John Benjamins Publishing Company

This is a contribution from *Interaction Studies 12:1*
© 2011. John Benjamins Publishing Company

This electronic file may not be altered in any way.
The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.
Permission is granted by the publishers to post this file on a closed server which is accessible to members (students and staff) only of the author's/s' institute, it is not permitted to post this PDF on the open internet.
For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: www.copyright.com).
Please contact rights@benjamins.nl or consult our website: www.benjamins.com

Tables of Contents, abstracts and guidelines are available at www.benjamins.com

Simulating the coevolution of compositionality and word order regularity

Tao Gong
Department of Linguistics, University of Hong Kong

This paper proposes a coevolutionary scenario on the origins of compositionality and word order regularity in human language, and illustrates it using a multi-agent, behavioral model. The model traces a 'bottom-up' process of syntactic development; artificial agents, by iterating local orders among lexical items, gradually build up basic constituent word order(s) in sentences. These results show that structural features of language (e.g. syntactic categories and word orders) could have coevolved with lexical items, as a consequence of general learning mechanisms (e.g. pattern extraction and sequential learning) initially not language-specific.

Keywords: Computer simulation; language origin; coevolution; compositionality; word order regularity

1. The coevolutionary scenario of language origin

We humans, as the only species possessing language, have been wondering when, where, and how human language (henceforth simply "language") *emerges*, *changes*, and *dies out*. Among these stages, *language emergence*, as an intriguing topic in *evolutionary linguistics* (MacWhinney, 1999; Ke & Holland, 2006), has two distinct senses (Wang, 1991): ontogenetically, it refers to the process whereby an infant acquires a language from its environment (*language acquisition*, Clark, 2003); phylogenetically, it refers to the process whereby *Homo sapiens* made the gradual transition from a pre-linguistic communication system, perhaps similar to those of our ape contemporaries, to a communication system with the languages of the sort we use today (*language origin*, MacWhinney, 1999).

In evolutionary linguistics, there are two mainstream hypotheses on language origin. *Linguistic nativism* states that humans have a set of genetically-determined capacities by which to master and use a natural language (*the language faculty*, Chomsky, 1972; Jackendoff, 2002; Pinker & Bloom, 1990), though the contents

Interaction Studies 12:1 (2011), **63–106**. DOI 10.1075/is.12.1.03gon
ISSN 1572–0373 / E-ISSN 1572–0381 © John Benjamins Publishing Company

of this faculty and whether it is language or human specific remain uncertain (Hauser et al., 2002; Jackendoff & Pinker, 2005; Larson et al., 2010; van der Hulst, 2010). Language learning is a process whereby learners conceive, assess, and reject hypotheses on the properties of the exposed language (Mellow, 2008b). Some nativists (e.g. Chomsky, 1972) suggest that language emerged abruptly due to remarkable, massive mutations in early hominins' gene pool, others (e.g. Pinker & Bloom, 1990) believe that language resulted from a series of gradual steps caused by mutation or recombination. *Linguistic emergentism* (MacWhinney, 1999), however, argues that the existence of *linguistic universals* (the particular features or principles of language structure and use that hold across most but not necessarily all languages of the world, Christiansen & Kirby, 2003) does not necessarily imply a prefiguration in the brain (Croft, 2001). Language is an *interface* among a variety of basic abilities (Wang, 1982), some of which may underlie other aspects of human cognition and social life, present in other species in different degrees, or emerge earlier than language in hominin evolution. Based on these fundamental abilities across modules, early hominins gradually developed in their primitive communication system some universal features resembling those in modern languages. Acquisition is not fundamentally different from other types of learning, and can be accounted for by the rich experiences an individual obtains during acquisition and the similar mechanisms used by an individual to interact with the environment in general (Elman et al., 1996; Elman, 2005).

There are phylogenetic scenarios respectively following these two hypotheses. Based on the nativism, Bickerton proposed *a bootstrapping scenario* (Bickerton, 1990). Language is assumed to start from a *protolanguage* (the prototypical language first used by our last common ancestor in the hominin evolution), which consists of a few words with no syntactic structure (Bickerton, 2008). Genetic changes led to a language faculty for individuals to innovate modern languages (Kirby et al., 2007). The transition from words to word concatenation, and to hierarchical syntactic structure, was *synthesis*, with predefined linguistic functions being applied to protolanguage. Following the emergentism, Wray proposed *a formulaic scenario* (Wray, 2002). Protolanguage is assumed to be *holistic*, composed of 'holophrases' or 'unitary utterances' symbolizing frequently occurring situations (Arbib, 2005, 2008). Segmentation arose from recognition of chance associations between phonetic segments in holistic utterances and objects or events to which these segments related. Via segmentation, individuals divided unitary utterances into sub-units and acquired rules regulating recombination of these subunits. Linguistic functions were developed mainly via cultural training, on the basis of institutionalized *post hoc* rationalization (Wray, 2007). The transition from holistic utterances to lexical items and syntax was *analytic*, with novel linguistic features being acquired based on linguistic instances and general learning mechanisms.

© 2011. John Benjamins Publishing Company
All rights reserved

Both of these scenarios are problematic in some aspects (Arbib & Bickerton, 2008). The bootstrapping scenario disregards the differences in linguistic environments between phylogeny and ontogeny. As for early hominins, the exchanged meanings in primitive communications would have mainly included frequent events in their environment, and these events could have been encoded by holistic messages with generic references (Arbib, 2008; Wray, 2002; Smith, 2008); as for modern children, the linguistic input in the early stage of acquisition is often composed of isolated words with specific references and accompanied by gestures to reinforce the conveyed messages (Volterra et al., 2005). Regarding this, the 'ontogeny recapitulating phylogeny' analogy has to be cautiously applied in evolutionary linguistics (Mufwene, 2008). Meanwhile, the formulaic scenario commits two mistakes. First, some proposed holistic examples[1] involve complex semantic structures, whose meanings are difficult for modern humans to handle, not to mention early hominins with limited semantic and/or syntactic competences (Tallerman, 2007, 2008). Only highly salient, relatively simple meanings can be associated with unitary, unstructured utterances, and be reconstructed with few linguistic or non-linguistic cues (Smith, 2008). Second, it is misleading to use idioms from modern languages to characterize features of protolanguage (Croft, 2001), because these idioms can be syntactically analyzed (Kirby, 2007; Tallerman, 2008) and their literal meanings are activated at least in production (Cutting & Bock, 1997).

Considering these, we propose *a coevolutionary scenario* to modify the drawbacks in these scenarios. Our scenario states that:

1. Language origin proceeded in *a mosaic fashion* (Wang, 1982), with general abilities from nonlinguistic modules contributing significantly to linguistic universals. In particular, the abilities of pattern extraction and sequential learning triggered the acquisition of lexical items and simple word orders;
2. Compositionality and syntactic regularity *coevolved* during the transition from a holistic protolanguage to a compositional one, driven by mutual understanding in primitive communications exchanging meanings with simple predicate-argument structures.

Linguistic compositionality refers to the principle on how the meaning of a complex expression is built from its subparts via regulating rules (Ehrlich & Raven, 1964; Krifka, 2001). It concerns both lexical and grammatical aspects of language. In order to separate these aspects and explore their relations, we clarify *compositionality* as the building of integrated meanings in sentential expressions from the meanings of their lexical constituents, and *regularity* as *the basic constituent word order* (Tomlin, 1986), which is the prototypical order of *subject* (*S*, the primary syntactic relation with respect to the verb), *object* (*O*, the secondary syntactic relation with respect to the verb), and *verb* (*V*) in simple declarative sentences.

© 2011. John Benjamins Publishing Company
All rights reserved

Coevolution in biology denotes the mutual reciprocal or cooperative influence between two or more species (Thompson, 1994). In languages mainly adopting word order to clarify semantic roles, such as English or Chinese, lexical items and word orders also form a coevolutionary relation. First, linguistic understanding requires exchanged utterances to be formed by identical lexical items that follow similar orders. Second, both lexical items and word orders can be viewed as recurrent patterns in different linguistic domains: lexical items are identical syllables in utterances that have come to be associated with identical atomic meanings in semantics; and word orders are identical sequential relations among lexical items in utterances. Finally, evidence on online language processing (Bates et al., 1996) and language acquisition (Bates & Goodman, 1997; Fillmore, 1979) has revealed an inseparable development of lexical items and word orders, together with relevant processing mechanisms.

This scenario assumes a holistic protolanguage capable of describing simple integrated events that involve probably a single predicate and its related arguments, e.g. "a wolf is running" or "a lion is chasing a gazelle". Such meanings can be reliably associated with unitary utterances and reconstructed with little linguistic or nonlinguistic information. Holistic utterances encoding these meanings are sufficient for individuals to develop preliminary linguistic knowledge based on socio-cognitive intelligence (Smith, 2005). With such unstructured utterances being segmented into sub-units, structural information, such as the order relations of sub-units, could emerge. Along with this analytic process, a synthetic process also occurs; individuals, following simple orders, can combine available lexical items to encode integrated events.

This scenario suggests that some abilities exapted from other cognitive modules could have served as the precursors of linguistic abilities to process lexical items and word orders. For example, *the pattern extraction ability* (the ability to note and abstract recurrent patterns occurring in instances from multiple sources) helps acquire lexical items. It is present in vocal learning of singing species (e.g. Bengalese finches, Okanoya & Merker, 2007), shown in humans during some cognitive tasks (Monchi et al., 2001), and revealed in pre-linguistic children when they extract regularities across exemplars (Saffran et al., 1996; Marcus et al., 1999). Meanwhile, *the sequential learning ability* (the ability to encode and represent the order of discrete elements occurring in a temporal sequence, Christiansen & Chater, 2008; Reali & Christiansen, 2009) helps manipulate word orders. It is found in chimpanzees or monkeys (Fitch & Hauser, 2004; Hauser et al., 2001; Terrace, 2002) in grooming (Aldridge & Berridge, 2003) or other memory-based planning activities (Shima et al., 2007). The brain-imaging evidence on humans shows that the processing of language (Friederici et al., 2002; Petersson et al., 2004) shares similar sequential learning mechanisms with the processing of music (Maess et al., 2001; Patel et al., 1998; Reali & Christiansen, 2009), and

© 2011. John Benjamins Publishing Company
All rights reserved

the perisylvian area of the human brain houses anatomically neuronal wirings for forming sequence detectors among word pairs in language (Pulvermüller & Knoblauch, 2009). Together with the pattern extraction ability, the sequential learning ability makes possible the acquisition of *local orders* (orders between two lexical items either adjacently or separately) and the iteration of local orders into *global orders* (orders between more lexical items at the sentence level). The neural bases of the iteration involve cerebellum and basal ganglia that regulate motor control (Braitenberg et al., 1997; Lieberman, 2006).

Such 'from local to global' process of word order development indicates that word order regularity could have emerged in a constructivist manner. Humans refer to local or partial information when processing language and judging the grammaticality of sentences (Munn, 1999; O'Grady, 2008). In addition, global order knowledge has to be acquired based on local, piecemeal information. This is evident by the fact that children, in their early stage of language acquisition, cannot directly generalize the global order from one verb to another, nor use it as a cue to comprehend sentences with novel verbs (Akhtar & Tomasello, 1997). Furthermore, word order variations in natural languages imply that global orders are subject to changes at the local order level (Steele, 1978). All these indicate that local processing underlies our language capability and a full-fledged grammar is preceded by an intermediate, linear structure based on local syntactic constraints (Lyon et al., 2007).

Finally, the 'tinkerer' view *à la* Jacob (1977) suggests that other than producing novelties from scratch evolution usually transforms an available system to give it new functions or combines available approaches to produce a more elaborate one. The existent structures, systems, or mechanisms previously for one purpose can be exapted to take on new tasks, though the results are not necessarily simple or elegant (Lieberman, 2006). From this point of view, language could be "a complex reconfiguration of ancestral systems that have been adapted in evolutionarily novel ways" (Fisher & Marcus, 2005, p.9); the abilities of language users could have resulted from general competences originally developed from other cognitive modules, and the accumulation of piecemeal, communicative strategies could gradually lead to the full syntax of language via generalization (Arbib, 2005).

2. Simulation studies on evolution of compositionality and regularity

Due to lacking direct evidence, research in language origin, has *an interdisciplinary nature*; knowledge, findings, and approaches from other relevant fields all contribute significantly to our understanding of language and its evolution. Apart from theoretical argumentation as in the previous section, *computer simulation* has recently joined the endeavor to tackle problems of language evolution (see Gong, 2009; Lyon et al., 2007; Wagner et al., 2003 for overviews). By creating

© 2011. John Benjamins Publishing Company
All rights reserved

abstract versions of language evolution or simulating individual behaviors during communications, computer models can help evaluate the internal coherence of linguistic scenarios, and even generate new theories on language evolution (Oudeyer, 2006). This line of research dated back to the work of Hurford (1989). He built up a computer model showing that a 'Saussurean' learning procedure (agents used their learned responses to signals to decide what to send) could lead to a coordinated signaling system. This procedure, also abstracted in mathematical models (e.g. Nowak et al., 1999), has been widely used in later studies on lexical evolution.

An early study on the origin of compositionality was done by Batali (1998). In his model, semantic constituents in phrases such as "me hungry" or "you tired" could be encoded by characters, and individual linguistic knowledge on how to process these characters were stored in each agent's artificial neural network. After a number of communications, agents started to use fixed characters to encode particular constituents and certain character(s) became a root in utterances encoding meanings with identical constituents. Later on, Munroe & Cangelosi (2002) proposed the mushroom foraging model, also using artificial neural networks to encode linguistic knowledge. This model simulated how agents learned to clarify mushrooms encountered in their lifetime. During learning, visual representations of encountered mushrooms and/or linguistic descriptions of them used by their parents were input to learners' networks. These inputs helped to develop descriptions and appropriate actions for edible mushrooms, which formed the outputs of learners' networks. After generations of repeated learning, a "noun plus verb" language emerged; one part of it was used to clarify the edibility of mushrooms, and the other part to indicate the proper actions for edible mushrooms. Both of these models showed that a compositional language with structural regularity could emerge via distributed communications. However, since the underlying mechanisms in artificial neural networks remain unclear, this way of encoding linguistic knowledge limits our understanding on the evolution of compositionality and relevant learning abilities.

