point at picture–point at jar, *picture [resembles] jar*). David produced a sixth type of stative sentence, one in which he indicated the similarity between two objects (e.g., point at jar 1–point at jar 2). These combinations are excluded from our analyses here because it is impossible to tell which of the two points is functioning as the subject of the sentence (the O-type item), and which is functioning as the predicate (the C-type item). These stative combinations are described in detail in Ozcaliskan et al., 2009.

4.1.4. Coding reliability

Reliability was determined by having two independent coders transcribe portions of the videotapes (Goldin-Meadow and Mylander 1984). Agreement between coders was 91% for isolating gestures from the stream of motor behavior, 93% for determining boundaries between signs, 95% for determining boundaries between sentences, 93% for assigning meanings to pointing and iconic gestures, 94% for deciding whether an iconic gesture was a noun, verb, or adjective, and 94% for classifying sentences according to proposition type.

4.2. Results

Our goal was to determine whether David's homesigns can be described as a productive system. To do so, we examined 12 different types of combinations and determined how often each combination actually occurred in David's corpus, compared to how often that combination would be expected to occur using Yang's (2013) analytic technique.

In the first comparison, which focused on the nominal constituent, we asked whether David was equally likely to use a noun gesture to indicate a particular entity as he was to use a demonstrative gesture for the same entity. The first comparison is thus not an analysis of how often two forms combine, but rather an analysis of how often a C-type form (i.e., demonstrative vs. noun) is used to refer to an O-type entity (i.e., an open-ended set of entities). As an example, we determined how many times David referred to an apple using a *demonstrative* form (a pointing gesture at the apple), and compared that number to the number of times he referred to an apple using a *noun* form (an iconic noun gesture, APPLE). In a truly combinatorial system, references to the apple should occur in both the demonstrative form and the noun form. In a modest sample, however, some references will occur in both forms, whereas others will occur in only one of the two forms. In general, the probability that an entity will be referred to using one form is not necessarily equal to the probability that the entity will be referred to using the alternate form. As in spoken language, an apple is sometimes referred to demonstratively in homesign (e.g., a point at an apple, meaning "that [apple]") and sometimes referred to generically (e.g., a categorical pointing gesture or an iconic noun gesture for *apple*, meaning "apple"). To calculate *bias* (B) for a given sample, we empirically measure the number of times that the demonstrative form was used to refer to a particular entity, and the number of times that the noun form was used to refer to the same entity. We then take the larger of the two values (the dominant value) for each entity, sum the dominant values over all entities, and divide that sum by the total number of times both forms were used. The larger the value of B, the less likely that both forms will be used to refer to the same entity.

Our test for the Demonstrative vs. Noun comparison contrasts two diversity values: The empirical diversity value, and the expected diversity value. The empirical diversity value is calculated from the sample as the percentage of references to entities that appear in both forms, out of references to entities that appear in either form. In this instance, David referred to 392 (N) different entities, using either a demonstrative form or a noun form for a total of 2297 times (S). Of the 392 entities, 88 appeared in both the demonstrative and the noun form, for an empirical (i.e., observed) diversity measure of 22.4% (88 divided by 392). The statistical expectation of diversity, which is calculated from S (392), N (2297), and the bias value B (0.924), predicts a diversity value of 19.4%, the expected diversity measure under the assumption that an entity can be freely referred to using both forms.

In addition to the *Demonstrative* vs. *Noun* contrast, we conducted one other comparison within the nominal constituent: *Demonstrative or Noun vs. Noun Phrase*. We ask whether David was equally likely to use a gestural noun phrase to indicate a particular entity as he was to use either a demonstrative or noun gesture for that entity (that is, does the larger unit substitute for the smaller units?). For example, we compared the number of times David referred to an apple using either a *Demonstrative* or *Noun* with the number of times he referred to the apple using a *Noun Phrase* (i.e., using both a *Demonstrative* and a *Noun* within the same sentence).

