

Original Paper

The Unifying Theory of Bootstrapping: The Collaboration of Multiple Language Acquisition Mechanisms

Zequn Harry Song^{1*}

¹ Zequn H. Song, Kent School, Kent, CT, USA

* Zequn Harry Song, Zequn H. Song, Kent School, Kent, CT, USA

Received: September 17, 2020 Accepted: October 6, 2020 Online Published: October 17, 2020

doi:10.22158/eltls.v2n4p16

URL: <http://dx.doi.org/10.22158/eltls.v2n4p16>

Abstract

Scientists have long been exploring the possibility of Universal Grammar (UG), a linguistic gene that engenders our language acquisition process. What UG may comprise is debated: if semantics is innate, or syntax, or some other linguistic aspects. Moreover, no definitive evidence has surfaced to attest to its existence. Therefore, a first-language acquisition process without the prerequisite of UG is called for. In this paper, we tear apart the incompatibilities among different language acquisition hypotheses and combine them into a theory in which language learning does not require UG. We contend that a unification of the current hypotheses (i.e., pragmatic, prosodic, semantic, and syntactic bootstrapping) is substantial for first-language acquisition, and that scientists should be wary of oversimplifying matters with UG.

Keywords

language acquisition, syntactic bootstrapping, prosody, universal grammar, semantics, pragmatic bootstrapping, english

1. Introduction

First language acquisition has been under keen investigation by linguists around the world. Indian philosophers and scientists pondered, for more than twelve centuries, if language is bestowed by God (innate) or learned (Matilal, 1992). The debate still lingers: if language acquisition is an innate or acquired ability or both, and the interrelation among linguistic input, children's neurological development, and their congenital capabilities. However, despite decades of research, the understanding of how the essential components interact and consummate to the acquisition of language remains fragmentary.

The central question is a chicken-or-egg dilemma—if syntax comes before semantics or vice versa, and if

there are any other components in play. Many theorists have developed hypotheses in an attempt to disentangle the problem, including prosodic bootstrapping (**I**), pragmatic bootstrapping (**II**), semantic bootstrapping (**III**), and syntactic bootstrapping (**IV**). Many theories are considered mutually antagonistic, feuding over their plausibility and credibility. The most notable one is the wrangle between semantic bootstrapping and syntactic bootstrapping (cf. **IV**).

The wrangle pivots on Universal Grammar (UG), a natural instinct definitive for linguistic development. In syntactic bootstrapping, it is posited that children have an innate sense of syntax and can link syntactic categories to semantic categories; in semantic bootstrapping, this sense is reversed—children can link semantic categories to syntactic categories without external instructional input. Whether UG exists and in what form determines how we acquire language. Despite its promise, unresponsive evidence surfaces and calls for a theory in which UG is not a pre-determining factor. In our theory, children acquire language using instincts unlimited to language, such as pattern recognition.

In this paper, we tear apart the theories' incompatibilities and review the problem from a holistic perspective: how the bootstrapping hypotheses work in tandem during the developmental period and ultimately lead to the mastery of a language. Imagine a piano. Four different notes are played separately and consecutively; one initial note starts off the melody, and each other note gradually adds to the consonance; at last, each pitch sounds and contributes to the overall harmony until the note is complete. In the Unifying Theory of Bootstrapping, there is a dynamic referential relation that links all the bootstrapping mechanisms: pragmatic bootstrapping initiates the learning process of basic vocabulary; the word repertoire built from pragmatic bootstrapping undergirds semantic bootstrapping, which, with the help of prosodic bootstrapping, engenders a basic comprehension of syntax; the understanding enables syntactic bootstrapping to breed the acquisition of more words—semantics. In the end, all the bootstrapping mechanisms function in a multidirectional fashion and result in language comprehension. Here we primarily focus on English.

Such an agreement cannot be reached without certain compromises of the original four hypotheses. Hence, their definitions are altered in this paper. Please refer to each bootstrapping hypothesis's respective section for their modified definition as well as reasoning and evidential research.

Lastly, in the Discussion section, we review the prospect and function of Universal Grammar (UG) in language acquisition and suggest further research be done on UG to test the credibility of the Unifying Theory of Bootstrapping.

2. Prosodic Bootstrapping

Prosodic bootstrapping hypothesizes that primary language learners use prosodic features from fluent speakers' elocution as cues to identify grammatical properties of the language. Plausible prosodies include amplitude, tempo, pitch, rhythm, and, etc. In this paper, prosodic bootstrapping is defined as a "sound bootstrapping" mechanism: the various qualities of sound confer not only a language's grammatical properties but also semantics.

Evidence has shown that prosody plays a large part in syntax learning in adults, and that those prosodic cues are exaggerated during a toddler's developmental period. Experiments conducted by Morgan et al. show that adult subjects were entirely successful in learning syntax only when the input included some cue marking the phrase structure of sentences (Morgan et al., 1987). Morgan proposes that prosodic cues delineate sentence domains within which distributional analysis is most easily and efficiently pursued. Such distributional analysis identifies each grammatical component of the sentence, and as linguistic experience accumulates, a mode is observed regarding the sequence of these grammatical components, precipitating grammar comprehension.

Relevant investigations demonstrate that prosodic cues are more exaggerated in Infant-Directed Speech (IDS) than in Adult-Directed Speech (ADS) (Fernald & Mazzie, 1991). The most observed IDS is motherese, where adults hyper articulate vowels at a high pitch and a slow speech rate. Motherese is essential for infants to amplify, discern, and utilize prosodic cues for language acquisition (Kuhl et al., 2007). Seven experiments in 1992 show that young infants are highly sensitive to syntax-prosody relations (Jusczyk et al., 1992). Collectively, these pieces of evidence, although indirect, speak to infant's application of prosody for syntax learning: If 1) an adult is utilizing prosodic cues to learn syntax, 2) prosody is emphasized in IDS, 3) infants are acute to syntax-prosody relations, they must take advantage of such speech features.

