

Adaptive Behavior

<http://adb.sagepub.com/>

Exploring the Roles of Horizontal, Vertical, and Oblique Transmissions in Language Evolution

Tao Gong

Adaptive Behavior 2010 18: 356

DOI: 10.1177/1059712310377241

The online version of this article can be found at:

<http://adb.sagepub.com/content/18/3-4/356>

Published by:



<http://www.sagepublications.com>

On behalf of:

ISAB

International Society of Adaptive Behavior

Additional services and information for *Adaptive Behavior* can be found at:

Email Alerts: <http://adb.sagepub.com/cgi/alerts>

Subscriptions: <http://adb.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations: <http://adb.sagepub.com/content/18/3-4/356.refs.html>

Exploring the Roles of Horizontal, Vertical, and Oblique Transmissions in Language Evolution

Tao Gong

Department of Linguistics, Max Planck Institute for Evolutionary Anthropology

This article proposes an acquisition framework that involves horizontal, vertical, and oblique transmissions. Based on a lexicon–syntax coevolution model, it discusses the relative roles of these forms of cultural transmission on language origin and change. The simulation results not only reveal an integrated role of oblique transmission that combines the roles of horizontal and vertical transmissions in preserving linguistic understandability within and across generations of individuals, but also show that both horizontal and oblique transmissions are more necessary than vertical transmission for language evolution in a multiagent cultural environment.

Keywords cultural transmission · language evolution · computational simulation

1 Introduction

Cultural transmission is the process by which information is passed from individual to individual via social learning mechanisms such as imitation, teaching, or language (Mesoudi & Whiten, 2008). In terms of linguistic interactions, cultural transmission can be viewed as the process of language adaptation in a community via various kinds of communication among individuals of the same or different generations (Christiansen & Kirby, 2003). Generally speaking, there are three major forms of cultural transmission (Cavalli-Sforza & Feldman, 1981): (1) *horizontal (H) transmission*, communications among individuals of the same generation; (2) *vertical (V) transmission*, a member of one generation talking to a biologically related member of a later generation; and (3) *oblique (O) transmission*, a member of one generation talking to a biologically unrelated member of a later generation. There have been many empirical studies from sociolinguistics and historical linguistics that discuss the influ-

ence of these forms of transmission on linguistic variation and change (e.g., Boyd & Richerson, 1985; Labov, 2001; Mufwene, 2001). These studies have largely extended the one-speaker-one-listener framework of language evolution (Chomsky, 1972).

Apart from the empirical approach, the simulation approach that incorporates various forms of transmission has recently joined the discussion on the roles of cultural transmission in language evolution. For example, H transmission was first considered in the study of social conventions (Lewis, 1969) and cultural dissemination (Axelrod, 1997), and adopted in some mathematical models on the evolution of signaling systems (e.g., Lenaerts, Jansen, Tuyls, & De Vylder, 2005; Skyrms, 2009). Many language models that explore the lexical or grammatical evolution (e.g., Baronchelli, Felici, Loreto, Caglioti, & Steels, 2006; Ke, Minett, Au, & Wang, 2002; Steels, 2005) have also adopted H transmission as a default way of cultural transmission. V transmission was first adopted in the *Iterated Learning Model* (ILM; Kirby, 1998, 2000; K. Smith, Brighton,

Correspondence to: Tao Gong, Department of Linguistics, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany. *E-mail:* gtojt@gmail.com. *Tel.:* +49 341 3550319, *Fax:* +49 341 3550333

Copyright © 2010 International Society for Adaptive Behavior (2010), Vol 18(3–4): 356–376.

DOI: 10.1177/1059712310377241

Figures 5–7, A.1–A.3, A.5 appear in color online: <http://adb.sagepub.com>

& Kirby, 2003) and accepted as a fundamental way of transmission across generations of individuals in laboratory experiments (e.g., Cornish, 2010; Mesoudi & Whiten, 2008) and Bayesian learning models (e.g., Griffiths & Kalish, 2007; Nowak & Komarova, 2001; Nowak, Plotkin, & Krakauer, 1999; K. Smith & Kirby, 2008). Finally, O transmission was partially involved, together with H or V transmission, in some studies (e.g., Lenaerts et al., 2005; K. Smith & Hurford, 2003), and the effect of O transmission on language evolution has not been systematically studied.

In order to simulate a realistic cultural environment, some recent studies (e.g., Acerbi & Parisi, 2006; Vogt, 2005) began to incorporate more than one form of transmission or combine cultural transmission with genetic transmission (e.g., Lenaerts et al., 2005). For example, Vogt's (2005) guessing-game model defined two parameters to manipulate the probabilities of choosing speakers and listeners from the adult generation. In this way, not only H, V, and O transmissions but also the child-talking-to-adult transmission were involved, but the latter form could contaminate the effects of the other three (Gong, Minett, & Wang, 2010). Noticing this, Gong et al. (2010) proposed an acquisition framework that included *intra-generational* (H transmission among children) and *inter-generational* transmissions (V and O transmissions from adults to children). In this framework, V and O transmissions were entangled (each adult had an equal chance to talk to any child), which made it unable to examine the respective roles of these forms in language evolution. In addition, both adults and children had the same level of language learning abilities. Considering the *critical period hypothesis*, there seems to be a stage in the maturation of a human being in which language acquisition is possible in a natural fashion, before or after which this process is much more difficult (Penfield & Roberts, 1959; Lennberg, 1967). Despite of the debate on the extent to which language acquisition abilities are linked to age (e.g., Friederici, Steinhauer, & Pfeifer, 2002; Pinker, 1994; Singleton & Lengyel, 1995), it is generally more difficult for an adult, compared with a young child, to acquire a language. A realistic acquisition framework should note this difference between adults and children.

In this article, we modify the acquisition framework in Gong et al. (2010) to examine the roles of H, V, and O transmissions in language evolution. Three

parameters are defined to control the probabilities of these forms of transmission in language acquisition respectively. During these transmissions, only children actively acquire linguistic knowledge. Based on a lexicon–syntax coevolution model, we conduct a series of simulations under different combinations of these forms of transmission to reveal their respective roles in language evolution, and illustrate under what cultural settings a communal language can be triggered efficiently and maintained largely across generations of individuals.