Other models adopted rule-based systems to encode linguistic knowledge and simulated individual behaviors to process different types of linguistic rules. For example, Briscoe (1997) developed an acquisition model that traced the origin of not only a communal language, but also learners whose parameter-settings favored the spoken language in their environment. Kirby and colleagues (Kirby, 1999; Smith et al., 2003) designed the Iterated Learning Model (ILM) and its extended versions. These models simulated the pattern extraction ability for agents to acquire lexical items from holistic utterances. Given a bottleneck that restricted linguistic instances transmitted among individuals across generations, a compositional language could emerge, which consisted of a set of lexical items and some grammatical structures encoding simple integrated meanings or complex

© 2011. John Benjamins Publishing Company
All rights reserved

ones with embedded structures. Vogt (2005) proposed the guessing game model to simulate the emergence of compositional structure in semantics. This *situated model* (Wagner et al., 2003) simulated how semantic concepts were acquired via categorization in different conceptual domains. Finally, Steels (2004) designed the *Fluid Construction Grammar* to explore the origin of morphosyntactic structures in an artificial language.

Apart from these behavioral models, Bayesian learning models were designed to study the evolution of compositionality. In these mathematical models, language was abstracted into different types of hypothesis and communication became a process of Bayesian learning, during which agents combined the *a priori* inductive biases for certain types of hypothesis with the evidence shown in the available data to compute an *a posteriori* distribution over hypotheses that generated the available data. Via repeated learning, agents gradually converged to a distribution similar to their *a priori* inductive biases (Griffiths & Kalish, 2007). Given a certain degree of bottleneck, if agents always acquired the type of hypothesis that could generate the most frequent type of the data in the available samples in each generation, the weak *a priori* bias for such type of hypothesis could be gradually magnified after generations (Kirby et al., 2007).

3. The compositionality-regularity coevolution model

We propose a multi-agent behavioral model to illustrate the coevolutionary scenario between lexical items and word orders. In a nutshell, *individuals* (artificial agents) in this model communicate with each other about integrated events such as “a wolf is running” or “a tiger is chasing a deer” by exchanging utterances. In a communication, a *speaker* produces some utterances to encode some integrated events, and a *listener* receives these utterances plus some unreliable environmental information to assist comprehension. Agents are equipped with the learning abilities of pattern extraction and sequential learning, and the mechanisms to produce and comprehend utterances. Based on these competences, agents can acquire, update and align their linguistic knowledge with each other. After a number of repeated communications, a common set of linguistic knowledge becomes shared among agents, and a communal language consisting of a set of common lexical items and consistent word order(s) emerges out of an initial holistic signaling system. This process, analogous to the origin of language in a population of individuals, recapitulates the major steps during the origin of compositionality and word order regularity, and the simulation results help evaluate if the incorporated learning abilities in agents are sufficient to develop the linguistic features in the communal language.

© 2011. John Benjamins Publishing Company
All rights reserved

In the following sections, the major components of this model are introduced, the biological bases of the adopted abilities are discussed, and some simulated mechanisms are compared with those in other models that explore similar questions.

3.1 Artificial language

Language in this model is represented by a set of mappings between meanings and utterances (*M-U mappings*). Meanings are from a semantic space and shared by all individuals in the population. *The semantic space* contains a finite set of *integrated meanings* denoted by simple predicate-argument structure. *Predicates* refer to actions that individuals can conceptualize (e.g. "run" or "chase") and *arguments* refer to entities on which and by which those actions are performed (e.g. "fox" or "tiger"). Some predicates take a single argument, e.g. "run<tiger>" (meaning "a tiger is running"). Others take two, e.g. "chase<tiger, fox>" (meaning "a tiger is chasing a fox"); in this case, the first constituent within < >, "tiger", denotes the *agent* (the entity that instigates the action) of the predicate "chase", and the second, "fox", denotes the *patient* (the entity that undergoes the action). Note that reflexive meanings with identical arguments, e.g. "chase<deer, deer>", are excluded.[2] These predicates form two types of integrated meanings: "*predicate<agent>*" and "*predicate<agent, patient>*".

An utterance that encodes an integrated meaning or its constituent(s) is formed by a string of syllables chosen from *a signaling space*. An utterance that encodes a whole integrated meaning can be segmented into subparts, each mapping a particular constituent; similarly, subparts that map constituents can combine to encode an integrated meaning, provided that these subparts specify each constituent in that integrated meaning only once. For example, subparts that respectively encode "chase<tiger, #>" and "fox" can combine to encode "chase<tiger, fox>", here "#" denotes the unspecified arguments of the predicate. However, subparts that respectively encode "chase<tiger, #>" and "fight<#, wolf>" cannot combine, since they specify the predicate twice. Using combinable syllables to denote linguistic utterances has been adopted in many *structured simulations* (Wagner et al., 2003).

Individuals can conceptualize semantic constituents and their thematic roles (predicate, agent, or patient) in integrated meanings. A primitive, structured conceptual system must have been established before language via sensory channels other than acoustics (Damasio & Damasio, 1992). Both pre-linguistic children (Tomasello, 2003) and social animals, such as crickets (Hoy & Paul, 1973), domestic chickens (Karakashian et al., 1988), vervet monkeys (Cheney & Sewfarth, 2005), and capuchin monkeys (Digweed et al., 2005), possess a rich cognitive representation system not fundamentally different from that of human adults (Fitch, 2005; Hurford, 2007; Cheney & Seyfarth, 2007). The neuroscience studies also reveal that the neural correlates of simple predicate-argument structure exist in non-human

© 2011. John Benjamins Publishing Company
All rights reserved

primates (Fitch et al., 2005). All these suggest that the mental organization of perceived events or situations is a fundamental aspect evolving before the origin of humans (Hurford, 2009).

3.2 Individual linguistic knowledge and acquisition mechanisms

In this model, based on equipped learning mechanisms, individuals can acquire linguistic knowledge from M-U mappings obtained in previous communications. Linguistic knowledge is encoded by three types of rules: *lexical rules,*[3] *syntactic rules*, and *syntactic categories*. Some examples are shown in Figure 1.

Lexical rules

Holistic rules:

(a) “chase<wolf, bear>”←→/a b/ (0.5)
(b) “hop<deer>”←→/c/ (0.4)
(b) “hop<deer>”←→/d e/ (0.6)

Compositional rules:

(d) “wolf”←→/f/ (0.6)
(e) “run<#>”←→/c/ (0.7)
(f) “chase<#, bear>”←→/e f * g/ (0.7)

Syntactic rules

(1) {“chase<#, bear” ←→/e f * g/ (0.7)} ▼{“fox”←→/a/ (0.5)} (0.7)
(2) {“wolf” ←→/f/ (0.6)}▲{“chase<#, bear”←→/e f * g/ (0.7)} (0.5)
(3) Category 1 (S)<< Category 2 (V) (0.8)
(4) Category 3 (O)<< Category 2 (V) (0.4)

Categories

Category 1 (S): *List of lexical rules:*
{“fox”←→ /a/ (0.5)} [0.5]
{“wolf”←→ /b c / (0.7)} [0.5]
List of syntactic rules:
Category 1 (S)<< Category 2 (V) (0.8)
Category 3 (O)<< Category 2 (S) (0.4)

Category 2 (V): *List of lexical rules:*
{“run<#>”←→ /d/ (0.4)} [0.5]
{“fight<#,#”←→ /e / (0.7)} [0.5]
List of syntactic rules:
Category 1 (S)<< Category 2 (V) (0.8)

Category 3 (O): *List of lexical rules:*
{“fox”←→ /a/ (0.5)} [0.7]
{“sheep”←→ /h/ (0.6)} [0.6]
List of syntactic rules:
Category 3 (O)<< Category 1 (S) (0.4)

Figure 1. Examples of lexical rules, syntactic rules, and syntactic categories. Lexical rules are itemized by letters, and syntactic rules are itemized by numbers. “Category 1”, “Category 2” and “Category 3” are categories; “S”, “V”, and “O” are their syntactic roles. Numbers enclosed by ( ) are rule strengths, and those by [ ] are association weights. “<<” is the local order “before”, “>>” is “after”, “▼” is “surround”, and “▲” is “between”

3.2.1 *Lexical rules*

A lexical rule includes an M-U mapping. In production, it can be activated if its meaning fully or partially matches the speaker’s integrated meaning; in comprehension, it can be activated if its utterance fully or partially matches the listener’s heard utterance. Some lexical rules are *holistic*, each mapping an integrated meaning

© 2011. John Benjamins Publishing Company
All rights reserved

onto an utterance (*a sentence*) (rules (a)–(c) in Figure (1); others are *compositional*, each mapping particular semantic constituent(s) onto subpart(s) of an utterance. Compositional rules include *word rules* (e.g. rules (d) and (e) in Figure 1, each mapping a single constituent onto an utterance) and *phrase rules* (e.g. rule (f) in Figure 1, each mapping two constituents that cannot form an integrated meaning onto an utterance).

Natural languages contain *synonyms* (expressions having different pronunciations but the same meaning, e.g. /freedom/ and /liberty/ in English)[4] and *homonyms* (expressions having the same pronunciation but different meanings, e.g. [yuán yīn] meaning either /原因/ (reason) or /元音/ (vowel) in Chinese), which respectively correspond to *synonymous* (e.g. lexical rules (b) and (c) in Figure (1) and *homonymous rules* (e.g. lexical rules (b) and (e) in Figure (1). These kinds of lexical items increase the complexity of language processing. Several mechanisms to avoid them, such as *the principle of contrast* (Clark, 1987) and *mutual exclusivity* (Markman, 1989), were simulated in some models (e.g. Smith, 2005). But due to these mechanisms, the emergent communal languages were free of synonyms or homonyms, which contradicted what was observed in natural languages. In this model, no direct avoiding mechanisms towards synonyms or homonyms are adopted, and the results show that only some kind of homonyms can be preserved in the communal language.

3.2.2 *Syntactic rules*

Holistic rules allow individuals to directly produce meaningful sentences, but compositional rules have to be sequentially regulated to form a meaningful sentence. The orders of words or phrases in utterances are regulated by *syntactic rules*, each containing a relative order between two lexical items respectively from two categories. Four types of local orders are defined: (1) *before*, a word rule or a continuous (without unspecified syllables in between) phrase rule precedes (not necessarily adjacently) another; (2) *after*, a word rule or a continuous phrase rule follows (not necessarily adjacently) another; (3) *surround*, a discontinuous (with unspecified syllables in between) phrase rule surrounds a word rule; and (4) *between*, a word rule is surrounded by a discontinuous phrase rule. As shown in Figure 1, syntactic rule (1) means that the utterance of the phrase rule surrounds the utterance of the word rule, and syntactic rule (3) means that the utterances of lexical rules from Category 1 are before those from Category 2. Some other models considered only adjacent orders (e.g. Pulvermüller & Knoblauch, 2009) or excluded categorical knowledge (e.g. de Pauw, 2006) when building order rules.

3.2.3 *Syntactic categories*

Categories are culture-dependent components characterized by a set of properties that are shared among category members (Barnes, 1984). Categorization of objects

© 2011. John Benjamins Publishing Company
All rights reserved

based on explicit or implicit criteria is a natural and adaptive process shown in many species. For example, pre-linguistic children can categorize visual stimuli based on colors or shapes (Tomasello, 2003), and honeybees can form concepts such as sameness or difference across different types of stimuli (Giurfa et al., 2001). Categorization eases effort of information processing, thereby providing evolutionary advantages (Kuczaj II & Hendry, 2003).

In this model, syntactic categories are formed in order for syntactic rules acquired for some lexical items to be applied productively to others that have the same thematic role. *A syntactic category* comprises a set of lexical rules and a set of syntactic rules that regulate the orders between these lexical rules and those from other categories. A syntactic category that associates lexical rules encoding the thematic role of agent is referred to as a subject (S) category, since the thematic role of agent corresponds to the syntactic role of subject. Similarly, patient corresponds to object (O), and predicate to verb (V). In other words, the language simulated here is nominative-accusative, and all sentences are in active voice.[5] A syntactic rule between two categories can be denoted by their syntactic roles, e.g. a local order before between lexical items from an S category and a V category can be denoted by S<<V, or simply SV.

A single lexical rule can be a member of multiple categories. For example, the lexical rule encoding "fox" can belong to an agent category and a patient category (see Figure 1), or two distinct agent or patient categories which have different local syntax. Once a lexical rule is associated to a category, its utterance will possess the syntactic role (S, V, or O) of that category in a sentence.

Following the general setting of a rule-based system, we give each lexical or syntactic rule *a strength*, indicating the probability of successfully using its M-U mapping or local order. A lexical rule also has *an association weight* to the category that contains it, indicating the probability of successfully applying the syntactic rules of this category to the utterance of that lexical rule. All strengths and association weights lie in [0.0 1.0]. The strength of a newly-acquired rule is 0.5, so is the association weight of a new association of lexical rule to a category. These numerical parameters make possible a strength-based competition in communications (see Section 3.4) and a gradual forgetting of linguistic knowledge (see Section 3.3).

Lexical and syntactic knowledge collectively encode integrated meanings. Based on the examples in Figure 1, if an individual wants to express "fight<wolf, fox>" using the lexical rules respectively from the three categories and the local orders SV and SO in these categories, the sentence can be either /bcea/ or /bcae/, following a word order SVO or SOV. In this constructivist manner, local orders among categories lead to global orders. This is an efficient way to generate global orders, but it also introduces a certain degree of imprecision, since combining local orders may lead to multiple global orders, e.g. SV plus SO leads to either SVO or SOV, and a global order may be formed by combining different local orders,

© 2011. John Benjamins Publishing Company
All rights reserved

e.g. SVO can be formed by SV plus VO or SO plus VO. This imprecision increases the difficulty in developing a consistent global order in the communal language, and induces order changes and heterogeneity among idiolects.

Different global orders can express the same type of integrated meanings, e.g. two possible orders (SV or VS) can express a "predicate<agent>" meaning and six possible orders (SVO, SOV, OSV, VSO, VOS, OVS) can express a "predicate<agent, patient>" meaning. A systematic mapping between semantic and syntactic structures (*the semantics-syntax correspondence*) requires consistent word orders to express meanings having identical semantic structures. In this model, *a communal language is said to emerge* when a set of common linguistic rules are shared among most individuals in the population; *the semantics-syntax correspondence is said to emerge* when most lexical rules encoding constituents having the same thematic roles are associated with the same categories having the corresponding syntactic roles, and their utterances consistently follow some global orders in sentences.

3.2.4 *Acquisition of linguistic knowledge*

Individuals acquire their linguistic knowledge through random creation and pattern extraction.