Researchers at Purdue University examined whether six-month-old infants could use prosodic cues for clause segmentation and if they weigh each prosodic cue (pause duration, pitch, and pre-boundary lengthening) differently (Seidl, 2007). Experimental outcome indicates that by six months of age, infants are already attuned to discerning prosodies for clause segmentation; and, they do assess such cues differently. This is a most direct evidence for prosodic bootstrapping. Hence, we concede the existence and utilization of prosodic cues in syntax acquisition.

Although the original theory of prosodic bootstrapping pertains to the comprehension of morph syntactic structures, theorists have propounded that a broader implication should be considered. Ramachandran and Hubbard suggest that, due to a biologically endowed capability to map and integrate multi-modal input, children as early as preverbal infants are highly sensitive to sound symbolism (Ramachandran & Hubbard, 2001). Sound symbolism enables children to connect speech sounds to their referents and thus, establish lexicon representations (Imai & Kita, 2014), an idea acknowledged and investigated by Edward Sapir (Sapir, 1929).

A convincing piece of evidence for the induction of word meanings from sound is an experiment done on infants reared in monolingual Spanish families (Peña et al., 2011). In Spanish, sound symbolism of vowels is associated with the size of an object. High/mid-frontal vowels (/i/, /e/) are associated with small objects; low/mid-posterior vowels (/o/, /a/) are with larger ones. Thirty-two trials were conducted on three-month-old babies from Spanish speaking families. They were presented with two geometric shapes that only differ in size while listening to a sound-symbolic syllable (e.g., "di" versus "da"). The babies demonstrated a sensibility for sound symbolism by looking at the sound-symbolically matching object

for a longer period of time (e.g., they had a longer attention span looking at the bigger object than at the smaller object when “da” is produced).

Similar experiments have been conducted on English-speaking infants. It was hypothesized that English vowels provide hints of an object’s shape: “kiki” sound-symbolically matches a spiky/angular shape; “bubu” matches a round/curvy shape (Ozturk et al., 2013). Four-month-old babies have been tested. They correctly identified the sound symbol’s matching object and therefore, demonstrated a sensitivity and utilization of sound symbolism.

However, many experiments have provided discrepant results, failing to signal the connection between sounds and shapes. A paper by Imai and Kita includes such an unsuccessful experiment and concluded that the effect of this type of sound symbolism in this age group should be deemed fragile (Imai & amp, n.d.; Kita, 2014). Employing the same method in the previously mentioned experiment, Fort et al. exposed prelexical infants to potential sound symbols (e.g., “bouba”, “kiki”) and observed their reactions. They found no significant differences in attention span for the two objects and therefore, no indication of them understanding sound symbolism (Fort et al., 2013).

Hence, in the Unifying Theory of Bootstrapping, we suggest that while prosodic (sound) bootstrapping does provide supplementary details for word learning via sound symbols, it mainly contributes to the comprehension of morph syntax.

3. Pragmatic Bootstrapping

Pragmatic bootstrapping refers to the process of word learning using both verbal and nonlinguistic cues. Gestures, eye movement, and focus of attention are some examples of the pragmatic attributes employed. For instance, a child is able to grasp the meaning, usage, and production of the word “sofa” from an instructor’s utterance of the word and their gesturing toward the sofa. Pragmatic bootstrapping is considered to be essential for language processing and acquisition in any context or modality (Oller, 2005).

The mechanism is quite intuitive. When an object is presented and an utterance is perceived, the brain automatically associates and registers the two, and a new word is acquired. The “sofa” instance above is representative of this process in which a gesture is the pragmatic cue in use. However, children are more sensitive to these cues than we would normally imagine. Below is an illustration where they take advantage of a micro-pragmatic signal.

Children are shown to be capable of following other’s gaze as a referential cue for word acquisition. In Baldwin’s research, the investigator showed two different objects to an infant and immediately concealed them into separate containers (Baldwin, 1993). The investigator peeked into one container and produced an utterance alien to the infant. The two objects were taken out and presented to the subject, who, when later requested, successfully picked out the target item the utterance denoted. Therefore, the participant followed the investigator’s gaze and associated the utterance with the object. Furthermore, Brooks and Meltzoff’s experiments revealed that by nine months of age, infants are highly sensitive to body

movement (in their case, head movement); by ten months of age, they are able to follow the experimenter's eye movements, which is correlated with their subsequent language scores at 18 months (subsequent language scores are used to quantify an infant's stage of the comprehension of a language) (Brooks & Meltzoff, 2005). The result suggests that pragmatic bootstrapping initiates at an early age.

There seems to be an instinct to map a referent to its articulation. Besides the early utilization of pragmatic bootstrapping, children are surprisingly good at recognizing the correct object despite distractions. In a typical scene, there are numerous possible referents, and therefore the paring of the target referent to a word is supposedly difficult. In the "sofa" example, when the speaker points toward the sofa, potential referents include the cushions, the backrest, the arms, the clock hanging near the sofa, the sofa itself, and etcetera. Infants and adults demonstrate no difficulty in associating the words with the intended referent. In Baldwin's research cited above (Baldwin, 1993), participants were able to ignore distractions from temporal contiguity and focus on the target objects. Whether the innate ability is exclusively linguistic warrants further inspection.

In the Unifying Theory of Bootstrapping, pragmatic bootstrapping is the main device for semantic learning in early language acquisition. We agree that children can instinctually recognize the target referent for an articulation, but do not avow to its exclusiveness to language; it might be a general learning and attention device ([Discussion](#)).