In the remaining 10 comparisons, we asked how often a nominal constituent for a particular referent (an O-type form) was combined with other forms that came from a constrained category (C-type forms); in other words, we focused on how often C and O forms were combined within a sentence. For example, we asked how often David referred to the apple (using any of the three nominal forms) in an *Action* vs. *Stative* proposition. Within *Action* propositions, we asked how often he referred to the apple in *Transitive* vs. *Intransitive* events, in *Crossing-Space* vs. *In-Place* events, and in cross-cutting combinations of transitive, intransitive, crossing-space and in-space events (see Table 1). Within *Stative* propositions, we asked how often David referred to the apple (again using any of the three nominal forms) in a *Locative* vs. *Possessive* event, in a *Naming* vs. *Describing* event, and in a *Naming vs. Identifying by Picture* event.

The expected diversity for each comparison is computed with the values of S, N, and B in Table 1 according to [1]. Yang (2013) compares the same construction (the determiner-noun combination) across multiple language samples and therefore assumes a constant value of the bias. Because the present study compares observed and expected combinatorial diversity across different combinations, the bias factor B must be calculated empirically for each combination. For instance, "ice-cream", an inanimate nominal, is more often used intransitively, whereas "boy", an animate nominal, is more often used transitively. The bias factor B, then, is the average probability of the more favored predicate across all nominals. If "ice-cream" is used intransitively 5 times and 0 times transitively, and "boy" is used intransitively 2 times and 10 times transitively, then the bias for these nominals combined with transitive/intransitive predicates is $(5 + 10)/(5 + 0 + 10 + 2) = 0.88$. This value, empirically measured, is then used to calculate the expected diversity using [1].

We found no significant difference between the expected and empirical values for the twelve comparisons in Table 1. The concordance correlation coefficient test (Lin 1989), which is appropriate for testing identity between two sets of continuous variables, confirms this conclusion ($\rho_c = 0.975$; 95% confidence interval 0.926-0.992). Fig. 1 displays the two sets of values graphically and makes it clear that the expected and empirical values are nearly identical. This finding suggests that David's homesigns are consistent with the output of a system that freely combines gestures.

It is also informative to examine combinatorial diversity across different comparisons in David's homesigns. The more times an entity is referred to, the more likely both forms of C will be used to refer to it. And indeed Yang (2013) found that S/N (the number of different nouns in a sample divided by the total number of nouns in the sample, a measure of how often an entity is referenced) was correlated with usage diversity in determiner-noun combinations; the low value of S/N thus helps explain the relatively low usage diversity score for determiner-noun combinations in the (adult) Brown Corpus (Kucera and Francis 1967), which was lower than in some child language corpora. Another factor, the bias for the two alternative forms of C (B, which is the probability of the more likely form), will also affect usage diversity. The more heavily one form of C is favored over the other form (i.e., the bigger the difference between f_1 and f_2), the lower the diversity value will be. Keeping to the coin toss analogy, a heavier bias (toward either side) will reduce the diversity of heads and tails in a sample of coin tosses. The bias values B shown in Table 1 vary considerably over the range of constructions in our comparison tests simply because the constructions have very different semantic and pragmatic conditioning factors. Taken together, these observations lead to a composite predictor, S/(NB), for usage diversity. The results in Table 1 confirm our reasoning: The value S/(NB) is strongly correlated with the empirical value of diversity ($r = 0.76$, $p < 0.005$). Thus, the usage diversity in a linguistic corpus can be well accounted for by the statistical nature of linguistic combinations, supporting a grammar-based approach not only to early child language, but also to homesign.

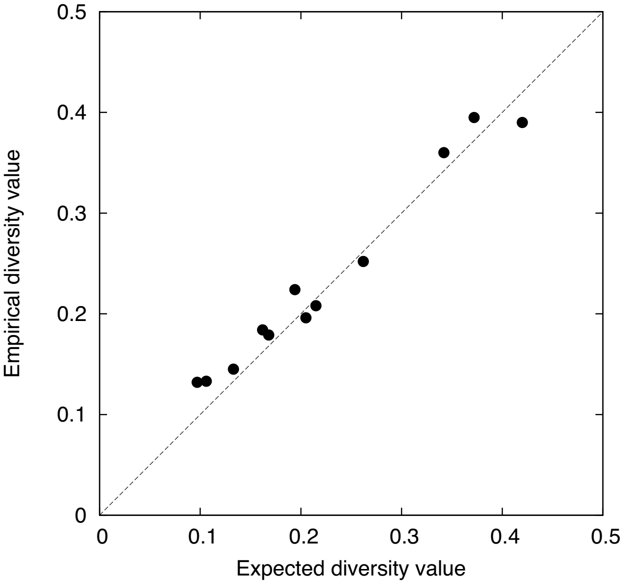

Fig. 1. The expected diversity value for each of the 12 constructions agrees well with the empirical diversity value for that construction (dotted line indicates identity).