Random creation takes place in production. If the speaker (henceforth referred as "she") lacks lexical rules to encode some or all constituents of the intended meaning, under *a creation rate*, she will create randomly a continuous utterance to map the whole meaning and acquire this mapping as a new holistic rule. In the early stage of language origin, new instances are introduced mainly via random creation, without which a language could never get off the ground. Sometimes, all constituents of the intended meaning are expressible, but the available local order rules are insufficient to generate a coherent global order. In this situation, a temporary local order will be used to form a sentence. For example, if the available local order to encode a "predicate<agent, patient>" meaning is SV, the temporary local order can be SO, OS, VO, or OV. Such temporary local order is not acquired immediately as a syntactic rule. New syntactic knowledge is obtained only via pattern extraction.

Other models adopted other forms of creation mechanisms, which were problematic for studying the origin of compositionality. For example, the guessing game (Vogt, 2005) simulated an *exploitation mechanism*; if the speaker could only encode part of an intended meaning, she would invent a lexical rule to encode the inexpressible part of that meaning. This mechanism builds in compositionality; before noting that holistic expressions are decomposable, individuals can already create and synthesize compositional materials commensurate with the subparts of the semantic representation (Kirby, 1999). In addition, the ILM (Kirby, 1999) adopted an *invention mechanism*; the speaker first selected the utterance that she could produce and the meaning of this utterance was closest to the

© 2011. John Benjamins Publishing Company
All rights reserved

intended one, and then, replaced the parts of this utterance that was inconsistent with the intended meaning with some randomly created strings. This mechanism builds in the semantics-syntax correspondence; individuals know that if encoded meanings share similar semantic structures, their expressions should also share similar syntactic structures. The empirical evidence shows that children in the early stage of language acquisition are conservative, sticking to available knowledge and not freely generalizing without positive evidence (Tomasello, 2003). The semantics-syntax correspondence must emerge after sufficient linguistic instances are available and preliminary syntactic structures are formed (Clark, 2003). Similarly, even if early hominins were capable of detecting correspondence across domains, without sufficient linguistic information, they could not instantly note the semantics-syntax correspondence. Therefore, a realistic model studying the origin of syntax should preclude such correspondence.

Pattern extraction is the internal force that triggers segmentation of holistic expressions. *A recurrent pattern* is an identical feature re-occurring in at least two M-U mappings. There are three types of identical features:

1. Identical meaning(s) that correspond to identical syllable(s);
2. Identical local order between one set of lexical rules and another;
3. Identical thematic roles of lexical items.

Each individual has a buffer (see Section 4.3) storing *previous experience* (a finite number of M-U mappings obtained from previous communications in which this individual was the listener). New mappings are compared with those in the buffer before they too are inserted into the buffer. During comparison, under *an extraction rate*, the listener will extract the detected recurrent patterns as linguistic knowledge.

Some examples of pattern extraction are shown in Figure 2. In Figure 2 (a), the listener (henceforth referred as "he") has M–U mapping (1) in his buffer. With the insertion of M–U mappings (2)–(8), he gradually acquires lexical rules (a)–(f). For example, by comparing M–U mappings (1) and (2), he notes that the meaning "fox" is consistently mapped to the syllable /d/, based on which he forms lexical rule (a). Similarly, by comparing M–U mappings (4) and (5), he forms lexical rule (d) to record the recurrent patterns "fight<wolf, #>" and /acb/. However, although M–U mappings (6) and (7) share the meaning "deer" and the syllables /fk/, they are not recurrent patterns, since "deer" has different thematic roles in (6) and (7).

After a number of lexical rules are acquired, categories and syntactic rules can be formed based on recurrent patterns of types (2) and (3). For example, evident in M–U mappings (2), (3) and (7) in Figure 2 (a), /d/ of lexical rule (a), /ac/ of lexical rule (c), and /k/ of lexical rule (f) all precede /m/ of lexical rule (b). Since "wolf", "fox", and "deer" are agents in these meanings, lexical rules (a), (c) and (f) are associated with an S category (Category 1 (S) in Figure 2 (b)). The local order (before)

© 2011. John Benjamins Publishing Company
All rights reserved

with respect to lexical rule (b) is acquired as syntactic rule (I) in Category 1. Similarly, based on M-U mappings (5), (6) and (8), syntactic rule (II) with respect to lexical rule (e) is acquired. Moreover, as shown in M–U mappings (2) and (6), /m/ of lexical rule (b) and /b/ of lexical rule (e) both follow /d/ of lexical rule (a). Since lexical rules (b) and (e) encode predicates in these meanings, these lexical rules are associated with a V category (Category 2 (V)), and syntactic rule (III) is acquired in Category 2 to record the local order (after) with respect to lexical rule (a). Now, since Category 1 and Category 2 respectively associate lexical rule (a) and lexical rules (b) and (e), those syntactic rules can be updated as "Category 1 (S) << Category 2 (V)", which indicates that lexical rules from Category 1 should precede, in utterance, lexical rules from Category 2.

Available M-U mappings	**Newly acquired lexical rules**
(1) "hop<fox>"←→/d h/	(a) "fox"←→/d/ (0.5)
(2) "run<fox>"←→/d m/	(b) "run<#>"←→/m/ (0.5)
(3) "run<wolf>"←→/a c m/	(c) "wolf"←→/a c / (0.5)
(4) "fight<wolf, deer>"←→/a c b e/	(d) "fight<wolf, #>"←→/a c b/ (0.5)
(5) "fight<wolf, gazelle>"←→/a c b m/	(e) "fight<#, #>"←→/b/ (0.5)
(6) "fight<fox, deer>"←→/d f k b/	
(7) "run<deer>"←→/f k m/	(f) "deer"←→/k/ (0.5)
(8) "fight<deer, gazelle>"←→/n k b/	

(a)

Acquired categories and syntactic rules

Category 1 (S): *List of lexical rules:*
Lexical rule (a) [0.5]
Lexical rule (c) [0.5]
Lexical rule (f) [0.5]
List of syntactic rules:
(I) Category 1 (S) << Lexical rule (b) (0.5)
(II) Category 1 (S) << Lexical rule (e) (0.5) ⇨ (I) Category 1 (S) << Category 2 (V) (0.5)

Category 2 (V): *List of lexical rules:*
Lexical rule (b) [0.5]
Lexical rule (e) [0.5]
List of syntactic rules:
(III) Category 2 (V) >> Lexical rule (a) (0.5) ⇨ (I) Category 1 (S) << Category 2 (V) (0.5)

(b)

Figure 2. Examples of pattern extraction: (a) acquisition of lexical rules from M-U mappings in the buffer; (b): acquisition of categories and syntactic rules. M-U mappings are itemized by numbers, lexical rules by letters, and syntactic rules by Roman numerals. Numbers within ( ) are rule strengths, those within [ ] are association weights

Categories with identical syntactic roles can merge, if (a) they share at least one lexical member and (b) there is evidence in the buffer that some of their lexical members are similarly used. The merged category comprises all lexical

© 2011. John Benjamins Publishing Company
All rights reserved

and syntactic members of the original categories. For example, in Figure 3, two S categories share lexical rule (b), though with different association weights. Based on M-U mappings (1) and (2), lexical rule (d) is acquired. Now, in these mappings, /d/ of lexical rule (a) and /k/ of lexical rule (c) both precede /b/ of lexical rule (d), which indicates the similar usage of some lexical members of those S categories. This evidence triggers the merge of these two S categories into one, and a new syntactic rule with respect to lexical rule (d) is acquired.

The formation and merge of categories can induce novel use of lexical rules. In Figure 3, after merge, following syntactic rule (I), lexical rule (c) should precede any lexical rules from Category 3, but there could be no instance in the buffer showing such use, or the available ones may be inconsistent with such use. If it is the latter case, lexical rule (c) may be associated with other S categories. Or, a new local order consistent with most instances involving this lexical item could be acquired in Category 1. With the support of those instances, this syntactic rule may eventually overwhelm other inconsistent ones. Both of these processes reflect the empirical evidence in language acquisition: children do overgeneralize occasionally, but inappropriate generalizations can be gradually corrected based on subsequent linguistic experiences (Brooks & Tomasello, 1999; Bybee, 1995; Goldberg, 1995; Marcotte, 2006). In addition, merge equips individuals with the competence to handle *combinatorial productivity* (van der Velde & Kamps, 2006). In Figure 3, although the individual never encounters an instance containing the constituent "deer" and others encoded by lexical rules from Category 3, this individual can regulate the local order between the utterances of these lexical rules based on the syntactic knowledge in the merged category. This competence makes possible the acquisition of general knowledge from limited instances, or sparse input.

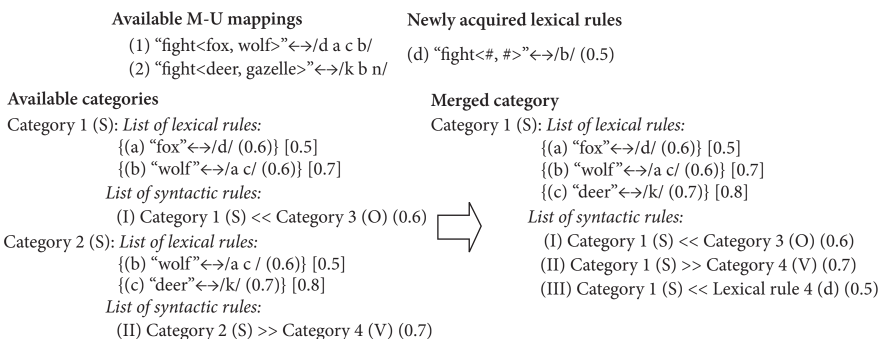

Figure 3. An example of category merge. M-U mappings are itemized by numbers, lexical rules by letters, and syntactic rules by Roman numerals. Numbers within ( ) are rule strengths, those within [ ] are association weights

© 2011. John Benjamins Publishing Company
All rights reserved

All these item-based learning mechanisms are traced in empirical studies (e.g. Mellow, 2008a), and the categorization process resembles *the "Verb-island" hypothesis* (Tomasello, 2003). This hypothesis states that children gradually build up item-based inventories, in each of which some nouns relative to certain verb predicate gradually cluster around that verb (*the verb-island*). The overlap of nouns in some verb-islands triggers merge of these islands, and a general verb category including nearly all frequent verbs can be eventually formed in this way. The mechanisms in our model are more general than this hypothesis. Verb and noun categories coevolve; based on similar use with respect to nouns (or verbs), V (or S and O) categories are acquired. Our mechanisms are also simpler than this hypothesis. The formation and merge of categories in our model are based on local order relations, whereas the criteria for merging verb-islands concern various pragmatic, semantic, and syntactic factors.

In this model, extracting recurrent patterns of type (1) helps acquire lexical rules, thus leading to compositionality; extracting recurrent patterns of types (2) and (3) helps acquire categories and local orders to regulate lexical members of these categories, thus leading to word order regularity; and the formation and merge of categories associate and adjust lexical and syntactic members within and across categories, thus leading to the semantics-syntax correspondence.

Other models adopted other forms of mechanisms to acquire compositional knowledge, which are problematic in some aspects. First, in Kirby (1999), once recurrent patterns in two M-U mappings were identified, compositional rules mapping the residual parts of these mappings were created, together with a grammatical construction that generalized these mappings. In this model, some grammatical constructions were not fully decomposed, even if all their semantic constituents were expressible using compositional rules. In other words, some syllables in these constructions became 'meaningless', having no semantic content. Sometimes, the number of these 'meaningless' syllables could 'explode', surpassing that of meaningful ones. Instead of morphological tags as claimed in Kirby (1999), these 'meaningless' syllables were actually redundant materials due to imperfect decomposition.

Second, in Smith et al. (2003), once a compositional rule was acquired, its utterance position in sentences was also recorded. In this way, lexical rules incorporated both lexical and syntactic information, and the evolution of lexical items and that of syntax became entangled. However, in language acquisition, the development of grammatical knowledge usually commences after sufficient lexical items are acquired (*the critical mass hypothesis*, Marchman & Bates, 1994). In artificial grammar learning, human subjects can transfer grammatical knowledge acquired from the training strings to the testing ones having the same underlying syntax but different surface vocabulary (Vokey & Higham, 2005). Moreover, neural imaging

© 2011. John Benjamins Publishing Company
All rights reserved

studies have showed that syntactic knowledge is represented separately from general verbal memory in different brain areas, and these two respective areas work together to process syntactically complex structures (Friederici et al., 2006).

Finally, in Vogt (2005), only recurrent patterns appearing at the beginning or the end of utterances could be detected. This position constraint implied that pre-linguistic individuals could not grasp other forms of recurrent patterns. However, the pattern extraction abilities present in humans and other species are much advanced. For example, pre-linguistic infants are sensitive to patterns repeatedly occurring in different positions throughout a stream of syllables (Marcus et al., 1999; Saffran et al., 1996). Cotton-top tamarins (Hauser et al., 2007), similar to pre-linguistic infants, are able to learn abstract patterns shown in stimuli. In addition, the mechanism in Vogt (2005) focused only on sentence boundaries, which restricted the possible forms of syntactic structures. If exchanged meanings contained more than three elements, this mechanism would become useless to distinguish patterns in medial positions.

The pattern extraction ability in our model is less sensitive to position, so can be extended to learning complex syntactic structures. And syntactic structures are gradually formed after compositional rules are available, because extracting recurrent patterns of type (1) is a prerequisite to extracting recurrent patterns of types (2) and (3). This developmental process is more realistic than incorporating directly lexical and syntactic information into lexical rules.

3.3 The memory system

Each individual has a *buffer*, storing M-U mappings comprehended in previous communications, and a *rule list*, storing linguistic knowledge derived from instances in the buffer (see Figure 4). This memory system enables a process of 'learning from experience'; in communications, listeners may insert M–U mappings into their buffers, and acquired linguistic rules accordingly into their rule lists, and those rules will be referred to in future communications in production and comprehension. Besides the buffer and rule list, each individual has a *working memory* (Anderson, 1976), which stores activated rules in production or comprehension, and is immediately emptied after those rules are updated by the individual.

Both the buffer and rule list have finite sizes. When a buffer is full, newly inserted mappings replace those acquired earlier; when a rule list is full, newly acquired linguistic rules replace those with weaker strengths. Linguistic rules in the rule list also undergo a *forgetting* process, which occurs regularly after a number of communications (scaled to the population size). During forgetting, all individuals deduct a fixed value from their strengths and association weights. After that, rules with negative strengths are removed from the rule list; lexical rules with negative

© 2011. John Benjamins Publishing Company
All rights reserved

association weights to some categories are removed from those categories; and categories having no lexical members are removed from the rule list, together with their syntactic members.