4. Semantic Bootstrapping

Semantic bootstrapping posits that to set off language learning, children acquire words which are first conceptualized as objects or actions (semantic categorization) and, based on their vocabulary, construct the syntax of the target language. In other words, children develop their semantic categories into syntactic categories (i.e., word classes), such as nouns and verbs, and then acquire syntax.

A well-remarked occurrence in children's early language usage is that nominals, a type of noun, are generally the first words learned, followed by an acquisition of both nouns and verbs with a vaster vocabulary base for nouns. This phenomenon is documented in a project sponsored by the National Institute of Education (Gentner, 1982). With the acquisition of verbs, children originate phrases and short sentences, demonstrating a preliminary understanding and usage of the language's grammatical properties. The observation is in concord with the Critical-Mass Hypothesis: developments within morph syntax are triggered by an increase of the lexicon a child possesses.

The strong dependence of grammatical development on children's lexicon is bolstered by Bate, Bretherton, and Snyder's research. They revealed a correlation between the infant's vocabulary size at 20 months of age and their state of morph syntactic development at the age of 28 months (Bates et al., 1988). They further proposed that the state of grammatical development at 28 months is best indicated by the vocabulary size eight months earlier. This correlation has been replicated cross-linguistically in studies on the acquisition of English (Fenson et al., 1994), Swedish (Berglund & Eriksson, 2000), German (Szagun et al., 2006), and also for bilinguals (Conboy & Thal, 2006). The dependence strongly suggests

that grammar is somehow acquired using observation of semantics and essentially, semantic bootstrapping.

According to Pinker, there are two preconditions for semantic bootstrapping to hold true (Pinker, 1984). First, a child must be able to understand word meanings and categorize them semantically as objects, or actions, or other sorts in reality. Second, they must be aware of the correspondence between semantic and syntactic categories and able to observe, integrate, and develop morph syntactic rules from their semantic knowledge. In the following passage, we posit a learning mechanism through which the two prerequisites can be met and set semantic bootstrapping in motion.

We propound that the aforementioned pragmatic bootstrapping (II) serves to obtain a basic vocabulary repertoire. Pinker's first rule implies that before the onset of semantic bootstrapping, a vocabulary repository needs to be secured, and children must be capable of using words to describe real-world objects and activities. In pragmatic bootstrapping, children retain word meanings by directly observing worldly events. Repeatedly looking at a "sofa" while hearing the pronunciation of the word automatically commits the term to memory along with the object it denotes. A child learns verbs—at this stage, words for actions—in a similar fashion. For the word "pat", repetitively observing someone gently caressing a dog with a petting motion and perceiving its diction helps the toddler associate the real-life event with the oral and/or auditory signal. Since pragmatic cues are bound to the physical world and ubiquitously accessible—they exist in daily social interactions where verbal languages are accompanied by body language, attention, eye focus and, etc.—kids are able to build strong connections between utterances and real-life objects since the very beginning of language acquisition and should have no problem using language to describe reality.

Nevertheless, words for actions and objects describe different referents, which may result in them being cognized differently. If so, it is a distinction a child can subconsciously and even consciously pick up to segregate vocabulary into semantic categories. An action word is generally more abstract than an object word. Most often, when relating to an object, the realistic referent is concrete and directly perceivable; when relating to an action, the referent is not physical and therefore elusive: an action word denotes an activity of objects, which exists not as an entity or designated to a specific object, but transiently as a perceived dynamic concept. For instance, "pat" as a verb can be used under various circumstances. The referred action prevails beyond the instantiated mother-to-dog relation. To accurately memorize and utilize the word, a child must realize that it does not denote the objects they perceive but the interaction between objects. This process is sensibly longer and more laborious than establishing a cognitive concept for a material object. The extra difficulty and elaborateness for action-word acquisition help distinguish them from object words, and hence, bestow the ability to construct different semantic categories. This theory explains the slower rate of verb acquisition relative to nouns (Gentner, 1982).

This strategy based on pragmatic bootstrapping and children's cognition suffices the first prerequisite for Pinker's theory.

Second, we propose a distributional model that institutes the correspondence between semantic and

syntactic categories and enables children to bootstrap syntax from their semantic knowledge.

Separating syntactic categories from semantic categories are their grammatical properties. Take verb inflections for example. Inflections are strapped onto the general terms/dictionary forms of verbs, a syntactic category, to indicate tenses, such as “-ed” for the simple past tense. According to semantic bootstrapping, a child extrapolates syntax from their vocabulary repertoire, and this paper suggests they do so by observing the commonalities among each word category. During the process of acquiring basic vocabulary, children are exposed to different forms of words—for nouns, singular and plural; for verbs, third-person singular, past tense, past participle, and present participle. In most cases, the overlap among different forms is evident enough for base-form recognition, and children would register them as “closely related” (e.g., *talk - talks, talked, talking*). As a child’s lexicon increases, they encounter more words with homogeneous inflections and subsequently recognize a pattern; with the assist of pragmatic and prosodic cues, learners can ascertain words and inflections’ grammatical meanings. Terms semantically categorized as objects generally have two forms, and the “-s” inflection is used to indicate plurality; action words predominantly enjoy more forms, and “-ed” is used for past tenses and participles while “-ing” is used for gerunds. Children, with the help of informative cues, can therefore extrapolate grammar from semantic categories, which are now developed into syntactic categories.