The rest of the article is organized as follows: Section 2 introduces the lexicon–syntax coevolution model and the acquisition framework; Section 3 describes the simulation setup and analyzes the results, whose robustness is further evaluated in the Appendix; and finally, Section 4 gives a general discussion on these results and concludes the article.

2 The Lexicon–Syntax Coevolution Model and the Acquisition Framework

This model aims to show whether a population of interacting *individuals* (artificial agents) can, based on some general learning mechanisms, develop a compositional language out of a holistic signaling system. These learning mechanisms include the ability of pattern extraction, which helps individuals to extract recurrent patterns from exchanged utterances into lexical items, and the ability of sequential learning, which helps individuals to develop knowledge concerning the order relations of lexical items in exchanged utterances and associate lexical items into syntactic categories. Following local orders among categories, individuals can regulate lexical items to encode expressions having a simple predicate–argument structure. The evolved language, consisting of a set of common lexical items and consistent word order(s), shows a certain degree of systematicity.

This model improves the previous ones on language evolution in many aspects. Compared with the ILM and the naming-game based models (e.g., Baronchelli et al., 2006; Ke et al., 2002; Vogt, 2005), it explicitly defines word order rules and syntactic categories, and clearly traces a simultaneous acquisition of both lexical and syntactic knowledge. This clarification of linguistic knowledge allows it to further examine the word order bias shown in human lan-

languages (Gong, Minett, & Wang, 2009) and the linguistic ambiguity (Minett & Gong, 2010). In addition, its communication scenario involves no direct meaning transfer (A. D. M. Smith, 2003); linguistic comprehension is aided by both linguistic and nonlinguistic information. This makes it suitable for exploring the coordination of linguistic and nonlinguistic information during language acquisition. Furthermore, compared with the artificial neural network models (e.g., Batali, 1998; Christiansen & Ellefson, 2002), it explicitly defines individual learning mechanisms and illustrates their interactions in language processing. All these features are briefly illustrated in the following sections. A detailed description of this model can be found in Gong (2009).

2.1 Language and Individuals

Language in this model is represented by a set of mappings between meanings and utterances (*M-U mappings*). All meanings are from a semantic space and shared by individuals. 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 can be performed (e.g., “fox” or “tiger”). Some predicates take a single argument, for example, “run<tiger>” (meaning “a tiger is running”). Others take two, for example, “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). These predicates form two types of integrated meanings: “*predicate<agent>*” and “*predicate<agent, patient>*.” Each utterance is formed by a string of syllables chosen from a signaling space. An utterance that encodes an integrated meaning can be segmented into subparts, each mapping a particular semantic constituent; similarly, subparts that map semantic constituents can be combined to encode an integrated meaning. Such predicate–argument structures and combinable syllables have been adopted in many *structured simulations* (Wagner, Reggia, Uriagereka, & Wilkinson, 2003) such as the ILM.

In this model, individuals are simulated as artificial agents. Based on equipped learning mechanisms,

they can acquire linguistic knowledge from M-U mappings obtained in previous communications (see Section 2.2). In communications, they can produce utterances to encode integrated meanings and comprehend heard utterances (see Section 2.3). As language evolves across generations, some individuals may be replaced by new ones (children) having no linguistic knowledge, and different forms of cultural transmission are implemented by manipulating the participants of communications during language acquisition (see Section 2.4).

2.2 Linguistic Knowledge and Acquisition Mechanisms

Individual linguistic knowledge is represented by three components: a *lexicon*, a *syntax*, and a set of *syntactic categories*.

The lexicon of an individual consists of a set of *lexical rules* (see Figure 1 for examples). Some lexical rules are *holistic*, each mapping an integrated meaning onto an utterance (a *sentence*). A holistic rule is denoted, for example, by “run<tiger>” \leftrightarrow /abcd/, which indicates that the meaning “run<tiger>” may be produced by the string /abcd/, and also that /abcd/ may be interpreted as “run<tiger>.” Other lexical rules are *compositional*, each mapping particular semantic constituent(s) onto subpart(s) of an utterance. A compositional rule is denoted, for example, by “fox” \leftrightarrow /ef/ (a *word rule*) or “chase<wolf, #>” \leftrightarrow /gh*i/ (a *phrase rule*, here # refers to an unspecified semantic constituent and * unspecified string syllable(s)).

Holistic rules allow individuals to directly produce meaningful sentences, but the use of compositional rules requires that these rules be sequentially regulated so that they can form a meaningful sentence. The order of words or phrases in an utterance is regulated by a set of *syntactic rules* (see Figure 1 for examples), each specifying a relative order between two lexical items. For example, “tiger” \ll “fox” denotes that the constituent “tiger” should be produced in an utterance *before*—but not necessarily immediately before—the constituent “fox.” A local order between two lexical items helps to express a “predicate<agent>” meaning; similarly, two or three local orders among three lexical items help to express a “predicate<agent, patient>” meaning.

Syntactic categories are formed in order for syntactic rules acquired for some lexical items to be

Lexical rules	
Holistic rules:	Compositional rules:
(a) "chase<wolf, bear>" \leftrightarrow /a b/ (0.5)	(c) "wolf" \leftrightarrow /f/ (0.6)
(b) "hop<deer>" \leftrightarrow /c/ (0.4)	(d) "chase<#, bear>" \leftrightarrow /e h * g/ (0.7)
Syntactic rules	
(1) Category 1 (S) << Category 2 (V) (SV) (0.8)	
(2) Category 3 (O) >> Category 2 (V) (VO) (0.4)	
Categories	
Category 1 (S): <i>List of lexical rules:</i>	
{"wolf" \leftrightarrow /b c/ (0.7)} [0.5]	
<i>List of syntactic rules:</i>	
Category 1 (S) << Category 2 (V) (0.8)	
Category 3 (O) >> Category 1 (S) (0.4)	
Category 2 (V): <i>List of lexical rules:</i>	
{"fight<#, #>" \leftrightarrow /e/ (0.6)} [0.5]	
<i>List of syntactic rules:</i>	
Category 1 (S) << Category 2 (V) (0.8)	
Category 3 (O): <i>List of lexical rules:</i>	
{"fox" \leftrightarrow /a/ (0.5)} [0.7]	
<i>List of syntactic rules:</i>	
Category 3 (O) >> Category 1 (S) (0.4)	