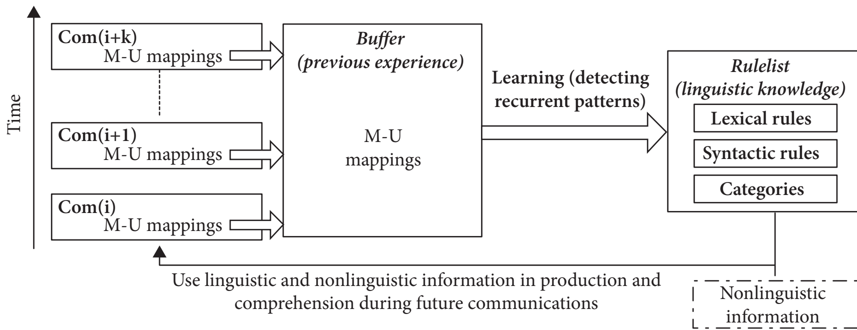

Figure 4. The memory system. Com(i), Com(i+1), Com(i+k) are communications. Dotted rectangle indicates that nonlinguistic information is unreliable

3.4 The communication scenario and rule competition

Individuals apply their linguistic knowledge in communications. In the early stage of language origin, nonlinguistic information acquired via *joint attention* (Herrmann et al., 200; Tomasello et al., 2005) must have helped to grasp exchanged meanings in communications (Tomasello, 2008). In this model, nonlinguistic information is simulated as *environmental cues*, each consisting of an integrated meaning plus a fixed cue strength. Considering *the referential indeterminacy* (Quine, 1960), early hominins, like modern humans, could not always attend to the relevant information in communications. And linguistic communications would become unnecessary if cues always contain the exchanged meanings. Therefore, cues should be unreliable. We define *Reliability of Cue* (*RC*) as the probability for a cue to contain the speaker's meaning. For example, if *RC* is 0.6, there is a 60-percent chance for a correct cue that contains the speaker's intended meaning to be available to the listener; otherwise, *a wrong cue* that contains a meaning distinct from the speaker's intended one is presented to the listener. Such wrong cues were usually disallowed in language game models (e.g. Baronchelli et al., 2006; Puglisi et al., 2008).

A *random communication framework* is adopted in this model. Each communication involves two individuals (a speaker and a listener) randomly chosen from the population, who perform many rounds of *utterance exchange* (see Figure 5). An utterance exchange proceeds as follows:

© 2011. John Benjamins Publishing Company
All rights reserved

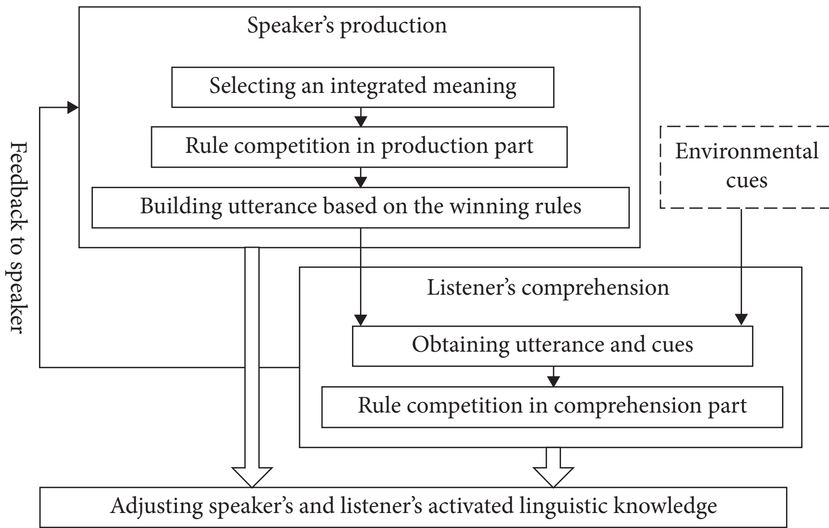

Figure 5. Utterance exchange in a communication. The dotted block indicates that the unreliable information

In production, the speaker first randomly selects an integrated meaning from the semantic space to produce. She then activates lexical rules and their syntactic categories that can encode some or all constituents of this meaning, as well as syntactic rules that can regulate these lexical rules to form a sentence. These rules comprise candidate sets for production, each allowing encoding the whole meaning into a sentence. After that, following Equation 1 (in which, *Avg* means taking average, *aso* taking association weights, and *str* taking rule strengths. "LexRule", "SynRule", "Cat" represents lexical rules, syntactic rules, and categories), she calculates *the combined strength* ($CS_{production}$) of each set:

$$CS_{production} = Avg(str(LexRule(s))) + Avg(aso(Cats) \quad str(SynRule(s))) \qquad (1)$$

$CS_{production}$ is the sum of two components. The first component is the contribution of lexical knowledge, calculated as the average strength of the lexical rules in this set. The second is the contribution of syntactic knowledge, which is the average product of two elements: the first is the strengths of the syntactic rules that regulate these lexical rules, and the second is the association weights of these lexical rules to the categories in this set. After calculation, the speaker identifies the set of *winning rules* with the highest $CS_{production}$, builds up the sentence accordingly, and transmits it to the listener. Random creation (see Section 3.2.4) may occur, if she lacks linguistic knowledge to produce the meaning.

© 2011. John Benjamins Publishing Company
All rights reserved

An example of production is shown in Figure 6.

Meaning to express: “fight<wolf, fox>”

Situations	Activated lexical rules	Activated categories	Applicable syntactic rules	Combined Strength (CS)	Utterance created
1 holistic rule	“fight<wolf, fox>”↔ /a b/ (0.6)			$CS_1 = \{0.6 + Asso\ W_{def} \times SynStr_{def}\} = 0.850$	/a b/
3 word rules	“wolf”↔ /b/ (0.8) [0.6] “fight<#,#>”↔ /c e/ (0.7) [0.8] “fox”↔ /g/ (0.6) [0.7]	Cat1 (S) Cat2 (V) Cat3 (O)	VS (0.7) SO (0.8)	$CS_2 = \{{}^1/_3(0.8 + 0.7 + 0.6) + {}^1/_2({}^1/_2(0.6 + 0.8) \times 0.7 + {}^1/_2(0.6 + 0.7) \times 0.8\} = 1.205$	/c e b g/
1 word rule 1 phrase rules	*“wolf”*↔ /b/ (0.8) [0.6] *“fight<#,fox>”*↔ /e * f/ (0.7) [0.5]	Cat1 (S) Cat4 (VO)	S▲VO (0.8)	$CS_3 = \{{}^1/_2(0.8 + 0.7) + {}^1/_2(0.6 + 0.5) \times 0.8\} = 1.190$	/e b f/

Utterance built up: /c e b g/

Figure 6. An example of production. Numbers within ( ) are rule strengths, those within [ ] are association weights. Cat 1-3 are categories, $CS_{1\text{-}3}$ are $CS_{production}$ of different candidate sets. “▲” is the local order “between”. To calculate $CS_{production}$ of a holistic rule, the syntactic contribution is $AssoW_{def} \times SynStr_{def}$ and $AssoW_{def}$ and $SynStr_{def}$ are both 0.5. The underlined rules and categories form the winning set, since CS_2 is the highest among these sets

In comprehension, the listener receives the sentence from the speaker and an environmental cue. Based on his linguistic knowledge, he activates lexical rules whose syllables fully or partially match the heard sentence, categories that associate these lexical rules, and syntactic rules in those categories whose local orders match those of the lexical rules in the heard sentence. These linguistic rules form candidate sets for comprehension.

Both the environmental cue and the listener’s linguistic knowledge help comprehend the heard sentence. The environmental cue assists comprehension in the following conditions:

1. If the cue’s meaning matches exactly the meaning provided by some linguistic rules, the cue is put together with those rules to form a candidate set;
2. If the available rules fail to provide a complete integrated meaning, but the constituent(s) specified by these rules matches the corresponding one(s) in the cue’s meaning, the cue is put together with these rules to form a candidate set, and its meaning becomes the meaning of this set. For example, if the linguistic rules interpret the heard sentence as “chase<tiger, #>”, with unspecified patient, and the cue’s meaning is “chase<tiger, sheep>”, then, this cue is put together with those rules to form a candidate set, and “chase<tiger, sheep>” becomes the meaning of this set;
3. If there are no linguistic rules, or the available ones fail to form an integrated meaning and the constituent(s) specified by these rules does not match the corresponding one(s) in the cue’s meaning, the cue itself forms a candidate set. For example, the linguistic rules interpret the utterance as “chase<tiger, #>”,

© 2011. John Benjamins Publishing Company
All rights reserved

but the cue has the meaning “fight<tiger, sheep>”. In this case, the cue itself forms a candidate set;

After setting up the candidate sets, following Equation 2, the listener calculates the combined strength of each set. As for a set without a cue, the calculation is identical to that in production; as for a set with a cue, the cue’s contribution, in the form of cue strength, is added to $CS_{comprehension}$; and as for a set with only a cue, the cue strength itself becomes $CS_{comprehension}$.

$$CS_{comprehension} = Avg(str(LexRule(s))) + Avg(aso(Cats) \quad str(SynRule(s))) + str(Cue) \quad (2)$$

After calculation, the listener selects the set of winning rules with the highest $CS_{comprehension}$, and interprets the heard sentence accordingly. If $CS_{comprehension}$ of this set exceeds *a confidence threshold*, the listener adds the perceived M-U mapping to his buffer, and transmits a positive feedback to the speaker. Then, both individuals reward their winning rules by adding a fixed value to their strengths and association weights, and penalize other competing ones by deducting the same value from their strengths and association weights. Otherwise, without adding the perceived mapping to the buffer, the listener sends a negative feedback, and both individuals penalize their winning rules only. For activated rules having initial values of strengths and association weights, the contribution of linguistic (lexical and syntactic) information is 0.75 (0.5 + 0.5 × 0.5). In order to treat equally linguistic and nonlinguistic information, we set both the cue strength and the confidence threshold to 0.75.

Equations 1 and 2 exemplify a multi-level selection among lexical, syntactic, and nonlinguistic information (Steels et al., 2007), and illustrate how nonlinguistic information assists linguistic comprehension, by clarifying constituent(s) unspecified by linguistic knowledge and strengthening linguistic knowledge that leads to a similar interpretation. This communication scenario implements *indirect meaning transfer*: both participants of communications refer to their own linguistic knowledge in production and comprehension, there is no direct check if the speaker’s intended meaning matches the listener’s comprehended one, and comprehension is based on aggregation of linguistic and nonlinguistic information. Nonlinguistic information drives the early acquisition of linguistic materials. And via rule competition, individuals can develop their linguistic knowledge to such an extent that they can solely rely on it to correctly interpret heard sentences and to withstand interference of wrong cues. In this stage, a transition from relying on nonlinguistic information to relying on linguistic information is achieved.

All three types of linguistic rules participate in utterance exchange. An example is shown in Figure 7. In production, the speaker activates the lexical rules to compose “chase<lion, wolf>”, and based on thematic roles of the constituents in the intended meaning (“chase<#, #>” is predicate, “wolf” is agent, and “lion” is

© 2011. John Benjamins Publishing Company
All rights reserved

patient), she activates categories that contain these word rules and have the corresponding syntactic roles. After that, she activates the syntactic rules OS and VO in those categories by which these lexical rules can be regulated. Finally, she judges if this set of linguistic rules wins the competition against other sets. If so, she creates the utterance /abcdef/ accordingly. In comprehension, after the listener identifies the lexical rules whose syllables partially match the heard utterance, the local orders VS and OS that are consistent with the locations of the syllables in this utterance (/ab/ before /ef/ and /d/ before /ef/) are detected. If these orders match the syntactic rules of the categories to which these lexical rules belong, both the categories and their syntactic rules are activated. Then, the thematic roles of these lexical items are specified based on the syntactic roles of those categories ("fight<#, #>" is predicate (V), "lion" is agent (S), and "fox" is patient (O)). The calculation of $CS_{comprehension}$ may involve cue strength if the cue's meaning is also "fight<lion, fox>". Finally, the listener judges if this set of rules is the winning set; if so, he comprehends the heard utterance and determines the feedback accordingly.

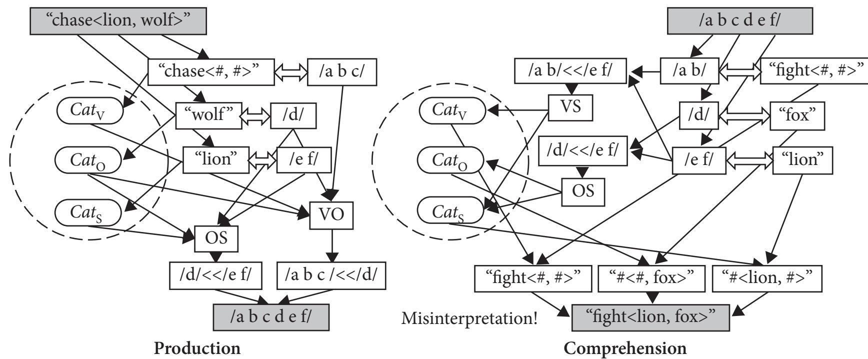

Figure 7. An example of interactions of different types of linguistic rules in production (on the left) and comprehension (on the right). Cat_S, Cat_V and Cat_O are categories with syntactic roles S, V and O. "<<" is the local order "before", and ">>" is "after"

This example illustrates how the conceptual-symbolic (lexical items) and regulatory (local orders and categories) systems work together closely to process linguistic expressions (O'Grady, 2008). In this example, the speaker's intended meaning "chase<lion, wolf>" encoded in /abcdef/ is misinterpreted by the listener as "fight<lion, fox>". Nevertheless, this example reveals the important role of syntactic categories: in production, they transduce semantic structures into syntactic structures, and based on local orders of their syntactic members, global orders are formed; in comprehension, they transduce syntactic structures into semantic structures, and based on local orders of their syntactic members, thematic roles of lexical rules are specified.

The strength-based competition adopted in this model is less dependent on language learning mechanisms. Other forms that can integrate lexical, syntactic,

© 2011. John Benjamins Publishing Company
All rights reserved

and nonlinguistic information in production and comprehension can be adopted as well. Recent neuroscience studies (e.g. Esterman et al., 2009) reveal that some brain areas can serve as a domain-independent source of cognitive control on categorization or stimulus-response mapping rules, and of voluntary shift of attention among perceptual inputs and memory representations. This evidence provides the neural bases for the coordination of various types of information.

Other models adopted different forms of strength-based competition. For example, in Vogt (2005), the score of a set was the product of the strength of each of its rules, and pointing was used to clarify the topic only if the listener failed to comprehend the heard utterance based on her linguistic knolwedge; in Smith et al., (2003), the score of applicable rules was their average association weight in an individual's association network, and speakers' intended meanings were directly transmitted to listeners; and in the Bayesian learning model (Kirby et al., 2007), the score of a linguistic hypothesis was determined by its likelihood based on observed linguistic data and *a priori* probability, leaving out nonlinguistic information. All these mechanisms did not explicitly consider the interaction between linguistic and nonlinguistic information in comprehension.