Admittedly, in English, there are plenty of special scenarios where children can acquire their usage even though the general rule does not apply. Irregular verbs, for instance, use uncommon inflections to indicate tenses, and some do not enjoy such design at all. These special cases can forsooth be learned through experience. If a child is immersed in reiteration and recurrence of an irregular verb’s inflected forms, their brain can memorize its unique pattern alongside the grammatical awareness they are accustomed to. Since to generalize a pattern requires more time and information than to memorize one specific set of inflections for an irregular verb, these irregularities should be learned earlier than the emergence of the “-ed” rule and continuously as children encounter novel words. This theory also entails a proclivity to apply the grammar rule to newly encountered verbs. Both phenomena are chronicled in research. Brown, by analyzing three children’s speech development over years, reported that some irregular verb inflections are learned prior to its regular counterpart, “-ed” (Brown, 1973), similar to the findings of other longitudinal research (cf. Cazden, 1968.) The overapplication of the regular inflection pattern is also observed in studies (Sherman, 1971; Slobin, 1973), including the two cited above—Young children often make errors such as *goed, singed, wented*, etcetera. These results show that a general rule is indeed developed, and irregular verb inflections are learned separately. Essentially, the learning strategy we propose is congruent with Pinker and Prince’s dual mechanism view that regular past tense forms are computed using the “-ed”-suffixation rule; irregular past tense forms are learned, stored in, and retrieved from associative memory (Pinker & Prince, 1988).

As explained, a connection between syntactic and semantic categories can be obtained, which satisfies the second prerequisite for semantic bootstrapping.

Therefore, children have already understood the semantics of some verbs and nouns and their

morphosyntactic qualities before semantic bootstrapping. When exposed daily to Subject-Verb-Object (SVO)-structured utterances with pragmatic and prosodic cues, they can decode the message easily and observe patterns. In the example “Mom pats the dog,” since the kid understands each component of the sentence (marked by prosodies) and perceives the image of “*Mom*”-“*patting*”-“*the dog*”, the subject-verb-object composition would be registered in their brain. As they encounter more sentences, a pattern can be generalized and constitutes syntax.

Moreover, an interesting statistical mechanism proffered by Misha Becker potentially explains how children learn to identify the subject and object in a sentence without pragmatics (Becker, 2016). She instantiates the subject/object identification in sentences with the main verb “hit”. In “John hits the chair,” one can easily grasp that “John” is the subject that initiates the “hit” action to the object “chair” since “chair” is an inanimate object that does not possess the ability to hit. In rare instances like “I was hit by a chair,” the “chair” is seen as a tool and object of preposition instead of the hitting subject. The technical subject is still “I,” an animate being, despite the passive-voice sentence composition. After hearing the usage of “hit” in several different sentences, a child can easily extrapolate that the object which the action is acted upon is usually localized after the verb, the object position. This observation is applied when they face an ambiguous case in which the subject and the object both have the ability to “hit.” “John hits Bill,” for example. A child would allege that “Bill” is the object/patient according to the mode they generalized through observations from linguistic experience. More research should be done to test the prospect of this hypothesis.

Either case, the semantics of the words are of tremendous help to analyze the sentences and generate an awareness of syntax.

Most critics against semantic bootstrapping question the innateness that connects semantic and syntactic categories. Ambridge, in his published perspective, opposes universal grammar and thence, semantic bootstrapping on the basis that semantic categories are not universal (i.e., different among languages) and ergo, not innate. (Ambridge et al., 2014) They argue that children do not inherit a mechanism to classify words, which is a prerequisite for the original semantic bootstrapping hypothesis. Here, we are not avowing to the nature of this mechanism, but simply suggesting a way in which in the absence of a preprogrammed instinct, such classification can be accomplished and in which semantics, with the help of pragmatics and prosody, can precipitate grammar comprehension.

Another opponent is Gleitman’s syntactic bootstrapping, which is explored below.

5. Syntactic Bootstrapping

Since its inception by Gleitman (Gleitman, 1990), syntactic bootstrapping has been deemed a rival hypothesis of semantic bootstrapping. It claims that children infer word meanings via observed syntax using an innate knowledge connecting syntactic and semantic categories, in contrast to semantic bootstrapping. In this paper, syntactic bootstrapping is redacted to simply represent the concept that children can use their observations of syntax to deduce verb meanings and further enrich their vocabulary.

It is here purported that the link between syntactic and semantic categories is not an innate ability but learned through other bootstrapping mechanisms during previous linguistic experience.

In years, research regarding the titular topic has focused heavily on verb acquisition. Hence, its ostensible opposition against semantic bootstrapping—the acquisition of syntax—does not necessarily exist, and the two mechanisms could co function reciprocally, which is acknowledged by Pinker in his response to Gleitman’s syntactic bootstrapping (Pinker, 1994). While research has shown that semantic meanings lead to a comprehension of morph syntax (the aforementioned Semantic Bootstrapping), the learned grammar, in return, is demonstrated to be informative for verb meanings.

Fisher et al., gave a structure-mapping account of syntactic bootstrapping (Fisher et al., 2010). When children encounter a sentence in which an unknown verb is juxtaposed with two known nouns (e.g., *She gorps her*), they can use syntax to deduct its meaning. First, they “map” the relations in the sentence—“*she*” and “*her*” are objects, and the unknown, “*gorp*”, is ostensibly a verb connecting the two. Children then assume that “*gorp*” is a transitive verb describing an action where the subject is the agent and the object the patient. Therefore, they eliminate some conceptual representations in which the action does not require a patient (e.g., *She sleeps.*) and retain the possible ones (e.g., *she feeds her; she slaps her; she loves her...*) Hence, the range of possible meanings is pinned down, enabling children to grasp the verb’s semantics.

However, it remains dubious if these clues are detailed enough for the specific definition of verbs. Using the syntax mentioned, it is impossible to further eliminate futile representations. The available syntactic clues are simply insufficient for toddlers to figure out if “feed”, “slap”, or “love” is the correct meaning for “gorp.” Hence, syntactic bootstrapping does not confer the ability to “zoom in” and grasp the accurate semantics, as Pinker stated (Pinker, 1994). Therefore, we concede that syntactic bootstrapping only helps to narrow down the meanings, enabling children to learn verbs with ease.