Figure 1 The examples of lexical rules, syntactic rules, and categories: # represents an unspecified semantic item, and * unspecified syllable(s); lexical rules are itemized by letters; S, V, and O denote the syntactic roles of categories; numbers enclosed by () denote rule strengths, and those by [] association weights; << is the local order before, and >> after. Compositional rules can combine if they specify each constituent of an integrated meaning exactly once. For example, the two compositional rules (c) and (d) can combine to form "chase<wolf, bear>," and the corresponding utterance is <ehfg>.

applied productively to others having the same thematic role. A syntactic category (see Figure 1 for examples) comprises both a set of lexical rules and a set of syntactic rules that may operate on these lexical rules and regulate the orders between these lexical rules and those from other categories. A syntactic category associating lexical rules that encode the thematic role of agent is referred to as a subject (S) category, because 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. A syntactic rule between two categories can be denoted by their syntactic roles. For example, a local order *before* between lexical items from an S category and a V category can be denoted by S<<V, or simply SV.

Lexical and syntactic knowledge collectively help to 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 from these categories, the sentence can be either /bcea/ or /bcae/, following a word order SVO or SOV.

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 its local order. In addition, a lexical rule 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 new association weight of a lexical rule to a category. These numerical parameters make

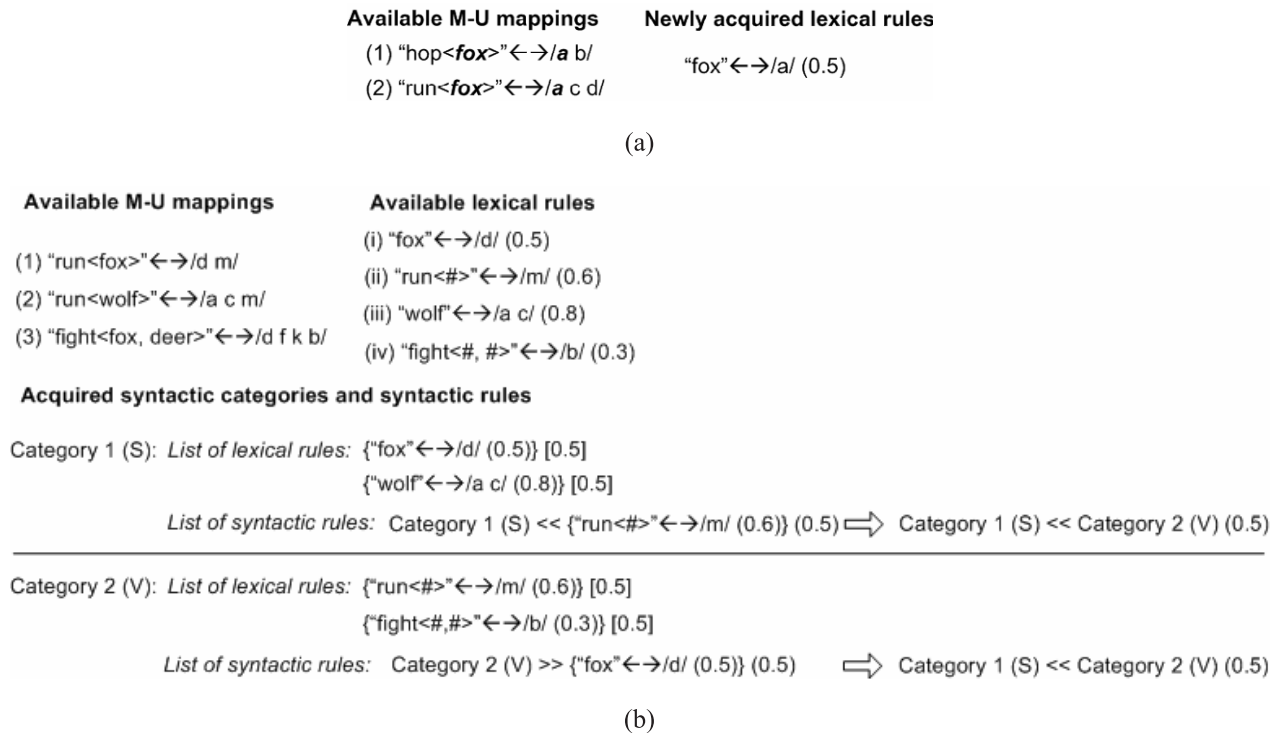


Figure 2 The examples of acquisition of linguistic knowledge: (a) the acquisition of lexical rules; (b) the acquisition of syntactic categories and syntactic rules. # represents unspecified semantic constituent; S, V, and O denote the syntactic roles of syntactic categories; numbers enclosed by () denote rule strengths, and those by [] association weights; << is the local order before, and >> after. M-U mappings are itemized by Arabic numbers, and lexical rules by Roman numbers.

possible a strength-based competition in communications (see Section 2.3) and a gradual forgetting of linguistic knowledge. *Forgetting* 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 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. Based on different linguistic instances, individuals may acquire various forms of lexical or syntactic knowledge, or use different ways to encode the same meaning. Through competition and forgetting, frequently used linguistic knowledge is strengthened.

Individuals apply some general learning mechanisms to acquire linguistic knowledge.

Lexical rules are acquired by detecting recurrent patterns in M-U mappings. A *recurrent pattern* is one or more meanings and one or more syllables appear-

ing recurrently in at least two M-U mappings. Each individual has a *buffer* storing a fixed number of M-U mappings obtained from previous communications. New mappings are compared with those in the buffer before they are inserted into the buffer. For example, in Figure 2a, by comparing the M-U mappings "hop<fox>" \leftrightarrow /ab/ and "run<fox>" \leftrightarrow /acd/, an individual can detect the recurrent pattern "fox" and /a/, and map it as a lexical rule "fox" \leftrightarrow /a/ with initial strength 0.5.