3.5 The indices to test the performance

We define the following indices to evaluate the coevolution of compositionality and word order regularity:

1. *Rule Expressivity* (*RE*), the average proportion of integrated meanings that all individuals can produce.
2. *Understanding Rate* (*UR*), the average proportion of integrated meanings that each pair of individuals can correctly communicate based only on their linguistic knowledge and without referring to cues. *Disp* is also defined, whose calculation is similar to that of *UR*, except that in each utterance exchange a wrong cue containing a randomly chosen meaning distinct from the speaker's intended one is available to the listener for assisting comprehension. This wrong cue has the same cue strength of 0.75. In other words, *Disp* reflects if individual linguistic knowledge can withstand the interference of incorrect nonlinguistic information.
3. *Understandability of a global order* (UR_{GloOrd}), the average proportion of "predicate<agent, patient>" meanings that each pair of individuals can correctly communicate based on lexical knowledge and a particular global order (SVO, SOV, OSV, VSO, VOS, or OVS), without referring to cues.
4. *Understandability of a local order* (UR_{LocOrd}), the average proportion of "predicate<agent, patient>" meanings that each pair of individuals can correctly communicate based on lexical knowledge and a particular local order (SV, VS, VO, OV, SO, or OS), without referring to cues.

© 2011. John Benjamins Publishing Company
All rights reserved

High UR_{GloOrd} of a global order or high UR_{LocOrd} of a local order indicates the *prevalence* of these orders in the emergent language. As for "predicate<agent>" meanings, the local and global orders are identical, hence UR_{GloOrd} and UR_{LocOrd} are not distinguished.

4. The simulation results

Table 1. The parameter setting. The effects of these parameters on language origin are discussed in Gong (2009)[6]

Parameters	Values
semantic space size	64
utterance space size	30
buffer size	40
rule list size for lexical rules	60
rule list size for categories	20
rule list size for syntactic rules	20
creation rate	0.25
extraction rate	0.25
amount of rule adjustment in competition	0.1
amount of rule adjustment in forgetting	0.01
reliability of cue	0.6
population size	10
number of communications	6000
number of utterance exchange per communication	20

Table 1 lists the parameter setting in the simulations reported in this paper. The 64 integrated meanings in the semantic space are formed by 12 constituents (4 as agents or patients, 4 as single-argument predicates, and 4 as double-argument predicates). There are in total 16 (4×4) "predicate<agent>" and 48 (4×4×(4–1)) "predicate<agent, patient>" meanings. Individuals have an equal chance to select any of these meanings in production. In each simulation, the initial signaling system consists of 8 holistic rules shared by all individuals. These rules can express 8 distinct integrated meanings that involve all 12 constituents, and their utterances are selected randomly from the signaling space.[7] In a simulation, there are 6000 (600 rounds per agent) communications among pairs of individuals, and forgetting occurs after each round of communications.

Twenty simulations are conducted under this setting, and the results of two simulations are shown respectively in Figure 8 and Figure 9, in which panel (a) traces the origin of compositionality using *RE* and *UR*, and panels (b)–(f) trace the origin of word order regularity using UR_{GloOrd} and UR_{LocOrd}. In all twenty

© 2011. John Benjamins Publishing Company
All rights reserved

simulations, a communal language consisting of a set of common lexical rules and consistent global orders emerges around 486 (±29.963, standard error) rounds of communications, and its *UR* is around 0.91 (±0.029), indicating a good communication accuracy among individuals.

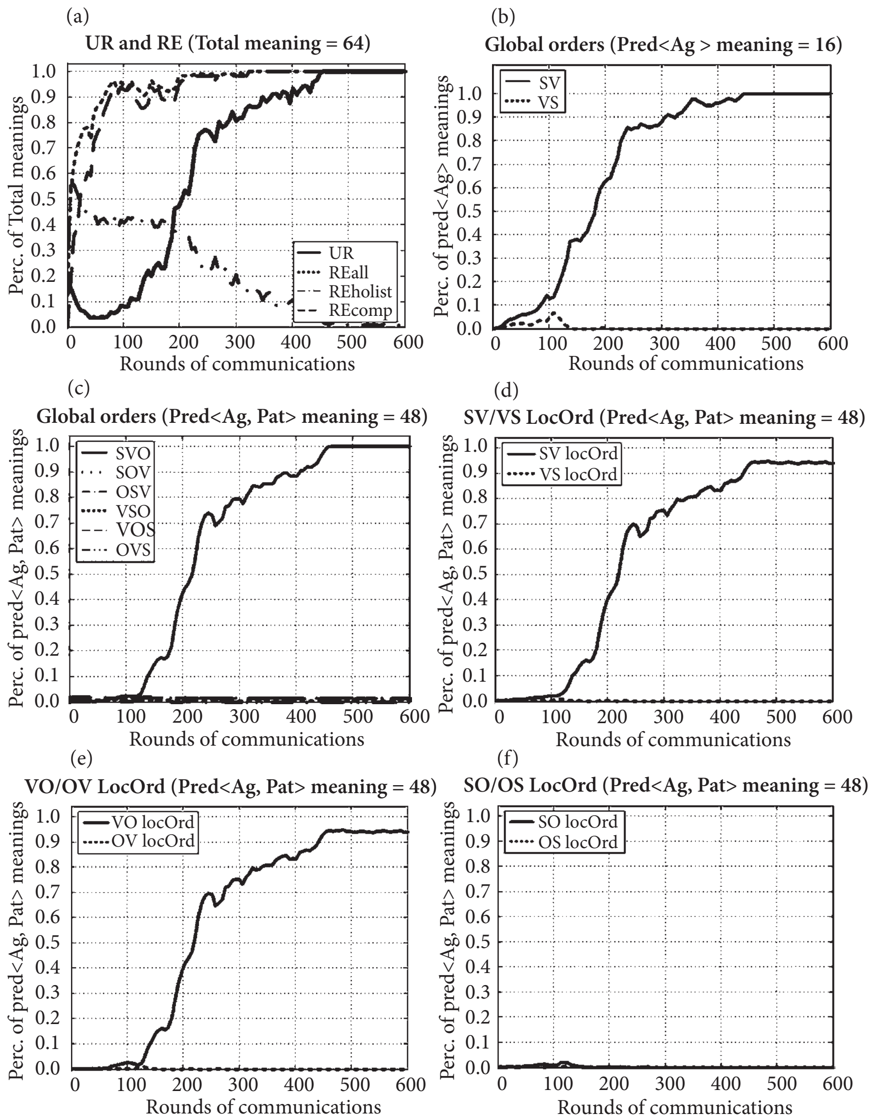

Figure 8. The results in Simulation 1: (a) *UR* and *RE*; (b) UR_{GloOrd} for "predicate<agent>" meanings; (c) UR_{GloOrd} for "predicate<agent, patient>" meanings; (d)–(f) UR_{LocOrd} for "predicate<agent, patient>" meanings

© 2011. John Benjamins Publishing Company
All rights reserved

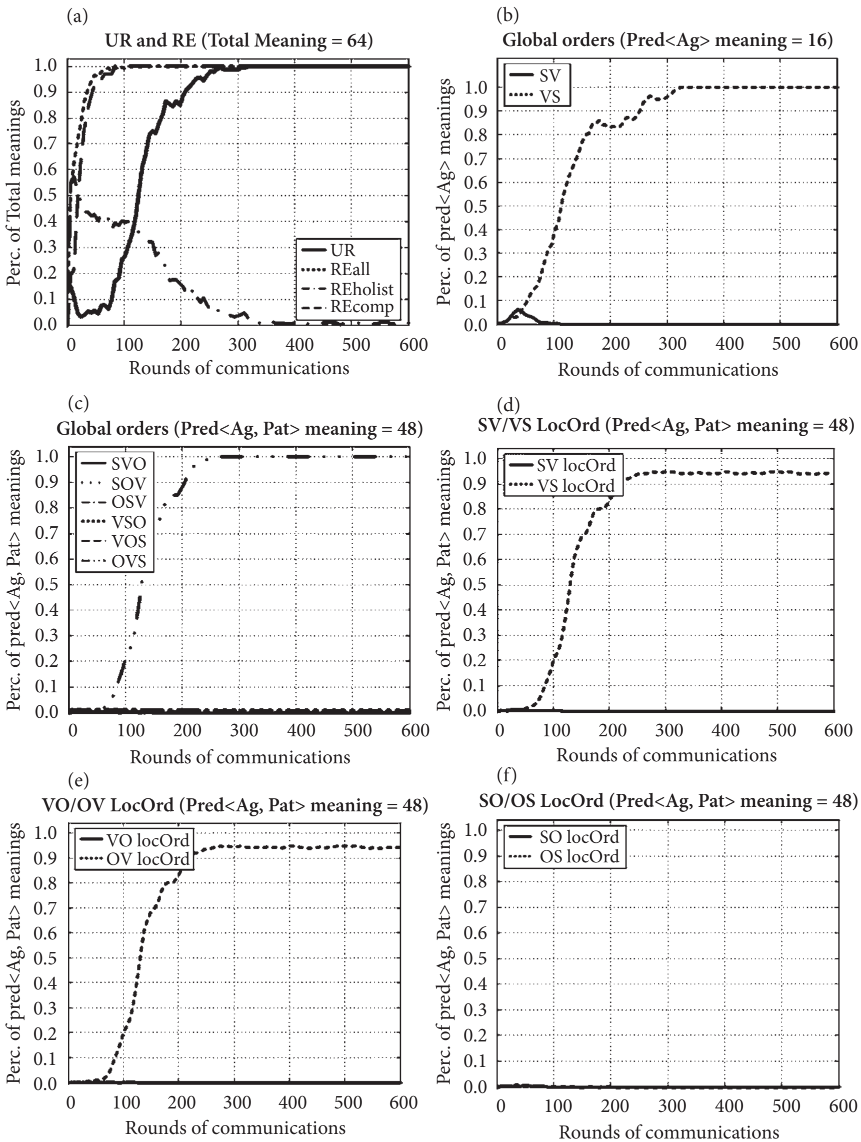

Figure 9. The results in Simulation 2: (a) *UR* and *RE*; (b) UR_{GloOrd} for "predicate<agent>" meanings; (c) UR_{GloOrd} for "predicate<agent, patient>" meanings; (d)–(f) UR_{LocOrd} for "predicate<agent, patient>" meanings

As shown in Figure 8 (a) and Figure 9 (a), *RE* of holistic rules (the dash-dotted line) starts from 0.125 (based on the eight initial holistic rules), and increases along with *RE* of all lexical rules (the dotted line) in the early few rounds of communications, which indicates that the exchanged sentences in this stage are

© 2011. John Benjamins Publishing Company
All rights reserved

mainly formed by randomly created holistic rules. The increased scope of linguistic instances causes recurrent patterns to appear more easily, which allows individuals to extract some of these patterns as compositional rules. This causes the increase in *RE* of compositional rules (the dashed line) from 0.0. Then, the competition between compositional and holistic rules begins, both inside and among idiolects. A holistic rule can only express one integrated meaning, whereas a compositional rule, due to combination, can potentially express many integrated meanings that involve the constituent(s) encoded by this compositional rule. This makes compositional rules be referred to more frequently than holistic rules in communications. Then gradually, compositional rules win the competition against holistic rules. As shown in these two figures, after some rounds of communications, *RE* of holistic rules drops, but that of compositional rules increases and approaches 1.0.

These two figures also trace *UR* (the solid lines). In the early stage, *UR* increases slightly along with *RE* of holistic rules. This indicates that comprehension in this stage is mainly based on holistic rules, and low *UR* shows that few holistic rules are shared among agents. With the origin of compositional rules, an explicit drop of *UR* is observed, which reflects the competition between holistic and compositional rules. During competition, integrated meanings originally expressed by holistic rules are sometimes misunderstood as they begin to be expressed by compositional rules. After compositional rules win the competition, *UR* starts to increase sharply; once an additional compositional rule is shared, individuals can accurately exchange many additional integrated meanings. As more compositional rules come to be shared, *UR* gradually approaches 1.0, along with *RE* of compositional rules. This whole process matches the general pattern in language acquisition (Clark, 2003; Fillmore, 1979), and has also been traced in development of morphosyntactic features (e.g. the past tense in English verbs, Rumelhart & McClelland, 1987).

High compositional expressivity and high understanding rate jointly indicate the origin of compositionality, i.e. all individuals share a common set of compositional rules by which to accurately express most integrated meanings. In addition, the increase in *UR* always takes place after the increase in *RE* of compositional rules, which illustrates that high *RE* alone is not enough to indicate compositionality, since it is uncertain whether the expressions produced by compositional rules can be accurately comprehended.

As shown in Figure 8 (b)–(c) and Figure 9 (b)–(c), along with the origin of compositional rules, UR_{GloOrd} of some global orders gradually increases. These prevalent global orders reflect the sequential information specified by local orders. In Simulation 1, most individuals develop SV and VO to comprehend "predicate<agent, patient>" meanings (see Figure 8 (d)–(f)), thus leading to the prevalence of SVO (see Figure 8 (c)). Since lexical rules encoding single- and

© 2011. John Benjamins Publishing Company
All rights reserved

double-argument predicates are both associated with V categories, SV is also used as the 'global order' to comprehend "predicate<agent>" meanings (see Figure 8 (b)). Similarly in Simulation 2, most individuals develop VS and OV, thus leading to the prevalence of OVS and VS.

These figures collectively trace a coevolutionary origin of compositionality and word order regularity. The sharp increase in *UR synchronizes* with that in UR_{GloOrd} of the prevalent global order; both take place around similar rounds of communications, and the increasing tendencies are similar. These figures also reveal a 'bottom-up' syntactic development; the prevalent global orders are constructed based on the local sequential information encoded by the prevalent local orders.

The coevolution of compositionality and word order regularity is driven by the following forces:

1. *Nonlinguistic information*, which assists comprehension and triggers the acquisition of linguistic knowledge in the early stage of language origin;
2. *Semantic similarity* (integrated meanings may share some constituents having the same thematic roles), which helps trigger the segmentation of holistic expressions;
3. *Mutual understanding*, which requires exchanged utterances to be formed by common compositional rules that follow similar orders.