5. Discussion

Our findings have three important implications for our understanding of language learning and how language might have evolved as a biological capacity.

First, the findings confirm, using a stringent statistical analysis, that a child who is lacking input from a conventional language model (in this case, a homesigner) can nevertheless communicate using a productive combinatorial system. Adjusting for the quantitative property of language known as Zipf's (1949) Law and for statistical bias in syntactic combinations, we found that the homesigner freely combined gestures referring to a particular entity in different syntactic constructions (e.g., action vs. stative; transitive vs. intransitive; crossing-space vs. in-place; locative vs. possessive), and freely used a demonstrative gesture, a noun gesture, or a noun phrase gesture-combination to refer to the same entity. In other words, the child generated gesture sentences characterized by a grammar that combines independent and interchangeable linguistic units, the hallmark of a human syntactic system. Importantly, although David could have developed a language that is not structure dependent, he did not––confirming that there are constraints

Table 1
Expected and Empirical Diversity Measures for Each of the 12 Comparisons Conducted on Homesign.

Type of Comparison	Gesture type (N)	Sample size (S)	Bias (B)	Expected diversity	Empirical diversity	S/(NB)
Nominal Constituent						
Demonstrative vs. Noun	392	2297	0.924	0.194	0.224	6.341
Demonstrative/Noun vs. Noun Phrase	327	2550	0.943	0.205	0.196	8.271
Action vs. Stative	313	2199	0.780	0.420	0.390	9.008
Action						
Transitive vs. Intransitive	238	1300	0.878	0.262	0.252	6.221
Crossing-Space vs. In-Place	238	1300	0.764	0.372	0.395	7.149
Transitive: Crossing-Space vs. In-Place	211	1090	0.791	0.342	0.360	6.531
Intransitive: Crossing-Space vs. In-Place	87	210	0.862	0.162	0.184	2.800
Crossing-Space: Transitive vs. Intransitive	159	726	0.897	0.215	0.208	5.090
In-place: Transitive vs. Intransitive	173	574	0.923	0.133	0.145	3.595
Stative						
Locative vs. Possessive	83	153	0.895	0.106	0.133	2.059
Naming vs. Describing	136	291	0.918	0.097	0.132	2.331
Naming vs. Identifying by picture	140	652	0.933	0.168	0.179	4.991

on the kinds of communication systems human children create (and thus fewer languages than are logically possible).

Second, because the homesigner had no model (either from a conventional sign language or from co-speech gesture) for his gesture system, the gesture combinations he produced could not have been memorized as unanalyzed chunks. We know, however, that memorization plays an important role in language learning––if children are exposed to a model for a language, they use that model to determine the *particular* combinatorial patterns that characterize the language they are learning. But do they need a language model to arrive at the idea of introducing combinatoriality into their communications in the first place? Our data suggest that they do not––homesigners do not have a model for a combinatorial gesture system but construct a combinatorial system nonetheless. Introducing combinatoriality into a communication system, along with other properties of language found in homesign––for example, hierarchical structure (Hunsicker and Goldin-Meadow, 2012), recursion (Goldin-Meadow, 1982, 2005), displacement (Butcher et al., 1991; Morford and Goldin-Meadow 1997), negation and questions (Franklin et al., 2011)––does not require a language model.

Third, juxtaposing homesigners with language-trained chimpanzees provides insight into the evolution of language. Homesigners are deprived of a linguistic model but nevertheless generate a productive communication system. Language-trained chimpanzees do not generate a productive system despite the fact that they do have a model for language; instead, they imitate their trainers even after years of exposure to a language model. The urge to communicate using a productive combinatorial system is so weak in chimpanzees that they do not even see productivity in the combinatorial communication systems to which they are exposed. In contrast, this urge is so strong in human children that they will create a system with combinatorial productivity even if not exposed to one. Taken together, these findings provide evidence that combinatorial productivity may have been a defining step in the evolution of language.