Syntactic categories and syntactic rules are acquired based on the thematic roles of lexical rules and order relations of their utterances in M-U mappings. For example, in Figure 2b, evident in the M-U mappings (1) and (2), the syllables /d/ of rule (i) and /a c/ of rule (iii) precede /m/ of rule (ii). Since "wolf" and "fox" are both agents in these meanings, rules (i) and (iii) are associated into a new S category (Category 1). Similarly, in M-U mappings (1) and (3), the syllables /m/ of rule (ii) and /b/ of rule (iv) follow /d/ of rule (i), which leads to a V category (Category 2) associating

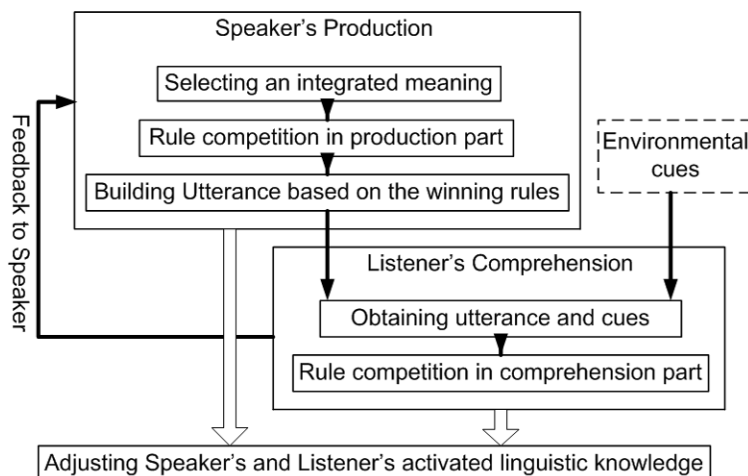


Figure 3 The diagram of utterance exchange. The dotted block indicates the unreliable information.

rules (ii) and (iv). Along with the acquisition of syntactic categories, the detected local orders are acquired as syntactic rules. During the acquisition of Category 1 (S), the local order *before* between rules (i) and (iii) and rule (ii) is acquired as a syntactic rule. Similarly, another syntactic rule is acquired in Category 2 (V). Now, since Categories 1 and 2 associate rules (i) and (iii), and rules (ii) and (iv), respectively, those two syntactic rules become “Category 1 (S) << Category 2 (V),” which indicates that the syllables of lexical rules in the S category should precede those of lexical rules in the V category.

These item-based learning mechanisms are traced in empirical studies (e.g., Mellow, 2008), and the categorization process resembles what is described in the *verb-island hypothesis* (Tomasello, 2003): individuals gradually form syntactic categories to regulate available lexical items and associate novel ones that are similarly used in utterances, and encode items having the same thematic roles.

2.3 Communication Without Direct Meaning Transfer

In the early stage of language acquisition, nonlinguistic information is important to help children to grasp the encoded meanings in communications (Tomasello, 2003). In this model, nonlinguistic information is simulated as *environmental cues*, each consisting of an integrated meaning plus a fixed strength. Cues are *unreliable*; otherwise, the exchanged meanings are

transparent without the need of linguistic communications. We define *Reliability of Cue* as the probability for a cue to contain the speaker’s encoded meaning. For example, if it is 0.6, there is a 60% chance for a *correct cue* that contains the speaker’s encoded meaning to be available to the listener; otherwise, a *wrong cue* that contains a meaning distinct from the speaker’s encoded one is presented to the listener.

Each communication involves two individuals (a speaker and a listener) who perform a number of *utterance exchanges*. An utterance exchange (see Figure 3) proceeds as follows. In production, the speaker (referred to hereafter as “she”) first randomly selects an integrated meaning from the semantic space to produce. She then activates some lexical rules and their categories that can encode some or all constituents of this meaning, and some syntactic rules that can regulate these lexical rules to form a sentence. These linguistic rules comprise candidate sets for production, each allowing her to encode the whole meaning into a sentence. Then, following Equation 1 (in which *Avg* means taking average, *aso* taking association weights, and *str* taking rule strengths), she calculates the combined strength ($CS_{production}$) of each set.

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

$CS_{production}$ 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 component 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 the lexical rules in this set, and the second is the association weights of those lexical rules to the categories in this set. For example, referring to Figure 1, the three categories, their lexical rules “wolf,” “fight,” and “fox”, and syntactic rules SV and SO form a candidate set to encode “fight<wolf, fox>”. Its $CS_{production}$ is: 0.6 (the contribution of lexical knowledge: $(0.7 + 0.6 + 0.5)/3$) + 0.38 (the contribution of syntactic knowledge: $(0.8 \times (0.7 + 0.6)/2 + 0.4 \times (0.7 + 0.5)/2)/2$) = 0.98.

After calculation, the speaker identifies the set of *winning rules* with the highest $CS_{production}$, builds up the sentence accordingly, and transmits the sentence to the listener. If she fails to construct a sentence to encode the whole meaning, she may occasionally (based on a *rate of random creation*) create a holistic rule to encode this meaning.

In comprehension, the listener (referred to hereafter as “he”) receives the sentence from the speaker and an environmental cue. Based on his linguistic knowledge, he activates some 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 rules form candidate sets for comprehension.

The environmental cue assists comprehension in several conditions. First, if the meaning in the cue matches exactly the meaning provided by some linguistic rules, the cue is put together with those rules to form a candidate set. Second, 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 those rules to form a candidate set, and its meaning becomes the meaning of this set. For example, if the linguistic rules only interpret the heard sentence as “chase<tiger, #>,” with the patient being unspecified, and the cue has the meaning “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. Finally, 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, #>,” 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 ($CS_{comprehension}$) of each set. For a set without a cue, the calculation is identical to that in production; for a set with a cue, the contribution of the cue, in terms of cue strength, is added to $CS_{comprehension}$; and for a set with only a cue, the cue strength itself becomes $CS_{comprehension}$.

$$\begin{aligned}
 CS_{comprehension} = & Avg(str(LexRule(s))) \\
 & + Avg(aso(Cats) \times str(SynRule(s))) \\
 & + str(Cue)
 \end{aligned} \tag{2}$$

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 competing ones in other candidate sets 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 strength and association weight, the contribution of linguistic (lexical and syntactic) information is 0.75 ($0.5 + 0.5 \times 0.5$). In order to treat linguistic and nonlinguistic information equally, we set both the cue strength and the confidence threshold to 0.75.