The “narrowing down” process, Fisher alleged, does not require prior syntax learning and only depends on children’s ability to construct a mental picture of the event narrated (Fisher et al., 2010). They argue that children have an innate ability to map nouns one to one onto an event’s participant-roles, and therefore, the SVO structure seems to be instinctual to them. We here argue the opposite: sufficient language exposure consists of other bootstrapping strategies is fundamental for syntax comprehension. Fisher quotes research done by Gertner et al., in an attempt to demonstrate that toddlers have a natural preference for word orders (Gertner et al., 2006), which is deemed a primitive language syntax that helps infer verb meanings. In four experiments, after hearing an SVO-structured sentence using a made-up verb, young children (25- and 21-month-olds) were shown two videos in which the roles of the agent and the patient differ. Even though they did not understand the verb, the participants exhibited a longer attention span to a scene in which the character named in the subject position was the agent, or the character named in the object position was the patient. The time course of the word-order effect supposedly suggests that children found it easy to use word order in this way and employ their natural preference to intuit the verb’s transitive quality and meaning. However, the results could be explained by

a misleading presupposition that underpins the research.

All the sentences were structured in a strict SVO composition: the subject comes first, then the verb, and lastly the object. The toddlers' reaction to the sentence (watching at the scene where the subject is the agent or where the object is the patient) might be a natural response looking at the first noun or pronoun they perceive. To rule out this possibility, researchers would need to employ non-SVO-structured sentences. Nevertheless, due to English's strict syntactic nature, such sentences may not make sense (e.g., "The apple ate Adam"). Even to adults, the noun that comes first would always be assumed to be the subject. Hence, experiments should be done in languages where high flexibility of word sequence is available, such as Greek or Russian. Since these languages usually employ case markers to signify their grammatical roles, research should be conducted on preverbal infants who have not been fully attuned to such signifiers. If the participant of the experiment could correctly identify the subject as the agent in an OVS- or VOS-structured sentence, that would connote an innate sensitivity that maps nouns one-to-one onto participant roles; otherwise, it would suggest that children learn word orders and other syntactic structures through linguistic experience, which we propose to be mainly semantic bootstrapping (III).

In 2004, ninety-six monolingual English children participated in an experiment to see if an "abnormal" word order can be instilled through language exposure (Matthews et al., 2004). The 48 children in the younger age group were of a mean age of 2:9 (range 2.3-3.2); the 48 older children were of a mean age of 3:9 (range 3.3-4.3). After the participants were acquainted with the target SOV structure, videos were played, and they were asked to describe the simple events recorded. Results have shown that the two-year-olds were significantly more likely to use the subject-object-verb composition that was taught before the experiment than the three-year-olds. The phenomenon evinces that three-year-olds are more accustomed than two-year-olds to the SVO structure in English, implying 1) new word order can be learned; 2) the comprehension of word order in English is a gradual process during language acquisition since the three-year-olds were considerably uninfluenced by their immersion in SOV-structured sentences than the two-year-olds.

Hakuta's research on Japanese-speaking children (Hakuta, 1982) shines a light on if they really have an inborn sense to map the nouns one-to-one onto participant-roles. Japanese speakers predominantly employ the subject-object-verb order. However, owing to postposed particles that signal grammatical roles, the language has high flexibility in word order. The children under investigation demonstrated their lack of ability to map nouns to agent-patient relations in any sentence without a clear marker of their cases even in the normal SOV structure. For them to interpret a noun as an agent, signifier "-ga" must be present; otherwise, they cannot take advantage of the sentence's canonical sequence. Hence, the subjects exhibited their reliance on external case markers and that their "natural instincts" to correctly interpret sentences may not exist. This behavior endorses our claim that the ability to identify subjects and objects in sentences as agents and patients is a result of linguistic experience and more accurately, semantic, prosodic, and pragmatic bootstrapping. We argue that syntax does not necessarily stem from a universal grammar designated for morphosyntax, but from a more general pattern recognition of signifiers and

word orders. How children learn syntax is detailed in Semantic Bootstrapping (III.)

In this section, we have established that 1) syntax does help learners grasp word meanings; 2) syntax itself is not sufficient enough to exact the meaning of a verb but only its grammatical qualities (e.g. transitive/intransitive); 3) the alleged innate ability through which children map nouns to agent-and-patient relations may not exist and is gradually learned through linguistic exposure. Further supported by the fact that two-year-olds are documented to have failed many syntactic-bootstrapping experiments (they could not correctly identify agent-patient relations) and that language acquisition initiates before two, we cling to the aforesaid Critical-Mass Hypothesis and stand by the view that syntactic bootstrapping happens subsequent to semantic bootstrapping. For learners who have obtained some syntactic and semantic knowledge, syntactic clues indeed help make out verb frames, contributing to language acquisition.

5.1 The Unifying Theory of Bootstrapping

Since there is much criticism against a genetic language faculty and no definitive evidence has been discovered, researchers need to be wary of to what extent instincts affect language learning (See Discussion). Therefore, a realistic account of language acquisition should rely on universal grammar as little as possible. In the Unifying Theory of Bootstrapping, we confer with the essence of the four main mechanisms. Furthermore, we establish a temporal connection that helps to attain the prerequisites of each bootstrapping theory through nurture. In this language acquisition strategy, prosodic, pragmatic, semantic, and syntactic bootstrapping are compatible and function multi-directionally, constituting a holistic linguistic experience.