6. Author contribution

SGM collected the homesign data and conducted the behavioral analyses; Yang applied the statistical technique to the data; both authors wrote the paper.

Acknowledgements:

Supported by grants from NIDCD (R01 DC00491) and from NSF (BNS 8497941) to Goldin-Meadow. We thank Sarah Fulton, Dea Hunsicker, and Carolyn Mylander for their help in coding and organizing the homesign data, and Lila Gleitman, Marie Coppola and John Goldsmith for their comments on an earlier draft of the manuscript. We also thank two anonymous reviewers for their helpful criticisms and suggestions.

References

Baroni, M., 2009. Distributions in text. In: Ludeling, A., Kyoto, M. (Eds.), Corpus Linguistics: An International Handbook. Mouton de Gruyter, Berlin.
Berwick, R., Chomsky, N., 2016. Why Only Us: Language and Evolution. MIT Press.
Bickerton, D., 1995. Language and Human Behavior. University of Washington Press.
Butcher, C., Mylander, C., Goldin-Meadow, S., 1991. Displaced communication in a self-styled gesture system: pointing at the non-present. Cognit. Dev. 6, 315–342.
Choi, S., Bowerman, M., 1991. Learning to express motion events in English and Korean: the influence of language-specific lexicalization patterns. Cognition 43, 83–121.
Chomsky, N., 1965. Aspects of the Theory of Syntax. MIT Press.
Coppola, M., Newport, E., 2005. Grammatical Subjects in homesign: abstract linguistic structure in adult primary gesture systems without linguistic input. Proc. Natl. Acad. Sci. 102, 19249–19253.
Feldman, H., Goldin-Meadow, S., Gleitman, L., 1978. Beyond Herodotus: the creation of language by linguistically deprived deaf children. In: Lock, A. (Ed.), Action, Symbol, and Gesture: The Emergence of Language. Academic Press, New York, pp. 351–414.
Fisher, C., 2002. The role of abstract syntactic knowledge in language acquisition: a reply to Tomasello (2000). Cognition 82, 259–278.
Flaherty, M., Hunsicker, D., Goldin-Meadow, S., 2017. The birth of a new language: gestural precursors to Nicaraguan Sign Language. Under Rev.
Franklin, A., Giannakidou, A., Goldin-Meadow, S., 2011. Negation, questions, and structure building in a homesign system. Cognition 118, 398–416.
Goldin-Meadow, S., Feldman, H., 1977. The development of language-like communication without a language model. Science 197, 401–403.
Goldin-Meadow, S., Mylander, C., 1983. Gestural communication in deaf children: the non-effects of parental input on language development. Science 221, 372–374.
Goldin-Meadow, S., Mylander, C., 1984. Gestural communication in deaf children: the effects and non-effects of parental input on early language development. Monogr. Soc. Res. Child Dev. 49, 1–121.
Goldin-Meadow, S., Mylander, C., 1998. Spontaneous sign systems created by deaf children in two cultures. Nature 391, 279–281.
Goldin-Meadow, S., Butcher, C., Mylander, C., Dodge, M., 1994. Nouns and verbs in a self-styled gesture system: what's in a name? Cognit. Psychol. 27, 259–319.
Goldin-Meadow, S., Mylander, C., Franklin, A., 2007. How children make language out of gesture: morphological structure in gesture systems developed by American and Chinese deaf children. Cognit. Psychol. 55, 87–135.
Goldin-Meadow, S., Brentari, D., Coppola, M., Horton, L., Senghas, A., 2015. Watching language grow in the manual modality: nominals, predicates, and handshapes. Cognition 135, 381–395.
Goldin-Meadow, S., 1979. Structure in a manual communication system developed without a conventional language model: language without a helping hand. In: Whitaker, H., Whitaker, H.A. (Eds.), Studies in Neurolinguistics, Vol. 4. Academic Press, New York.
Goldin-Meadow, S., 1982. The resilience of recursion: a study of a communication system developed without a conventional language model. In: Wanner, E., Gleitman, L.R. (Eds.), Language Acquisition: The State of the Art. Cambridge University Press, New York, pp. 51–77.