Equations 1 and 2 exemplify quantitatively a *multilevel selection* (Steels, van Trijp, & Wellens, 2007) among lexical, syntactic, and nonlinguistic information, and illustrate how nonlinguistic information assists linguistic comprehension by clarifying constituent(s) not specified by linguistic knowledge. This strength-based competition is independent of language-learning abilities. Other forms of competition that can integrate lexical, syntactic, and nonlinguistic information in production and comprehension can be adopted

as well. In this way, the linguistic knowledge of individuals that participate in communications tends to be similar, thus leading to *conventionalization* (a social agreement for meaning–utterance associations; Burling, 2005; Chater & Christiansen, 2009) among idiolects. In addition, throughout an utterance exchange, there is no check whether the speaker’s encoded meaning matches the listener’s decoded one. These features characterize linguistic communications based on evolving knowledge and unreliable cues.

2.4 The Acquisition Framework

This framework is shown in Figure 4. Figure 4a shows what happens during generation turnover, and Figure 4b defines three forms of transmission in the learning stage. H transmission refers to *Child-to-Child transmission* between two children; V transmission refers to *Parent-to-Child transmission* from a parent to its offspring; and O transmission refers to *Adult-to-Child transmission* from an adult to a child who is not the

offspring of that adult. All these forms of transmission are randomly interwoven, and their ratios in the total number of transmissions during the learning stage in each generation are manipulated respectively by $CCrate$, $PCrate$, and $ACrate$, the sum of which is 1.0.

Strictly speaking, this is a purely cultural evolution setup, without genetic inheritance in reproduction. Following the critical period hypothesis, adults and children are distinguished based on their learning behaviors in transmissions: During utterance exchanges, only children update their linguistic knowledge based on feedbacks, whereas adults do not. A similar manipulation was adopted in other studies (e.g., Hurford & Kirby, 1999). In addition, V and O transmissions are distinguished: If there is more than one V transmission, a child keeps sampling from its parent; however, if there is more than one O transmission, a child can sample from more than one adult. Considering that this child may also interact with the offspring of those adults in future H transmission, O transmission may play a role different from V transmission in language evolution. Furthermore, instead of actual number of transmissions in which a particular child is involved, $CCrate$, $PCrate$, and $ACrate$ control the ratios of these forms of transmission in the total number of cultural transmissions.

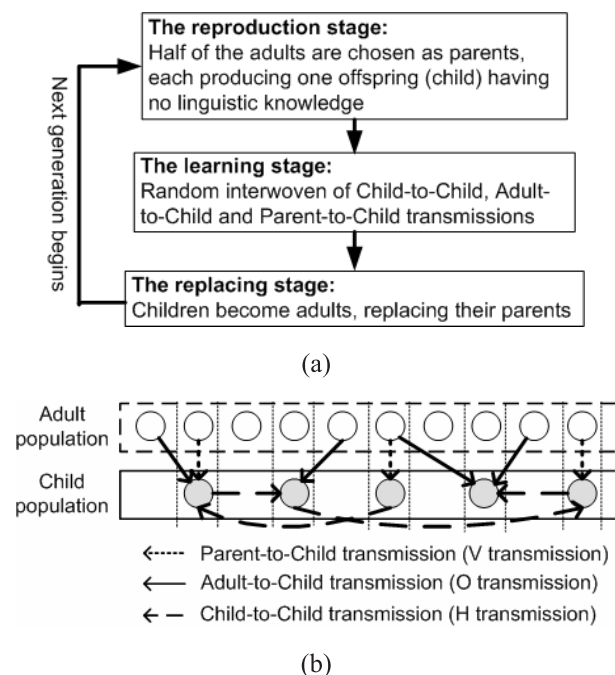


Figure 4 The language acquisition framework: (a) the three stages during generation turnover; (b) the different forms of transmission in the learning stage. Different types of balls represent adult and child individuals. Different types of arrows represent different forms of cultural transmission. Each arrow starts from the speaker and points to the listener.

3 The Simulation Results

3.1 The Simulation Setup

Table 1 summarizes the parameter setting in the simulations reported in this article. The 64 integrated meanings in the semantic space are formed by 12 constituents (four as agents or patients, four as single-argument predicates, and four as double-argument predicates). There are in total 16 (4×4) “predicate<agent>” and 48 ($4 \times 4 \times (4 - 1)$) “predicate<agent, patient>” meanings. Individuals have an equal chance to select any of these meanings in production. The population contains 10 adults, 5 of which are randomly chosen for reproduction in each generation. In the learning stage, the rule forgetting occurs every five times of transmission (scaled to the size of the child population). The robustness of the results with respect to changes in some of these parameters is evaluated in the Appendix.

Similar to Gong et al. (2010), we use three indices to evaluate linguistic understandability in different generations:

Table 1 The general parameter setting, in which the upper part concerns the learning mechanisms and communication and the lower part concerns the acquisition framework.

Parameter	Value
Size of semantic space	64
Size of signaling space	30
Size of individual buffer	40
Size of rule list for lexical rules, categories, and syntactic rules	60
Number of utterance exchange per transmission	20
Rate of random creation of holistic rules	0.25
Adjusting value on rule strength and association weight during competition	0.1
Deducting value on rule strength and association weight during forgetting	0.01
Reliability of cue	0.6
Population size	10
Forgetting frequency	5
Number of generations	100
Number of transmission in the learning stage of each generation	200

1. *Understanding Rate (UR)*, calculated in Equation 3 as the average percentage of integrated meanings understandable to each pair of adults in one generation. During the calculation, no cue is available to listeners.

$$UR = \frac{\sum_{i,j,i \neq j} \left(\begin{array}{c} \text{No. accurately understood} \\ \text{integrated meanings between} \\ \text{agents } i \text{ and } j \end{array} \right)}{\left(\begin{array}{c} \text{Size of semantic space} \\ \times \text{Population size} \\ \times (\text{Population size} - 1) \times 0.5 \end{array} \right)} \quad (3)$$

2. *UR* between two consecutive generations (UR_{con}), calculated as *UR* between adults in generation i and those in generation $i + 1$. High UR_{con} indicates that a communal language is well understood by individuals from consecutive generations.
3. *UR* between the first and later generations (UR_{mi}), calculated as *UR* between adults in generation 1 and those in generation i . High UR_{mi} indicates that

individuals in a later generation can well understand the language of the first generation; in other words, an initial language is largely preserved in a later generation.