Throughout twenty simulations, the emergent language usually develops SVO (in 11 simulations) or OVS (in 9 simulations) to comprehend "predicate<agent, patient>" meanings. These orders are formed by two local orders, one regulating S and V, and the other regulating V and O. The origin of SV/VS is due to the fact that agent arguments and predicates are more frequent than patient arguments in the semantic space. This non-uniform distribution of semantic constituents triggers an early development of S and V categories and local orders that regulate their lexical members. The origin of a VO/OV is caused by both internal and external factors. The internal factor deals with communication mechanisms: in comprehension, to clarify thematic roles (especially agent and patient arguments), listeners have to explicitly or implicitly develop a local order to regulate S and O (*the clarifying requirement*); in production, speakers tend to economically use two local orders to form a global order to produce intended meanings (*the economy requirement*). The external factor concerns the fact that every individual has a roughly equal chance to be a speaker and a listener. Then, both the clarifying and economy requirements can sufficiently affect the development of word order regularity.

Combining these factors, the bias toward certain prevalent global orders can be predicted as follows: Suppose that SV wins the competition over VS. Then, there

© 2011. John Benjamins Publishing Company
All rights reserved

are 4 sets of possible local orders, as shown in Table 2: (a) SV plus VO; (b) SV plus OV; (c) SV plus SO; and (d) SV plus OS. Among these sets, only sets (a) and (d) match both the clarifying and economy requirements. Set (b) does not match the clarifying requirement; it is impossible to clarify which is S and which is O based on its orders. Therefore, it is really difficult for such set to emerge or remain stable. Set (c) matches the clarifying requirement, but it does not match the economy requirement. Under the drive of economy requirement, the local orders in set (c) may *shift* to those in set (a): SV plus SO triggers SVO, which allows individuals to develop the implied local order VO, and since SV plus VO match both requirements, these orders gradually prevail. During this shift, comprehension is not greatly affected, since the relative order between S and O in either SV plus SO or SV plus VO is consistent. Meanwhile, a reverse shift from set (a) to set (c) is rare, since set (c) does not match the economy requirement. To sum up, if individuals perceive utterances having each of these orders with equal probabilities, and have sufficient chances to be speakers and listeners, once SV wins the competition over VS, there is a higher probability for SV and VO to emerge as the prevalent local orders, and SVO as the prevalent global order. Similarly, if VS wins, VS and OV will become the prevalent local orders, and OVS as the prevalent global order.

Table 2. The possible local and global orders when SV wins the competition over VS, whether they match the two requirements, and whether they can shift to other orders

Possible local orders	Equivalent global orders	Match the clarifying requirement?	Match the economy requirement?	Shift
(a) SV + VO	SVO	Yes	Yes	/
(b) SV + OV	OSV or SOV	No	No	/
(c) SV + SO	SVO or SOV	Yes	No	to (a)
(d) SV + OS	OSV	Yes	Yes	/

SVO and OVS are not evenly distributed in natural languages; there are more languages having SVO than those having OVS (Dryer, 2008). The roughly same prevalence of these two orders in this model is mainly caused by semantic structures and socio-cultural factors. Based on this model, Gong et al. (2009) gave a systematic study on the word order bias shown in this model.

Besides high *UR*, the emergent language also has high *Disp* (0.88±0.13). Figure 10 traces *UR* and *Disp* in Simulations 1 and 2. These results show that the emergent language is *displaced*; it can reliably communicate about events not occurring in the current temporal or spatial environment of conversations (high *UR*, calculated without cues), and even withstand the interference of wrong cues (high *Disp*, calculated under wrong cues).

© 2011. John Benjamins Publishing Company
All rights reserved

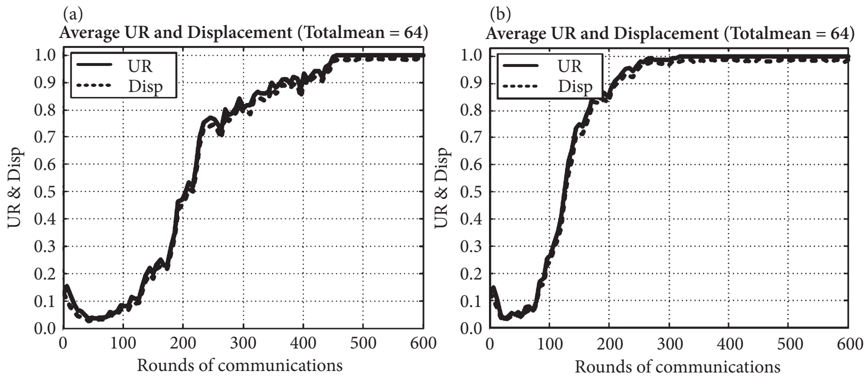

Figure 10. *UR* and *Disp* in Simulations 1 (a) and 2 (b)

(a) Com = 60000:

ComLex =12
Lex 0(Pr) (0.93): ‘10<#, #>’<->/17/
Lex 1(Pr) (0.93): ‘5<#>’<->/2/
Lex 2(Pr) (0.93): ‘8<#, #>’<->/27/
Lex 3(Pr) (0.93): ‘6<#>’<->/26/
Lex 4(Ag|Pat) (0.93): ‘3<#>’<->/8/
Lex 5(Pr) (0.93): ‘4<#>’<->/11/
Lex 6(Ag|Pat) (0.93): ‘1<#>’<->/5/
Lex 7(Ag|Pat) (0.93): ‘2<#>’<->/24/
Lex 8(Ag|Pat) (0.93): ‘0<#>’<->/2/
Lex 9(Pr) (0.93): ‘9<#, #>’<->/3/
Lex 10(Pr) (0.93): ‘11<#, #>’<->/28/
Lex 11(Pr) (0.93): ‘7<#>’<->/12/

Global orders:
SV(16.00 16.00); VS(0.00 0.00);
SVO(48.00 48.00);SOV(7.04 0.00); OSV(0.70 0.00);
VSO(6.10 0.00); VOS(0.70 0.00); OVS(6.40 0.00);

Local orders:
SV(44.88 44.88); VS(0.18 0.00);
VO(44.88 44.88); OV(0.25 0.00);
SO(1.37 0.00); OS(0.00 0.00);

(b) Com = 60000:

ComLex =12
Lex 0(Pr) (0.93): ‘4<#>’<->/17/
Lex 1(Pr) (0.95): ‘10<#, #>’<->/1/
Lex 2(Pr) (0.94): ‘9<#, #>’<->/26/
Lex 3(Ag|Pat) (0.95): ‘1’<->/3/
Lex 4(Ag|Pat) (0.95): ‘2’<->/25/
Lex 5(Pr) (0.94): ‘7<#>’<->/29/
Lex 6(Pr) (0.95): ‘6<#>’<->/23/
Lex 7(Ag|Pat) (0.95): ‘3’<->/7/
Lex 8(Pr) (0.95): ‘5<#>’<->/0/
Lex 9(Ag|Pat) (0.95): ‘0’<->/9/
Lex 10(Pr) (0.95): ‘8<#, #>’<->/28/
Lex 11(Pr) (0.95): ‘11<#, #>’<->/12/

Global orders:
SV(0.00 0.00); VS(16.00 16.00);
SVO(5.50 0.00);SOV(1.70 0.00); OSV(9.60 0.00);
VSO(0.30 0.00); VOS(7.60 0.00); OVS(48.00 48.00);

Local orders:
SV(0.49 0.00); VS(45.60 45.60);
VO(0.13 0.00); OV(45.60 45.60);
SO(0.00 0.00); OS(1.64 0.00);

Figure 11. The common linguistic knowledge emerged after 6000 communications in Simulations 1 (a) and 2 (b). As for lexical knowledge, ComLex records the number of shared lexical rules. In each lexical rule (Lex i), Ag (agent), Pat (patient), and Pr (predicate) are thematic roles, the number within ( ) is the average strength of this rule in all idiolects, integers within ‘ ’ are semantic constituents and those within// are syllables. The homonymous rules are underlined. As for syntactic knowledge, the first number within ( ) indicates the expressivity of that local or global order, and the second number indicates UR_{LocOrd} or UR_{GloOrd} of that order

© 2011. John Benjamins Publishing Company
All rights reserved

Finally, Figure 11 lists the common linguistic knowledge emerged after 6000 communications in Simulations 1 and 2. Some homonymous rules (underlined) are preserved in the communal language in Simulation 1, and the tiny values of UR_{GloOrd} and UR_{LocOrd} of non-prevalent orders indicate a degree of heterogeneity among idiolects. Both of these aspects are consistent with natural languages.

5. General discussions

Compared with the numerous simulation studies on language origin, this model has several important contributions to our understanding of language, language evolution, and simulation studies in this line of research.

5.1 Explicit representation of linguistic knowledge and behaviors

Language evolution involves biological evolution, individual learning, and cultural transmission (Christiansen & Kirby, 2003). Although some previous models using artificial neural networks to store linguistic knowledge (e.g. Batali, 1998; Munroe & Cangelosi, 2002) simulated in detail the process of individual learning, they failed to reveal the close interactions of various linguistic components in language processing, and the structure of the emerged language in those models was limited and sensitive to the adopted neural networks. Meanwhile, other previous models (e.g. the ILM) focused on cultural transmission and revealed some selective roles of cultural transmission in the origin of compositionality. However, the abstract grammatical structures (e.g. Kirby, 1999) and entangled definition of linguistic knowledge (e.g. Smith et al., 2003) in these models were far from what is generally believed to be linguistic knowledge.

Our compositionality-regularity coevolution model explicitly defines linguistic components including lexical items, word orders, and syntactic categories, and clearly traces a simultaneous acquisition of lexical and syntactic knowledge and a gradually-established semantics-syntax correspondence. These results support our coevolution hypothesis on language origin. Such clarification in linguistic knowledge also allows this model to examine other relevant linguistic questions, such as the word order bias (Gong et al., 2009) and linguistic ambiguity due to homonyms (Minett & Gong, 2010). Meanwhile, this model clearly simulates the relevant learning abilities for handling lexical items and word orders during acquisition, production and comprehension. These concrete behaviors help evaluate whether domain-general abilities, once exapted to handle linguistic materials, can trigger linguistic universals.

© 2011. John Benjamins Publishing Company
All rights reserved

5.2 Indirect meaning transfer in comprehension

Many previous models simplified the comprehension process based on direct meaning transfer (Smith, 2005), and even started to simulate comprehension of complex structures based on this assumption (e.g. Kirby, 1999). This assumption was seriously criticized by linguists (e.g. Smith, 2005; Tallerman, 2007), and before how individuals comprehend utterances with simple structures based on linguistic knowledge is understood, it is presumptuous to consider comprehension of complex structures.

In our model, we assume no 'mind reading'; individuals in comprehension refer to their own linguistic knowledge and unreliable nonlinguistic cues. This scenario traces the transition from relying on nonlinguistic information to relying on linguistic information, and reveals the driving force of nonlinguistic information in the early stage of language origin. This transition, unable to be studied based on direct meaning transfer, also paves the way for the emergence of *displacement* (Hockett, 1960) in language. In addition, the reliability of nonlinguistic cues reflects the level of joint attention ability in early hominins. Based on this scenario, we can explore the relation between the level of joint attention and the understandability of the emergent language, and discuss if linguistic understandability could reversely select the level of joint attention in humans. These could lead to an alternative, other than human-unique (Tomasello, 2008), explanation on the level difference of language-related abilities between humans and other primates.

5.3 'Bottom-up' syntactic development

This study illustrates that consistent global orders that regulate multiple items at the sentence level are not necessarily predefined; instead, they can emerge via iteration of local orders between lexical items. Besides the basic constituent word order, more complex orders that involve numerous items could be triggered in a similar way. The iteration of local orders is the underlying mechanism to process sentences with both simple and complex semantic structures. In addition, together with categorization mechanism that associates lexical items and regulating rules on these items, the selective iteration of local orders could naturally lead to semantic or syntactic hierarchy in language. This process could take place in three steps. First, as shown in the current simulations, the non-uniform distributions of semantic constituents trigger an early development of some categories and local orders that regulate members from these categories. Second, once consistent orders of some items are formed, these items could become more coherent and the whole structure involving them could be easily processed. Then, this structure tends to be treated

© 2011. John Benjamins Publishing Company
All rights reserved

as one unit. Finally, when the semantic complexity is increased and other lexical items are added to sentences, new orders could be formed between this coherent unit and other peripheral items. Now, a hierarchical structure emerges, with frequent or coherent items being layered and processed together. A repetition of the last two steps will further increase the complexity in hierarchy. Although the current model, due to predefined semantic space, demonstrates only the first step of this process, the whole process is plausible and requires no predefined concept of linguistic hierarchy.

This 'from local to global' and 'bottom-up' syntactic development senario bridges the gap between domain-general competences and language learning abilities, and between simple non-hierarchical structures and complex grammatical structures. Both aspects question the nativism about the human-uniqueness of language learning abilities and the singularity of language-specific features.

5.4 Semantics driving syntax

The semantic information in this model plays an important role in the emergent syntax, which manifests in three aspects:

1. Guided by the thematic roles of lexical items in integrated meanings, individuals develop categories (V, S or O) to associate lexical items and syntactic rules;
2. The relative frequencies of semantic constituents in integrated meanings and the necessity to clarify thematic roles (e.g. S and O) trigger bias towards certain global orders (e.g. SVO and OVS);
3. The emergent syntactic structures reflect the semantic structures. In the emergent language with good understandability, lexical rules having the same thematic roles are usually associated with categories having the corresponding syntactic roles, and expressions encoding meanings with the same semantic structure usually have the same syntactic structure (i.e. following consistent orders).

These aspects are commensurate with *the semantics driving syntax hypothesis* in cognitive grammar (Langacker, 1990; Schoenemann, 1999). This hypothesis states that universal features of grammar in language appear to be essential descriptions of aspects of our basic conceptual universe and changes in the conceptual complexity drive the evolution of grammar.

5.5 Effects of socio-cultural factors on language evolution

Language origin takes place in a socio-cultural environment (Feldman, 2006). As discussed before, the random, pair-wise communications in this model can

© 2011. John Benjamins Publishing Company
All rights reserved

affect the emergent prevalent global order. Apart from such *horizontal transmission* (communications among members of the same generation, Cavalli-Sforza & Feldman, 1981), *vertical* and *oblique transmissions* (communications among members of different generations, *ibid.*) could also cast their influences on language evolution. Based on this model, the roles of these major forms of *cultural transmission* (Christiansen & Kirby, 2003) in language evolution can be examined (Gong et al., 2010; Gong, in press). Similarly, social factors, by restricting participants in communications, could also affect language evolution, and there could be mutual influence between language and social structures (Gong et al., 2008). All these suggest that a systematic study of language evolution should synthesize not only the language processing mechanisms in communications but also the socio-cultural factors that could affect communicative patterns.