Pragmatic bootstrapping commences language acquisition. Proficient speakers use body language, eye movement, and focus alongside verbality to inform the toddler of the connection between utterances and realistic referents. During this process, action words, object words, and other semantic categories are conceptualized differently based on their difficulty to be mentally represented. Different forms/inflections of a word are perceived and registered as closely related to its common form. As the vocabulary base builds up, patterns are detected within each semantic category, which points to grammar. Action words are predominantly verbs that have third-person singular, past tense, participle, and gerund forms; object words are majoritively nouns that demonstrate plurality via the “-s” inflection. The comprehension of these rules, a primary morphosyntax, help the learner develop semantic categories into syntactic categories.

Auxiliarily, sound symbols help the child with word meanings. Sound symbolism is detailed in the Prosodic Bootstrapping section (I). Besides examining similarities among word categories for grammar, the child also searches for commonalities between words under the same category, which accounts for their understanding of sound symbolism. They detect common features among terms whose referents have a same characteristic. The observed common feature becomes a sound symbol that is associated with that specific feature. “Kiki”, for example. sounds sharper and perkier than “Bubu”. Noticing sound symbols in an utterance helps narrow down its potential meaning.

At the same time, prosodies segmentize a sentence into syntactic components. These syntactic components include the subject, verb, and object positions, clause, sentential intention (statement/question/exclamation), and, etc. Children encounter similarly constructed sentences, commit the structure to memory, and therefore, acquire syntax.

Besides hearing the sentential composition, children also benefit from pragmatic and semantic bootstrapping for syntax acquisition. In the “Mom pats the dog” example, the observer understands each word meaning from their word depository and their respective syntactic category. When they observe the motion of “Mom-patting-the dog” while the sentence is articulated, they sense the “noun-verb-noun” sequence. With more experience and understanding of morph syntax, the kid develops the “noun-verb-noun” sequence into the SVO structure in English. A preliminary syntax is thus acquired.

When semantics exists absent of pragmatics, the statistical model proffered by Misha Becker may be used for semantic bootstrapping. As a child encounters more SVO-structured sentences using the same main verb, they analyze and compare the characteristics of the subjects and the objects. The observation bestows on children the awareness of the agent-patient relationship apropos of the subject-object relationship, which constitutes basic syntax. For detailed description, visit *Semantic Bootstrapping (III)*. With the introduction of syntax, syntactic bootstrapping kicks in. When a child faces an unknown verb in a sentence, whether an object is present could help determine the verb’s transitivity and eliminate futile possible representations. With its range of potential meanings narrowed down, the child can grasp the verb meaning more quickly and effectively. Semantics precipitates syntax acquisition, which, here as a reciprocation, helps word acquisition.

Meanwhile, all the other bootstrapping mechanisms are still at work. Pragmatic bootstrapping continues introducing new words; prosodies advance insights into sentential details, components, and syntax; semantics now not only helps with grammar comprehension, but also word acquisition. When an unknown noun is presented in a sentence with a familiar noun and a familiar verb, using preexisted semantic knowledge, a child efficiently pins down the potential meaning of the unknown. The meaning of the verb tells them the relationship between the two objects. With one object already known—either the agent or the patient—the child can then make an educated guess about the possible role the target noun plays and infer its meaning. Verbs can be learned similarly. Children estimate the relationship or the action between the agent and the patient, which is what the verb denotes.

All the mechanisms subsequently work in tandem and lead to the mastery of English.

6. Discussion

The paper probes into toddlers’ primary language acquisition process, especially in English. Since this process’ early inception, there could sensibly be a facilitative “language gene”. Newborns are able to discriminate phonetic contrasts of all languages; by 6 months of age, language-specific perception for vowels occurs; and they produce their first words at around 12 months (Kuhl et al., 2007). Universal Grammar (UG) strives to explain the extraordinary flair for language by suggesting that humans are

pre-programmed for its acquisition. Allegedly, certain languages' structural rules are encoded in our brain, and their existence is independent of postnatal linguistic experience. Linguistic stimuli build detailed grammatical rules upon these instincts. However, a prevailing number of scientists disavow universal grammar.

Chomsky, a postulator of UG, asserts his belief on saying human languages are “essentially identical” (Chomsky & Huybregts, 2004). Since they end up similar in structure, there has to be an instinct dictating language's grammatical development. However, others argue the differences among languages are too drastic to be deemed similar. Evans and Levinson vehemently tried to disprove Chomsky's view with counterexamples to the proposed universals (Evans & Levinson, 2009), including phrase structure rules, verb inflections, grammar that appoints subjects and objects to agents and patients, auxiliaries, anaphora, major lexical categories, and etc. Many typologists share the same concern (cf. Croft, 2009).

Current hypotheses of language acquisition more or less rely on UG. If UG does exist, the learning process of any language has to develop upon its foundation. To put it more clearly, UG conceptualizes a device built within the infant's brain and is definitive to access language. It determines the knowledge we naturally inherit even before the inception of language acquisition.

However, linguists suggest many instincts unrestricted to languages can potentially substitute. Michael Tomasello demurred against UG that language acquisition is forsooth dominated by processes constituted by these non-linguistic human instincts—cognition, perception, communication, etc. (Tomasello, 2010) Hence, without abundant evidence vouching for UG, the focus of the language acquisition process should be shifted from pre-existing knowledge to the potential source of these knowledge. Syntactic bootstrapping, for example, is originally based on the supposition that infants have an innate sense of syntax and can naturally link syntactic and semantic categories. However, the origin where a toddler obtains syntax is debated. A plethora of experiments and theories suggest its postnatal acquisition via linguistic experience, contrary to Gleitman's belief. Please refer to the Syntactic Bootstrapping section for details (IV.)