We conduct two sets of simulations:

1. *Language origin simulations*, in which adults in the first generation share a small number (eight) of holistic rules (with the highest strength 1.0) each expressing an integrated meaning. This setting is based on the assumption that individuals acquire all the semantic constituents from the initial holistic rules. In fact, simulations starting from no initial holistic rules show similar results.
2. *Language change simulations*, in which adults in the first generation share a communal language capable of expressing all integrated meanings in the semantic space. This language consists of 12 lexical rules (with the highest strength 1.0) each encoding a semantic constituent, three categories (S, V, and O) that associate these lexical rules

Table 2 The 54 cases of the cultural setting.

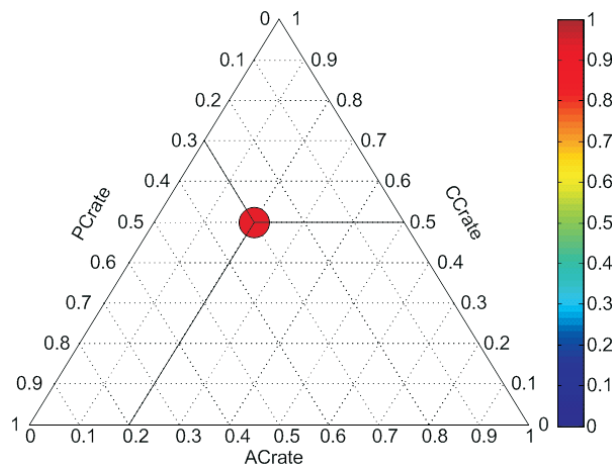
Cases	$PCrate$	$ACrate$	$CCrate$
1	0.0	0.1	0.9
2	0.0	0.2	0.8
3	0.0	0.3	0.7
⋮	⋮	⋮	⋮
9	0.0	0.9	0.1
10	0.1	0.0	0.9
11	0.1	0.1	0.8
⋮	⋮	⋮	⋮
18	0.1	0.8	0.1
19	0.2	0.0	0.8
⋮	⋮	⋮	⋮
51	0.7	0.2	0.1
52	0.8	0.0	0.2
53	0.8	0.1	0.1
54	0.9	0.0	0.1

(with the highest association weight 1.0), and three syntactic rules (SV, VO, and SO, with the highest strength 1.0) in these categories that form a global order SVO.

Based on $PCrate$, $ACrate$, and $CCrate$, we set up 54 cases in each set of simulations (see Table 2), which largely extend the limited conditions in Gong et al. (2010). Note that the case of purely H transmission ($PCrate = 0.0$, $ACrate = 0.0$, $CCrate = 1.0$) is excluded, because the communal language cannot be transmitted across generations in this case. In each case, 20 simulations are conducted for statistical analysis.

3.2 The Simulation Results

We use a surface ternary plot to illustrate the linguistic understandability under different combinations of $PCrate$, $ACrate$, and $CCrate$. Figure 5 gives an exam-

**Figure 5** An example of the surface ternary plot.

ple of this plot, in which the three axes trace $PCrate$, $ACrate$, and $CCrate$, the patch records the result in the case ($PCrate = 0.3$, $ACrate = 0.2$, $CCrate = 0.5$), and the color map beside indicates the value of this patch (1.0). Such a plot can reveal the tendencies of linguistic understandability with respect to changes in $PCrate$, $ACrate$, and $CCrate$.

Based on the surface ternary plot, Figure 6 shows the peak UR ($peak-UR$) and average UR_{con} ($avg-UR_{con}$) across 100 generations in the language origin simulations, and Figure 7 shows the $avg-UR_{con}$, the average UR_{ini} ($avg-UR_{ini}$) across 100 generations, and the UR at the end of 100 generations ($last-UR$) in the language change simulations. These results are analyzed from three aspects: the UR values in different regions, the tendencies of UR values along movements in these plots, and the regions having the highest UR values.

3.3 The UR Values in Different Regions

There are two regions in which the UR values are low. The first one is near the left angle, where $PCrate$ is high, but $ACrate$ and $CCrate$ are low, that is, V transmission is dominant in the learning stage. The low UR values here indicate that if cultural transmission is mainly vertical, it is difficult to trigger a communal language with good understandability or maintain an initial language across generations.

This region resembles Case 1 in Gong et al. (2010) and the results are similar. In the Appendix, we conduct further simulations in which adults can also