Goldin-Meadow, S., 2003. The Resilience of Language: What Gesture Creation in Deaf Children can Tell Us About How All Children Learning Language. Psychology Press, New York.

Goldin-Meadow, S., 2005. What language creation in the manual modality tells us about the foundations of language. Ling. Rev. 22, 199–225.

Hauser, M.D., Chomsky, N., Fitch, W.T., 2002. The faculty of language: what is it, who has it, and how did it evolve. Science 298 (5598), 1569–1579.

Hauser, M.D., Yang, C., Berwick, R.C., Tattersall, I., Ryan, M.J., Watumull, J., Chomsky, N., Lewontin, R.C., 2014. The mystery of language evolution. Front. Psychol. 5, http://dx.doi.org/10.3389/fpsyg.2014.00401.

Hunsicker, D., Goldin-Meadow, S., 2012. Hierarchical structure in a self-created communication system: building nominal constituents in homesign. Language 88 (4), 732–763.

Hurford, J., 2012. The Origins of Grammar: Language in the Light of Evolution II, Vol. 2. Oxford University Press.

Kucera, H., Francis, N., 1967. Computational Analysis of Present-Day English. Brown University Press, Providence, RI.

Lidz, J., Gleitman, H., Gleitman, L., 2003. Understanding how input matters: verb learning and the footprint of universal grammar. Cognition 87, 151–178.

Lillo-Martin, D., 1999. Modality effects and modularity in language acquisition: the acquisition of American Sign Language. In: Ritchie, W.C., Bhatia, T.K. (Eds.), The Handbook of Child Language Acquisition. Academic Press, New York, pp. 531–567.

Lin, L.I.-K., 1989. A concordance correlation coefficient to evaluate reproducibility. Biometrics 45, 255–268.

Morford, J.P., Goldin-Meadow, S., 1997. From here to there and now to then: the development of displaced reference in homesign and English. Child Dev. 68, 420–435.

Newport, E.L., Meier, R.P., 1985. The acquisition of american sign language. In: Slobin, D.I. (Ed.), The Cross-linguistic Study of Language Acquisition, Vol. 1: The Data. Erlbaum, Hillsdale,N.J, pp. 881–938.

Ozcaliskan, S., Goldin-Meadow, S., Gentner, D., Mylander, C., 2009. Does language about similarity play a role in fostering similarity comparison in children? Cognition 112, 217–228.

Pine, J., Lieven, E., 1997. Slot and frame patterns in the development of the determiner category. Appl. Psycholing. 18 (2), 123–138.

Pine, J.M., Freudenthal, D., Krajewski, G., Gobet, F., 2013. Do young children have adult-like syntactic categories? Zipf's law and the case of the determiner. Cognition 127, 345–360.

Studdert-Kennedy, M., 1998. The particulate origins of language generativity: from syllable to gesture. In: Hurford, J., Studdert-Kennedy, M., Knight, C. (Eds.), Approaches to the Evolution of Language. Cambridge University Press, Cambridge, pp. 202–221.

Tattersall, I., 2012. In: Macmillan Terrace, H.S. (Ed.), Masters of the Planet: The Search for Our Human Origins, 1979. Knopf, Nim. N.Y.

Terrace, H.S., Petitto, L.A., Sanders, R.J., Bever, T.G., 1979. Can an ape create a sentence. Science 206 (4421), 891–902.

Tomasello, M., 2000. Do young children have adult syntactic competence? Cognition 74 (3), 209–253.

Tomasello, M., 2003. Constructing a Language. Harvard University Press, Cambridge, MA.

Valian, V., Solt, S., Stewart, J., 2009. Abstract categories or limited-scope formulae? The case of children's determiners. J. Child Lang. 36 (4), 743–778.

Yang, C., Valian, V., 2014. Determining the productivity of determiners. In: Paper Presented at the 40th Boston University Conference on Language Development, Boston, MA.

Yang, C., 2013. Ontogeny and phylogeny of language. PNAS 110 (16), 6324–6327.

Zipf, G.K., 1949. Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology. Addison-Wesley, Reading, MA.