We agree that the dubiousness of UG discussed above calls for a language acquisition strategy that is not founded on linguistic instincts. The Unifying Theory of Bootstrapping relies on infants' more general, natural abilities unrestricted to language. Toddler's sensitivity to sound symbolism, for instance, is due to a biologically endowed **capability to generalize patterns** and associate them with an auditory input. This ability is not confined to language. Humans develop patterns everywhere in perception and cognition. For instance, in math, a general pattern and order dominates the progression of a sequence. In this paper, **this generalization ability is predominant**: in semantic bootstrapping, by discerning the commonalities in sentences and within word categories, children succeed at grasping the SVO structure and the grammatical properties of semantic categories. The development of semantic categories into syntactic categories and a preliminary comprehension of syntax are owed to pattern recognition.

It is natural for infants to recognize the different difficulty in cognition of semantic categories. As discussed earlier, action words may be harder to be mentally represented and committed than object

words, which underpins semantic segmentation. This instinct is unlikely exclusively linguistic. In terms of general cognition, humans memorize things at a different rate based on their difficulty to be mentally retained. The instinct should thus be deemed a cognitive ability.

We also acknowledge that infants can instinctually ignore distractions and map referents to utterances, but waver about its originality. In Baldwin's research (Baldwin, 1993), toddlers were able to ignore distractions from temporal contiguity and focus on the target objects. However, this phenomenon takes place under non-linguistic circumstances as well. When someone points at an object without any verbal communication, an observer can still identify the referent. Subjectively, they assume the biggest object they perceive at the direction of the pointer's attention is the target. The assumption is more or less accurate. Hence, this ability is more likely a derivative of the instinct to distribute attention proportionately to objects' sizes.

However, while further evidence for UG is needed, there is no absolute disrobement either. It is extremely difficult to falsify UG because of its recursive nature. The grammatical rules that are posited to be UG are direct observations of existing languages. It is defining itself in terms of itself. The recursion grants it infallibility. Geoffrey Sampson even called it "pseudoscientific" in his book (Sampson, 2005). Nonetheless, the possibility persists.

UG is crucial when it comes to different bootstrapping strategies. The original theory of semantic bootstrapping hinges on an innate knowledge of the link between semantic and syntactic categories; syntactic bootstrapping is based on the same instinct, but reversed; In the Unifying Theory of Bootstrapping, UG is neither a prerequisite nor a main factor for language acquisition. The core determinant in the mechanism debate is the prospect of UG.

Hence, further research should be conducted in pursuit of the existence of UG and the language-specificity of the aforementioned instinctual abilities.

Acknowledgment

Sincerest gratitude to Professor. Lindsay Whaley for his instruction and support and to Jessica McDermott for valuable suggestions.

References

- Ambridge, B., Julian, M. P., & Elena, V. M. L. (2014). Child language acquisition: Why universal grammar doesn't help. *Language*, 90(3). <https://doi.org/10.1353/lan.2014.0051>
- Baldwin, D. A. (1993). Early referential understanding: Infants' ability to recognize referential acts for what they are. *Developmental Psychology*, 29(5), 832-843. <https://doi.org/10.1037/0012-1649.29.5.832>
- Bates, E., Inge, B., & Lynn, S. S. (1988). *From first words to grammar: Individual differences and dissociable mechanisms*. Cambridge: Cambridge University Press.
- Becker, M. K. (2016). Essay. In *The acquisition of syntactic structure: Animacy and thematic alignment*

- (pp. 186-187). Cambridge: Cambridge University Press.
- Berglund, E., & Mårten, E. (2000). Communicative Development in Swedish Children 16-28 months old: The Swedish Early Communicative Development Inventory—Words and Sentences. *Scandinavian Journal of Psychology, 41*(2), 133-144. <https://doi.org/10.1111/1467-9450.00181>
- Brooks, R., & Andrew, N. M. (2005). The development of gaze following and its relation to language. *Developmental Science, 8*(6), 535-543. <https://doi.org/10.1111/j.1467-7687.2005.00445.x>
- Brown, R. (1973). *A First Language*. Harvard University Press. <https://doi.org/10.4159/harvard.9780674732469>
- Cazden, C. B. (1968). The Acquisition of Noun and Verb Inflections. *Child Development, 39*(2), 433. <https://doi.org/10.2307/1126956>
- Chomsky, N., & Huybregts, M. A. C. (2004). *The generative enterprise revisited: discussions with Riny Huybregts, Henk van Riemsdijk, Naoki Fukui and Mihoko Zushi*. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110902440>
- Conboy, B. T., & Donna, J. T. (2006). Ties Between the Lexicon and Grammar: Cross-Sectional and Longitudinal Studies of Bilingual Toddlers. *Child Development, 77*(3), 712-735. <https://doi.org/10.1111/j.1467-8624.2006.00899.x>
- Croft, W. (2009). *Radical construction grammar syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Evans, N., & Stephen, C. L. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences, 32*(5), 429-448. <https://doi.org/10.1017/S0140525X0999094X>
- Fenson, L., ..., Joan, S. (1994). Variability in Early Communicative Development. *Monographs of the Society for Research in Child Development, 59*(5), i. <https://doi.org/10.2307/1166093>
- Fernald, A., & Claudia, M. (1991). Prosody and focus in speech to infants and adults. *Developmental Psychology, 27*(2), 209-221. <https://doi.org/10.1037/0012-1649.27.2.209>
- Fisher, C., Yael, G., Rose, M. S., & Sylvia, Y. (2010). Syntactic bootstrapping. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(2), 143-149. <https://doi.org/10.1002/wcs.17>
- Fort, M., Alexa, W., Alexander, M., & Sharon, P. (2013). Looking for the bouba-kiki effect in prelexical infants. Ed.ac.uk. *INRIA, 71-76*. Retrieved from [https://www.research.ed.ac.uk/portal/en/publications/looking-for-the-boubakiki-effect-in-prelexical-infants\(4b388e5b-c20d-4229-8734-fe91b7fe330\).html](https://www.research.ed.ac.uk/portal/en/publications/looking-for-the-boubakiki-effect-in-prelexical-infants(4b388e5b-c20d-4229-8734-fe91b7fe330).html)
- Gentner, D. (1982). Tech. *Why nouns are learned before verbs: Linguistic relativity versus natural partitioning*. Retrieved from <https://www.scholars.northwestern.edu/en/publications/why-nouns-are-learned-before-verbs-linguistic-relativity-versus-n>
- Gertner, Y., Cynthia, F., & Julie, E. (2006). Learning Words and Rules. *Psychological Science, 17*(8), 684-691. <https://doi.org/10.1111/j.1467-9280.2006.01767.x>