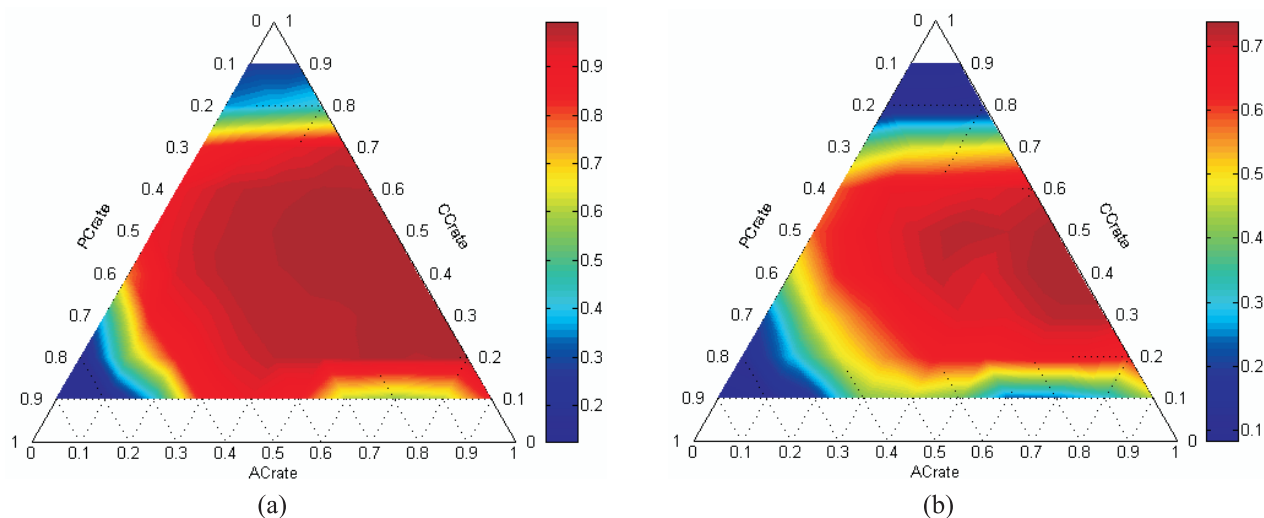


Figure 6 The results of the language origin simulations: (a) *peak-UR*, the highest value is obtained in Case 5 ($PCrate = 0.0$, $ACrate = 0.5$, $CCrate = 0.5$); (b) *avg-UR_{con}*, the highest value is obtained in Case 6 ($PCrate = 0.0$, $ACrate = 0.6$, $CCrate = 0.4$). In these simulations, individuals in the first generation share eight holistic rules that can express a small proportion of the integrated meanings in the semantic space.

update their linguistic knowledge in V and O transmissions, and the results remain similar. These results are distinct from those in the ILM and its later version (e.g., K. Smith & Hurford, 2003) in which V transmission alone led to a communal language with good understandability. Despite the different language evolution models, these difference results are caused mainly by indirect meaning transfer in our communication scenario. In our model, because children in each generation initially have no linguistic knowledge, their comprehension of adults' expressions relies mainly upon cues. The occasional "wrong" cues may cause children to develop salient knowledge not widely acceptable. Given only V transmission, children have no chances to further coordinate their knowledge with each other. After they replace adults and talk to new children in the next generation, the idiolects will continue diverging and the understandability of the communal language will remain poor. However, in the ILM, the explicit meaning transfer makes both the adults' utterances and their encoded meanings transparent to children. Based on similar learning mechanisms, children can develop similar linguistic knowledge to that of adults. Then, after a sufficient number of generations, a communal language with good understandability can be triggered given only V transmission. As shown in Table 4 in

Gong et al. (2010), once the reliability of cue is set to 1.0 (similar to the case of explicit meaning transfer), our model can produce results similar to the ILM.

The second region with low *UR* values is near the top angle, where $PCrate$ and $ACrate$ are low, but $CCrate$ is high, that is, H transmission is dominant. The low *UR* values here indicate that if cultural transmission is mainly horizontal, neither the origin of a communal language with good understandability nor the maintenance of an initial language is possible. This region resembles Case 3 in Gong et al. (2010) and the results are similar.

Unlike the top and left angles, the *UR* values near the right angle are relatively high, especially in the language change simulations. In this region, $PCrate$ and $CCrate$ are low, but $ACrate$ is high, that is, O transmission is dominant in the learning stage. This result reveals the role of O transmission, which is not discussed in Gong et al. (2010). Similarly to V transmission, O transmission allows children to sample the language of the previous generation, but more than one period of O transmission allows one child to sample from multiple adults. Therefore, O transmission is more efficient than V transmission to spread linguistic knowledge among multiple individuals across generations, as indicated by the higher *UR* values in the right angle than those in the left angle.