5.6 Language as a Complex Adaptive System (CAS)

A CAS usually features a number of interacting components, whose aggregate activity is nonlinear and exhibits self-organization (Steels, 2000; Lee, 2003). One type of nonlinearity is *phase transition*; the aggregate activity may change abruptly or dramatically due to small changes in some parameters or initial conditions (Bonabeau et al., 1999). *Self-organization* is "a process in which pattern at the global level of a system emerges from numerous interactions among the lower-level components of the system" (Camazine et al., 2001, p.8). In a CAS, macroscopic outcomes usually emerge from microscopic interactions among various components, and the global organizational properties are invisible at the local level (Ball, 2001; Oudeyer, 2006).

As discussed in the previous sections, language features a number of individuals that communicate with each other and a number of components (e.g. lexicon, syntax and categories) that interact with each other in language processing. In communications, self-organization leads to a *circularity* of signal usage and learning: one individual affects another when producing utterances that encode meanings, and is affected by others when comprehending utterances created by others; meanwhile, comprehension induces an adjustment in an individual's linguistic knowledge, thus also affecting production. In this model, such circularity triggers the interactions: (1) *between lexical items and ordering relations*, thus leading to a coevolution of compositionality and word order regularity; (2) *between semantic and emergent syntactic structures*, thus leading to the semantics-syntax correspondence; and (3) *between emergent language and its cultural environment*, thus leading to some bias in distributions of global orders. All these contribute to the 'sudden' (phase transition) origin of lexical and syntactic

© 2011. John Benjamins Publishing Company
All rights reserved

knowledge in the communal language, as shown by the S-shape sharp increase in *UR* in Figure 8 and Figure 9.

5.7 Limitations

A comprehensive evaluation of a computer model needs to consider both its advantages and its limitations. The current model has some limitations. First, there is *a default connection between semantics and syntax*: lexical items encoding semantic constituents as agents are always mapped to S, and those as patients mapped to O. However, in natural languages, the mapping between semantics and syntax is not always like this. For example, passive voice maps agents to O and patients to S. Note that passive voice in natural languages is rarely marked solely by word orders. In order to incorporate passive voice, we have to consider other relevant language processing mechanisms.

Second, the language processing mechanisms in this model are *offline*. For example, in comprehension, after the whole utterance is received, the listener starts to interpret it into an integrated meaning. Psycholinguists have suggested that semantic information and syntactic information are processed in parallel and immediately used once be available (MacDonald et al., 1994; Grodner & Gibson, 2005). Therefore, the offline mechanisms in this model and many others (e.g. Kirby, 1999; Steels, 2004) are not that realistic. Some online processing mechansims revealed from the psychological experiments of comprehension of sentences with temporary syntactic ambiguity (e.g. van Gompel et al., 2006) could be borrowed to modify the current model.

Finally, this model faces some *scale-up problems* (Cangelosi & Parisi, 2002); it considers only simple predicate-argument structures and mechanisms to process lexical items and simple word orders. Natural languages are more than lexicon plus word orders, and can express more complex meanings. This model alone fails to explain the complete picture of language evolution. However, to investigate a CAS as language, it is wise to separate the whole system into smaller subsets, disentangle complex processes into relatively simple sub-processes, and divide mixed results into separate ones dependent on specific conditions. The convincing conclusions obtained from these simplified cases pave the way for future studies with increased complexity, and will eventually lead to the comprehensive understanding of the target question. Considering this, the current model is still useful. By combing this model with others that address the acquisition of semantic concepts (e.g. Puglisi et al., 2008; Vogt, 2005) and the evolution of other grammatical aspects, such as morphology (e.g. Steels, 2004), it is promising to reveal the whole picture of language evolution towards its full complexity.

© 2011. John Benjamins Publishing Company
All rights reserved

6. Conclusions

This paper proposes that language origin is a process of coevolution of compositionality and word order regularity, and that syntax in language is developed in a constructivist manner. A multi-agent model is designed to illustrate this hypothesis. The simulation results show that given some general learning abilities, such as pattern extraction and sequential learning, a communal language showing a certain degree of systematicity can emerge in a population of individuals. This coevolutionary hypothesis inspires us to take an integrated view on language and note that language origin can be due to various factors. We certainly do not attempt to claim that coevolution is the definitive answer to the question of syntax evolution; instead, our purpose is to provide an exploratory discussion of possible scenarios of language origin, and to propose simpler, but still useful, forms of behavior that could bridge the gap between no language and the languages we use today. As shown in our study, coevolution may partially explain the origin of some linguistic universals and the singularity of language.

Acknowledgement

This work is supported by the Alexander von Humboldt Foundation in Germany and the Society of Scholars in the Humanities in the University of Hong Kong. The author thanks Prof. Bernard Comrie from the Max Planck Institute for Evolutionary Anthropology and Prof. William S-Y. Wang from the Chinese University of Hong Kong for the useful comments. The paper is also benefited a lot from the comments of anonymous reviewers.

Notes

1. Two such examples are given by Arbib (2005): */grooflook/* means "Take your spear and go around the other side of that animal and we will have a better chance together of being able to kill it" and */koomzask/* means "The alpha male has killed a meat animal and now the tribe has a chance to feast together. Yum, yum!"

2. Theoretically speaking, this model can handle reflexive meanings. Once lexical items are associated to syntactic categories, the categorical knowledge can help clarify which is S and which is O in reflexive meanings, just as it did to clarify *virtual homonyms* (whose meanings possess different thematic roles) in Minett & Gong (2010). Considering that additional

© 2011. John Benjamins Publishing Company
All rights reserved

mechanisms are needed to handle such clarification, I simply exclude reflexive meanings in this model, just as many other simulation studies did (e.g. the ILM).

3. In standard linguistic practice, lexical associations are not called "rules", but productive grammatical realizations are. However, a standard rule-based system treats any kinds of knowledge as rules, and in models that adopt rule-based systems to encode linguistic knowledge, both lexical and syntactic knowledge are usually defined as rules. In this model, we stick to calling lexical associations "lexical rules".

4. Some semanticists (e.g. Lyons, 1977) deny the existence of synonyms at the lexical level, but suggest that most synonyms are context-dependent. In our simulations, only some kind of homonyms can be preserved in the communal language, whereas synonyms cannot.

5. The model can also simulate ergative languages by associating lexical items encoding constituents as agent of intransitive meanings and those as patient of transitive meanings into a syntactic category. The possible emergent word orders in this situation were discussed in Gong et al. (2009).

6. As shown in Gong (2009), the slight change in some parameters (e.g. semantic space size, utterance space size, buffer size, and rule list sizes) cannot greatly affect the simulation results. Some parameters (e.g. creation rate, extraction rate, and amount of rule adjustment in competition) must have non-zero values to trigger a communal language with good understandability. Some (e.g. reliability of cue) must exceed certain threshold values to trigger a communal language with good understandability. The values of some parameters (e.g. population size, number of communications, and number of utterance exchange per communication) are correlated; if one of them is increased or decreased, the others must also be increased or decreased to achieve a communal language with good understandability.

7. The model simulates the evolution of lexical and syntactic knowledge, not that of semantic constituents. We assume that individuals acquire the semantic constituents from the initial holistic rules. In fact, simulations starting from no initial holistic rules show similar results.

References

Akhtar, N., & Tomasello, M. (1997). Young children's productivity with word order and verb morphology. *Developmental Psychology*, 33(6): 952–965.

Aldridge, J.W., & Berridge, K.C. (2003). Basal ganglia neural coding of natural action sequences. In A.M. Graybiel, S.T. Kitai, & M. Delong (Eds.), *The basal ganglia VI series: Advances in behavioral biology* (pp. 279–287). New York: Plenum.

Anderson, J.R. (1976). *Language, memory and thought*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Arbib, M.A. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences*, 28(2): 105–124.

Arbib, M.A. (2008). Holophrasis and the protolanguage spectrum. *Interaction Studies*, 9(1): 154–168.

© 2011. John Benjamins Publishing Company
All rights reserved

Arbib, M.A. & Bickerton, D. (Eds.) (2008). *Holophrasis vs. compositionality in the emergence of protolanguage*. Special issue of *Interaction Studies*, 9(1).
Ball, P. (2001). *The self-made tapestry: Pattern formation in nature*. New York: Oxford University Press.
Barnes, J. (Ed.) (1984). *The complete works of Aristotle: The revised oxford translation*. Princeton, NJ: Princeton University Press.
Baronchelli, A., Felici, M., Loreto, V., Caglioti, E., & Steels, L. (2006). Sharp transition towards shared vocabularies in multi-agent systems. *Journal of Statistical Mechanics: Theory and Experiment*, P06014. Doi: 10.1088/1742-5468/2006/06/P06014.
Batali, J. (1998). Computational simulations of the emergence of grammar. In J.R. Hurford, M. Studdert-Kennedy, & C. Knight (Eds.), *Approaches to the evolution of language: Social and cognitive bases* (pp. 405–426). Cambridge: Cambridge University Press.
Bates, E., Devescovi, A., Hernandez, A., & Pizzamiglio, L. (1996). Gender priming in Italian. *Perception and Psychophysics*, 58(7): 992–1004.
Bates, E., & Goodman, J.C. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real-time processing. *Language and Cognitive Processes*, 12(5–6): 507–584.
Bickerton, D. (1990). *Language and species*. Chicago, IL: The University of Chicago Press.
Bickerton, D. (2008). But how did protolanguage actually *start*? *Interaction Studies*, 9(1): 169–176.
Bonabeau, E., Dorigo, M., & Theraulaz, G. (1999). *Swarm intelligence: From natural to artificial systems*. New York: Oxford University Press.
Braitenberg, V., Heck, D., & Sultan, F. (1997). The detection and generation of sequences as a key to cerebellar function: Experiments and theory. *Behavioral and Brain Sciences*, 20(2): 229–245.
Briscoe, E.J. (1997). Coevolution of language and of the language acquisition device. In *Proceedings of the* 8th *Conference on the European Chapter of the Association for Computational Linguistics*, 418–427.
Brooks, P.J., & Tomasello, M. (1999). How children constrain their argument structure constructions. *Language*, 75(4): 720–738.
Bybee, J.L. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, 10(5): 425–455.
Camazine, S. Deneubourg, J-L., Franks, N.R., Sneyd, J., Theraulaz, G., & Bonabeau, E. (2001). *Self-organization in biological systems*. Princeton, NJ: Princeton University Press.
Cangelosi, A., & Parisi, D. (2002). Computer simulation: A new scientific approach to the study of language evolution. In A. Cangelosi, & D. Parisi (Eds.), *Simulating the evolution of language* (pp. 3–28). London: Springer-Verlag.
Cavalli-Sforza, L.L., & Feldman, M. (1981). *Cultural transmission and evolution: A quantitative approach*. Princeton, NJ: Princeton University Press.
Cheney, D.L., & Seyfarth, R.M. (2005). Constraints and preadaptations in the earliest stages of language evolution. *The Linguistic Review*, 22(2–4): 135–160.
Cheney, D.L., & Seyfarth, R.M. (2007). *Baboon metaphysics: The evolution of a social mind*. Chicago, IL: University of Chicago Press.
Chomsky, N. (1972). *Language and mind*. New York: Harcourt Brace Jovanovich.
Christiansen, M.H., & Kirby, S. (2003). Language evolution: Consensus and controversies. *Trends in Cognitive Sciences*, 7(7): 300–307.
Christiansen, M.H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, 31(5): 489–509.

© 2011. John Benjamins Publishing Company
All rights reserved

Clark, E.V. (1987). The principle of contrast: a constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 1–33). Hillsdale: Lawrence Erlbaum Associates.

Clark, E.V. (2003). *First language acquisition*. Cambridge: Cambridge University Press.

Croft, W. (2001). *Radical construction grammar: Syntactic theory in typological perspective*. New York: Oxford University Press.

Cutting, J.C., & Bock, K. (1997). That's the way the cookie bounces: Syntactic and semantic components of experimentally elicited idiom blends. *Memory and Cognition*, 25(1): 57–71.

Damasio, A.R., & Damasio, H. (1992). Brain and language. *Scientific American*, 267(3): 89–95.

de Pauw, G. (2006). Towards to a fixed word order in a society of agents: A data-driven baseline perspective. In A. Cangelosi, A.D.M. Smith, & K. Smith (Eds.), *The evolution of language: Proceedings of the* 6th *international conference* (pp. 43–50). Singapore: World Scientific Press.

Digweed, S.M., Fedigan, L.M., & Rendall, D. (2005). Variable specificity in the anti-predator vocalizations and behaviour of the white-faced capuchin (*Cebus capucinus*). *Behavior*, 142(8): 997–1021.

Dryer, M.S. (2008). Order of subject, object, and verb. In: M. Haspelmath, M.S. Dryer, D. Gil, & B. Comrie (Eds.), *The world atlas of language structures online* (Chapter 81). Munich: Max Planck Digital Library. Available at http://wals.info/feature/81. Accessed on July 14, 2008.

Ehrlich, P.R., & Raven, P.H. (1964). Butterflies and plants: a study of coevolution. *Evolution*, 18(4): 586–608.

Elman, J. (2005). Connectionist models of cognitive development: Where next? *Trends in Cognitive Sciences*, 9(3): 111–117.

Elman, J., Bates, E., Johnson, M.H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.

Feldman, J.A. (2006). From molecule to metaphor: A neural theory of language. Cambridge, MA: MIT Press.

Fillmore, L.W. (1979). Individual differences in second language acquisition. In C.J. Fillmore, D. Kempler, & W.S-Y. Wang (Eds.), *Individual differences in language ability and language behavior* (pp. 203–228). New York: Academic Press.

Fisher, S.E., & Marcus, G.F. (2005). The eloquent ape: Genes, brains and the evolution language. *Nature Reviews Genetics*, 7: 9–19.

Fitch, W.T. (2005). The evolution of language: A comparative review. *Biology and Philosophy*, 20(2–3): 193–230.

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

Fitch, W.T., Hauser, M.D., & Chomsky, N. (2005). The evolution of the language faculty: Clarifications and implications. *Cognition*, 97(2): 179–210.

Friederici, A.D., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing: Evidence challenging the critical period hypothesis. *Proceedings of the National Academy of Sciences of United States of America*, 99(1): 529–534.

Friederici, A.D., Bahlmann, J., Heim, S., Scubotz, R.I., & Anwander, A. (2006). The brain differentiates human and non-human grammar: Functional localization and structural connectivity. *Proceedings of the National Academy of Sciences of United States of America*, 103(7): 2458–2463.

Giurfa, M., Zhang, S., Jenett, A., Menzel, R., & Srinivasan, M.V. (2001). The concepts of 'sameness' and 'difference' in an insect. *Nature*, 410: 930–932.