- Gleitman, L. (1990). The Structural Sources of Verb Meanings. *Language Acquisition*, 1(1), 3-55.
https://doi.org/10.1207/s15327817la0101_2
- Hakuta, K. (1982). Interaction between particles and word order in the comprehension and production of simple sentences in Japanese children. *Developmental Psychology*, 18(1), 62-76.
<https://doi.org/10.1037/0012-1649.18.1.62>
- Imai, M., & Sotaro, K. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1651). <https://doi.org/10.1098/rstb.2013.0298>
- Jusczyk, P. W., Kathy, Hirsh-Pasek, Deborah, G. K. N., Lori, J. K., Amanda, W., & Julie, P. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology*, 24(2), 252-293. [https://doi.org/10.1016/0010-0285\(92\)90009-Q](https://doi.org/10.1016/0010-0285(92)90009-Q)
- Kuhl, P. K., Barbara, T. C., Sharon, Coffey-Corina, Denise, P., Maritza, Rivera-Gaxiola, & Tobey, N. (2007). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1493), 979-1000. <https://doi.org/10.1098/rstb.2007.2154>
- Matilal, B. K. (1992). *The word and the world: India's contribution to the study of language*. Delhi: Oxford University Press.
- Matthews, D., Elena, L., Anna, T., & Michael, T. (2004). The role of frequency in the acquisition of English word order. *Cognitive Development*, 20(1), 121-136.
<https://doi.org/10.1016/j.cogdev.2004.08.001>
- Morgan, J. L., Richard, P. M., & Elissa, L. N. (1987). Structural packaging in the input to language learning: Contributions of prosodic and morphological marking of phrases to the acquisition of language. *Cognitive Psychology*, 19(4), 498-550. [https://doi.org/10.1016/0010-0285\(87\)90017-X](https://doi.org/10.1016/0010-0285(87)90017-X)
- Oller, J. W. (2005). Common Ground between Form and Content: The Pragmatic Solution to the Bootstrapping Problem. *The Modern Language Journal*, 89(1), 92-114.
<https://doi.org/10.1111/j.0026-7902.2005.00267.x>
- Ozturk, O., Madelaine, K., & Athena, V. (2013). Sound symbolism in infancy: Evidence for sound–shape cross-modal correspondences in 4-month-olds. *Journal of Experimental Child Psychology*, 114(2), 173-186. <https://doi.org/10.1016/j.jecp.2012.05.004>
- Peña, M., Jacques, M., & Marina, N. (2011). The Role of Audiovisual Processing in Early Conceptual Development. *Psychological Science*, 22(11), 1419-1421.
<https://doi.org/10.1177/0956797611421791>
- Pinker, S. (1984). *Language learn ability and language development*. Cambridge: Harvard University Press.
- Pinker, S. (1994). How could a child use verb syntax to learn verb semantics? *Lingua*, 92, 377-410.
[https://doi.org/10.1016/0024-3841\(94\)90347-6](https://doi.org/10.1016/0024-3841(94)90347-6)
- Pinker, S., & Alan, P. (1988). On language and connectionism: Analysis of a parallel distributed

- processing model of language acquisition. *Cognition*, 28(1-2), 73-193.
[https://doi.org/10.1016/0010-0277\(88\)90032-7](https://doi.org/10.1016/0010-0277(88)90032-7)
- Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia—A window into perception, thought and language. *Journal of Consciousness Studies*, 8(12), 3-34.
- Sampson, G. (2005). *The “language instinct” debate*. New York: Continuum.
- Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology*, 12(3), 225-239.
<https://doi.org/10.1037/h0070931>
- Seidl, A. (2007). Infants’ use and weighting of prosodic cues in clause segmentation. *Journal of Memory and Language*, 57(1), 24-48. <https://doi.org/10.1016/j.jml.2006.10.004>
- Sherman, J. A. (1971). Imitation and Language Development. *Advances in Child Development and Behavior Advances in Child Development and Behavior Volume, 6*, 239-272.
[https://doi.org/10.1016/S0065-2407\(08\)60084-1](https://doi.org/10.1016/S0065-2407(08)60084-1)
- Slobin, D. I. (1973). *The ontogenesis of grammar: A theoretical symposium*. New York: Academic Press.
- Szagan, G., Claudia, S., Melanie, F., & Barbara, S. (2006). Development of vocabulary and grammar in young German-speaking children assessed with a German language development inventory. *First Language*, 26(3), 259-280. <https://doi.org/10.1177/0142723706056475>
- Tomasello, M. (2010). *Origins of human communication*. Cambridge, MA: MIT Press.
- Weissenborn, J., & Höhle, B. (2001). *Approaches to bootstrapping: phonological, lexical, syntactic and neurophysiologic aspects of early language acquisition*. Amsterdam: Benjamins.
<https://doi.org/10.1075/lald.24>
- Yuan, S., Cynthia, F., & Jesse, S. (2012). Counting the Nouns: Simple Structural Cues to Verb Meaning. *Child Development*, 83(4), 1382-1399. <https://doi.org/10.1111/j.1467-8624.2012.01783.x>