© 2011. John Benjamins Publishing Company
All rights reserved

Goldberg, A.E. (1995). Constructions: A construction grammar approach to argument structure. Chicago, IL: Chicago University Press.

Gong, T. (2009). Computational simulation in evolutionary linguistics: A study on language emergence. Taipei: Academia Sinica.

Gong, T., Minett, J.W., & Wang, W.S-Y. (2008). Exploring social structure effect on language evolution based on a computational model. *Connection Science*, 20(2–3): 135–153.

Gong, T., Minett, J.W., & Wang, W.S-Y. (2009). A simulation study on word order bias. *Interaction Studies*, 10(1): 51–76.

Gong, T., Minett, J.W., & Wang, W.S-Y. (2010). A simulation study exploring the role of cultural transmission in language evolution. *Connection Science*, 22(1): 69–85.

Gong, T. (in press). Exploring the roles of horizontal, vertical, and oblique transmissions in language evolution. *Adaptive Behavior*.

Griffiths, T.L., & Kalish, M.L. (2007). Language evolution by iterated learning with Bayesian agents. *Cognitive Science*, 31: 441–480.

Grodner, D., & Gibson, E. (2005). Consequences of the serial nature of linguistic input. *Cognitive Science*, 29(2): 261–291.

Hauser, M.D., Newport, E.L., & Aslin, R.A. (2001). Segmentation of the speech stream in a non-human primate: Statistical learning in cotton-top tamarins. *Cognition*, 78(3): B53–B64.

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: 1569–1579.

Hauser, M.D., Barner, D., & O'Donnell, T. (2007). Evolutionary linguistics: A new look at an old landscape. *Language Learning and Development*, 3(2): 101–132.

Herrmann, E., Call, J., Hernández-Lloreda, M.V., Hare, B., & Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science*, 317: 1360–1366.

Hockett, C.F. (1960). The origin of speech. *Scientific American*, 203: 88–96.

Hoy, R.R., & Paul, R.C. (1973). Genetic coding of song specificity in crickets. *Science*, 180: 82–83.

Hurford, J.R. (1989). Biological evolution of the Saussurean sign as a component of the language acquisition device. *Lingua*, 77(2): 187–222.

Hurford, J.R. (2007). *The origins of meaning*. New York: Oxford University Press.

Hurford, J.R. (2009). Diversity in language, genes and the language faculty. In R. Botha, & C. Knight (Eds.), *The Cradle of Language* (pp. 163–184). Oxford: Oxford University Press.

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

Jackendoff, R., & Pinker, S. (2005). The nature of the language faculty and its implications for evolution of language (Reply to Fitch, Hauser, and Chomsky). *Cognition*, 97(2): 211–225.

Jacob, F. (1977). Evolution and tinkering. *Science*, 196: 1161–1166.

Karakashian, S.J., Gyger, M., & Marler, P. (1988). Audience effects on alarm calling in chickens (*Gallus gallus*). *Journal of Comparative Psychology*, 102(2): 129–135.

Ke, J-Y., Minett, J.W., Au, C-P., & Wang, W.S-Y. (2002). Self-organization and selection in the emergence of vocabulary. *Complexity*, 7(3): 41–54.

Ke, J-Y., & Holland, J.H. (2006). Language origin from an emergentist perspective. *Applied Linguistics*, 27(4): 691–716.

Kirby, S. (1999). *Function, selection and innateness: The emergence of language universals*. Oxford: Oxford University Press.

© 2011. John Benjamins Publishing Company
All rights reserved

Kirby, S. (2007). The evolution of language. In R. Dunbar, & L. Barrett (Eds.), *Oxford handbook of evolutionary psychology* (pp. 669–681). Oxford: Oxford University Press.
Kirby, S., Dowman, M., & Griffiths, T.L. (2007). Innateness and culture in the evolution of language. *Proceedings of the National Academy of Sciences of United States of America*, 104(12): 5241–5245.
Krifka, M. (2001). Compositionality. In R.A. Wilson, & F.C. Keil (Eds.), *The MIT encyclopedia of the cognitive sciences* (pp. 152–157). Cambridge, MA: MIT Press.
Kuczaj II, S.A., & Hendry, J.L. (2003). Does language help animals think? In D. Gentner, & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 237–275). Cambridge, MA: MIT Press.
Langacker, R.W. (1990). *Foundations of cognitive grammar: Theoretical prerequisites*. Stanford, CA: Stanford University Press.
Larson, R.K., Déprez, V., & Yamakido, H. (Eds.) (2010). The evolution of human language: Biolinguistic perspectives. Cambridge: Cambridge University Press.
Lee, N. (2003). *Emergence of language as Complex Adaptive Systems*. Doctoral Dissertation, Los Angeles, CA: University of California.
Lieberman, P. (2006). *Toward an evolutionary biology of language*. Cambridge, MA: Belknap Press of Harvard University Press.
Lyon, C., Nehaniv, C.L., & Cangelosi, A. (Eds.) (2007). *Emergence of communication and language*. London: Springer-Verlag.
Lyons, J. (1977). *Semantics* (Vol. I & II). Cambridge: Cambridge University Press.
MacDonald, M.C., Pearlmutter, N.J., & Seidenberg, M.S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101: 676–703.
MacWhinney, B. (Ed.) (1999). *The emergence of language*. Mahwah, NJ: Lawrence Erlbaum Associates.
Maess, B., Koelsch, S., Gunter, T., & Friederici, A.D. (2001). Musical syntax is processed in Broca's area: an MEG study. *Nature Neuroscience*, 4: 540–545.
Marchman, V., & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language*, 21(2): 339–366.
Marcotte, J-P. (2006). Causative alternation errors and innate knowledge: Consequences of the "no negative evidence" fallacy. In E.V. Clark, & B.F. Kelly (Eds.), *Constructions in acquisition* (pp. 205–232). Stanford, CA: CSLI Publications.
Marcus, G.F., Vijayan, S., Bandi, R.S., & Vishton, P.M. (1999). Rule learning by seven-month-old infants. *Science*, 283: 77–80.
Markman, E.M. (1989). *Categorization and naming in children: Problems of induction*. Cambridge, MA: MIT Press.
Mellow, J.D. (2008a). The emergence of complex syntax: A longitudinal case study of the ESL development of dependency resolution. *Lingua*, 118(4): 499–521.
Mellow, J.D. (2008b). How big is minimal? *Lingua*, 118(4): 632–635.
Minett, J.W., & Gong, T. (2010). Exploring linguistic ambiguity from a simulation perspective. In W-Y. Pan, & Z-W. Shen (Eds.), *Festschrift for the* 75th *birthday of William S-Y. Wang* (pp., 112–131). Shanghai: Shanghai Education Press.
Monchi, O., Petrides, P., Petre, V., Worsley, K., & Dagher, A. (2001). Wisconsin card sorting revisited: Distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. *Journal of Neuroscience*, 21(19): 7733–7741.

© 2011. John Benjamins Publishing Company
All rights reserved

Mufwene, S.S. (2008). What do creoles and pidgins tell us about the evolution of languages? In B. Laks (Ed.), *Origin and evolution of languages: Approaches, models, paradigms* (pp. 272–297). London: Equinox.

Munroe, S., & Cangelosi, A. (2002). Learning and the evolution of language: The role of cultural variation and learning cost in Baldwin Effect. *Artificial Life*, 8(4): 311–339.

Munn, A. (1999). First conjunct agreement: Against a clausal analysis. *Linguistic Inquiry*, 30(4): 643–668.

Nowak, M.A., Plotkin, J.B., & Krakauer, D. (1999). The evolutionary language game. *Journal of Theoretical Biology*, 21(2): 147–162.

O'Grady, W.D. (2008). The emergentists program. *Lingua*, 118(4): 447–464.

Okanoya, K., & Merker, B. (2007). Neural substrates for string-context mutual segmentation: A path to human language. In C. Lyon, C.L. Nehaniv, & A. Cangelosi (Eds.), *Emergence of communication and language* (pp. 421–434). London: Springer-Verlag.

Oudeyer, P–Y. (2006). *Self-organization in the evolution of speech* (translated by J.R. Hurford). Oxford: Oxford University Press.

Quine, W.V. (1960). *Word and object*. Cambridge, MA: MIT Press.

Patel, A.D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P.J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience*, 10(6): 717–733.

Petersson, K.M., Forkstam, C., & Ingvar, M. (2004). Artificial syntactic violations activate Broca's region. *Cognitive Science*, 28(3): 383–407.

Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences*, 13(4): 707–784.

Puglisi, A., Baronchelli, A., & Loreto, V. (2008). Cultural route to the emergence of linguistic categories. *Proceedings of the National Academy of Sciences of United States of America*, 105(23): 7936–7940.

Pulvermüller, F., & Knoblauch, A. (2009). Discrete combinatorial circuits emerging in neural networks: A mechanism for rules of grammar in the human brain? *Neural Networks*, 22(2): 161–172.

Reali, F., & Christansen, M.H. (2009). Sequential learning and the interaction between biological and linguistic adaptation in language evolution. *Interaction Studies*, 10(1): 5–30.

Rumelhart, D.E., & McClelland, J.L. (1987). Learning the past tenses of English verbs: Implicit rules or parallel distributed processing. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 195–248). Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Schoenemann, P.T. (1999). Syntax as an emergent characteristic of the evolution of semantic complexity. *Minds and Machines*, 9(3): 309–346.

Shima, K., Isoda, M., Mushiake, H., & Tanji, J. (2007). Categorization of behavioral sequences in the prefrontal cortex. *Nature*, 445: 315–318.

Smith, A.D.M. (2005). Mutual exclusivity: Communicative success despite conceptual divergence. In M. Tallerman (Ed.), *Language origins: Perspectives on evolution* (pp. 372–388). Oxford: Oxford University Press.

Smith, A.D.M. (2008). Protolanguage reconstructed. *Interaction Studies*, 9(1): 98–114.

Smith, K., Kirby, S., & Brighton, H. (2003). Iterative learning: A framework for the emergence of language. *Artificial Life*, 9(4): 371–386.

© 2011. John Benjamins Publishing Company
All rights reserved

Steele, S. (1978). Word order variation: A typological study. In J.H. Greenberg (Ed.), *Universals of human language* (pp. 585–623). Stanford, CA: Stanford University Press.

Steels, L. (2000). Language as a Complex Adaptive System. In M. Schoenauer, K. Deb, G. Rudolph, X. Yao, E. Lutton, J.J.M. Guervós, & H–P. Schwefel (Eds.), *Proceedings of the* 6th *international conference on parallel problem solving from nature* (pp. 17–26). Berlin: Springer-Verlag.

Steels, L. (2004). Constructivist development of grounded construction grammar. In D. Scott, W. Daelemans, & M. Walker (Eds.), *Proceedings of* 45th *annual Meeting of the Association for Computational Linguistics* (pp. 9–16).

Steels, L., van Trijp, R., & Wellens, R. (2007). Multi-level selection in the emergence of language systematicity. In F. Almeida e Costa, L.M. Rocha, E. Costa, & I. Harvey (Eds.). *Proceedings of* 9th *European conference on artificial life* (pp. 425–434). Berlin: Springer-Verlag.

Tallerman, M. (2007). Did our ancestors speak a holistic protolanguage? *Lingua*, 117(3): 579–604.

Tallerman, M. (2008). Holophrastic protolanguage: Planning, processing, storage, and retrieval. *Interaction Studies*, 9(1): 84–99.

Terrace, H.S. (2002). Serial expertise and the evolution of language. In A. Wray (Ed.), *The transition to language* (pp. 64–90). Oxford: Oxford University Press.

Thompson, J.N. (1994). *The coevolutionary process*. Chicago, IL: University of Chicago Press.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Tomasello, M. (2008). *Origins of human communication*. Cambridge, MA: MIT Press.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28(5): 675–691.

Tomlin, R.S. (1986). *Basic word order: Functional principles*. London: Croom Helm.

van der Hulst, H. (Ed.) (2010). *Recursion and human language*. New York: De Gruyter Mouton.

van der Velde, F., & Kamps, M. (2006). Neural blackboard architectures of combinatorial structures in cognition. *Behavioral and Brain Sciences*, 29(1): 37–108.

van Gompel, R.P.G., Pickering, M.J., Pearson, J., & Jacob, G. (2006). The activation of inappropriate analyses in garden-path sentences: Evidence from structural priming. *Journal of Memory and Language*, 55(3): 335–362.

Vogt, P. (2005). On the acquisition and evolution of compositional languages: sparse input and productive creativity of children. *Adaptive Behavior*, 13(4): 325–346.

Vokey, J.R., & Higham, P.A. (2005). Abstract analogies and positive transfer in artificial grammar learning. *Canadian Journal of Experimental Psychology*, 59(1): 54–61.

Volterra, V., Caselli, M.C., Capirci, O., & Pizzuto, E. (2005). Gesture and the emergence and development of language. In M. Tomasello, & D.I. Slobin (Eds.), *Beyond nature-nurture: Essays in honor of elizabeth bates* (pp. 1–40). Mahwah, NJ: Lawrence Erlbaum Associates.

Wagner, K., Reggia, J.A., Uriagereka, J., & Wilkinson, G.S. (2003). Progress in the simulation of emergent communication and language. *Adaptive Behavior*, 11(1): 37–69.

W.S-Y. Wang (1982). Explorations in language evolution. In W.S–Y. Wang (Ed.), *Explorations in language* (pp. 105–131). Seattle, WA: Pyramid Press.

W.S-Y. Wang (1991). The three scales of diachrony. In W.S–Y. Wang (Ed.), *Explorations in language* (pp. 60–71). Seattle, WA: Pyramid Press.

Wray, A. (2002). *Formulaic language and the lexicon*. New York: Cambridge University Press.

Wray, A. (2007). 'Need only' analysis in linguistic ontogeny and phylogeny. In C. Lyon, C.L. Nehaniv, & A. Cangelosi (Eds.), *Emergence of communication and language* (pp. 53–70). New York: Springer-Verlag.

© 2011. John Benjamins Publishing Company
All rights reserved

Author's address

Tao Gong
Department of Linguistics, University of Hong Kong
Pokfulam Road, Hong Kong

email: gtojty@gmail.com

Author biography

Tao Gong received his Ph.D. from the Chinese University of Hong Kong in 2007. After two years study as an Humboldt Researcher in the Department of Linguistics, Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, he is now a Research Scholar in the Society of Scholars at the School of Humanities in the University of Hong Kong. By means of computational modeling, he carries out research on evolutionary linguistics, such as semantics-syntax coevolution, categorization, and bottom-up syntactic development.

© 2011. John Benjamins Publishing Company
All rights reserved