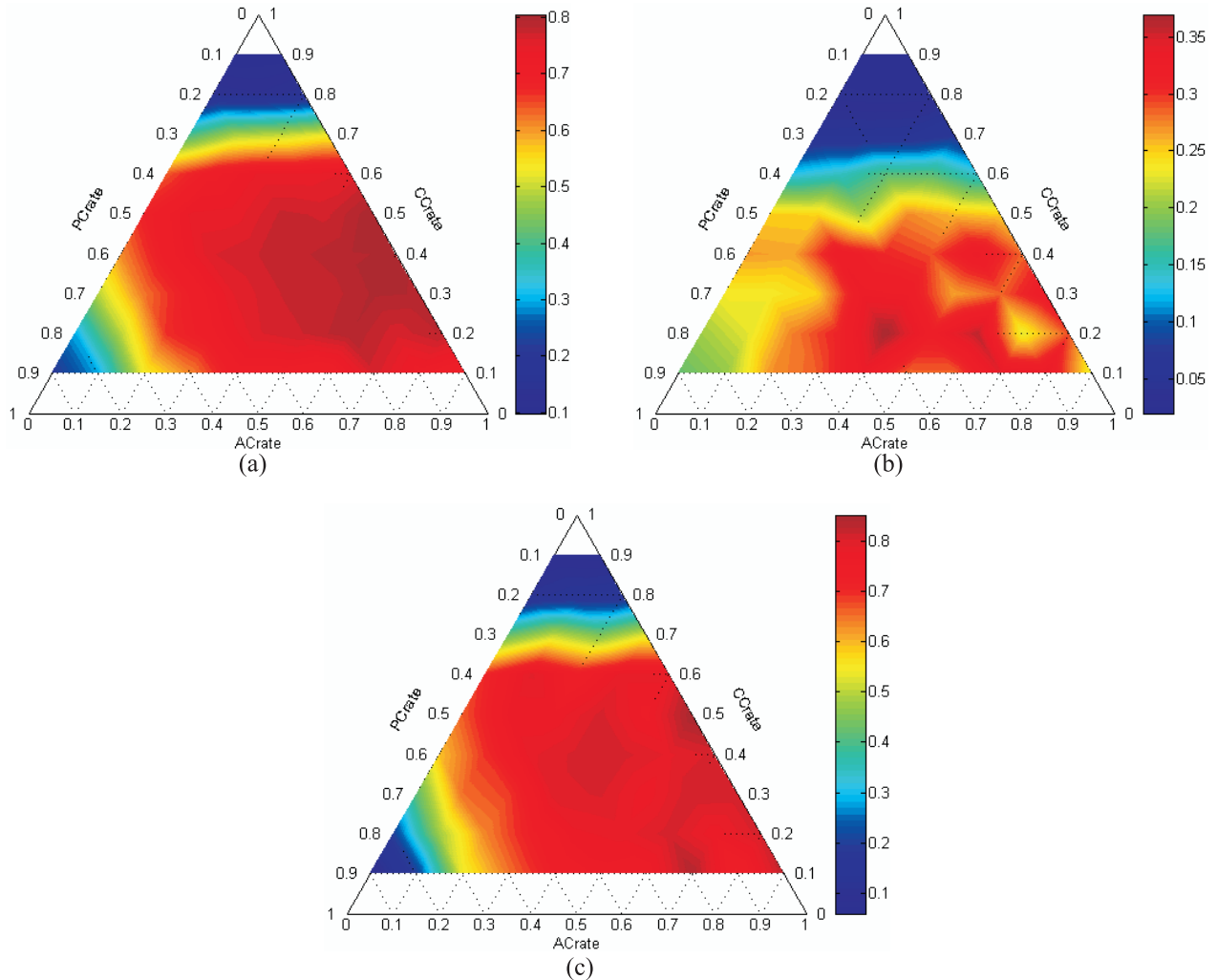


Figure 7 The results of the language change simulations: (a) $avg-UR_{con}$, the highest value is obtained in Case 5 ($PCrate = 0.0$, $ACrate = 0.5$, $CCrate = 0.5$); (b) $avg-UR_{inj}$, the highest value is obtained in Case 38 ($PCrate = 0.4$, $ACrate = 0.4$, $CCrate = 0.2$); (c) $last-UR$, the highest value is obtained in Case 5 ($PCrate = 0.0$, $ACrate = 0.5$, $CCrate = 0.5$). In these simulations, individuals in the first generation share a compositional language that can express all integrated meanings in the semantic space.

3.4 The UR Tendencies Along Movements

Apart from the three angles in which the three forms of transmission are dominant respectively, particular movements on these plots provide another way to analyze the relation between linguistic understandability and different combinations of H, V, and O transmissions.

First, along the movement from the top angle to the left angle following a line parallel to the axis of $PCrate$ (under a fixed $ACrate$, say 0.1), UR starts

from a low value when $PCrate$ is low and $CCrate$ is high, then increases, and finally decreases again when $PCrate$ is high and $CCrate$ is low. During this movement, H transmission is gradually replaced by V transmission. This movement is similar to Case 2 in Gong et al. (2010), where the ratio between intra- (H transmission) and intergenerational (V and O transmissions) transmissions changes from 160:40 to 80:120, and a similar increasing–decreasing tendency of UR is shown in that study. This tendency confirms the conclusion that either excessive H or excessive V trans-

mission fails to trigger a communal language with good understandability.

Second, along the movement from the left angle to the right angle following a line parallel to the axis of $ACrate$ (under a fixed $CCrate$, say 0.3), UR starts from a low value when $PCrate$ is high and $ACrate$ is low, and then increases when $PCrate$ is low and $ACrate$ is high. During this movement, V transmission is gradually replaced by O transmission, and the increase in UR confirms the conclusion that O transmission is more efficient than V transmission for spreading linguistic knowledge across generations of individuals.

Third, along the movement from the top angle to the right angle following a line parallel to the axis of $CCrate$ (under a fixed $PCrate$, say 0.2), UR starts from a low value when $CCrate$ is high and $ACrate$ is low, then increases, and finally decreases slightly when $CCrate$ is low and $ACrate$ is high. During this movement, H transmission is gradually replaced by O transmission. The initially low UR values indicate that without sufficient O transmission, linguistic understandability cannot be preserved across generations. Meanwhile, the final slight decrease in UR suggests that O transmission is less efficient than H transmission in conventionalization of idiolects within generations. This is mainly because H transmission gives children opportunities to not only acquire some knowledge (as listeners) but also use it in communications (as speakers), while O transmission only allows children to be passive learners. Both aspects indicate that given a fixed proportion of V transmission, both H and O transmissions are necessary to achieve mutual understanding within and across generations.

3.5 The Regions With the Highest UR

The highest UR in every plot is obtained in regions with particular values of $PCrate$, $ACrate$, and $CCrate$. The captions of Figures 6 and 7 record the cases with the highest UR values. Some nearby cases also have similarly high values.

UR and UR_{con} trace linguistic understandability within and across consecutive generations. The regions with the highest UR and UR_{con} are close to the axis of $CCrate$, in which both $ACrate$ and $CCrate$ are higher than $PCrate$. This indicates that in a multiagent cultural environment, O and H transmissions are more important than V transmission to spread linguistic

knowledge within and across generations of individuals. Similarly high $ACrate$ and $CCrate$ in these regions show that conventionalization of idiolects during H or O transmission is crucial to achieve mutual understanding. This conclusion is further confirmed in the Appendix based on the results of another language evolution model in which the critical period hypothesis is removed.

UR_{ini} reflects the extent to which an initial language is preserved after many generations. The regions with the highest UR_{ini} are close to the axis of $ACrate$, in which $PCrate$ and $ACrate$ are higher than $CCrate$. This suggests that in a multiagent cultural environment, V and O transmissions are more important than H transmission to preserve linguistic knowledge across many generations. During V and O transmissions, adults are speakers who keep transmitting their idiolects to children for them to develop linguistic knowledge. However, during H transmission, children can be speakers as well. Before acquiring sufficient linguistic knowledge from the adult population, children may create salient knowledge and spread some of it to others. Therefore, although children's creativity in H transmission helps to introduce new types (e.g., compositional) of linguistic knowledge (Vogt, 2005), it could also introduce changes in the communal language.

Considering these, we can see an integrated role of O transmission that combines the roles of H and V transmissions. This role makes O transmission more necessary than V transmission in a multiagent cultural environment for maintaining linguistic understandability across generations. As shown in regions with either high UR and UR_{con} , or high UR_{ini} , $ACrate$ is always high. In addition, H transmission is also necessary for maintaining linguistic understandability within generations. Unifying the results of this model and those of another one in the Appendix, we can see that both H and O transmissions are more necessary than V transmission for language evolution in a multiagent cultural environment.

Finally, as shown in Figure 7, in regions with high UR and UR_{con} , UR and UR_{con} are higher than UR_{ini} . This indicates that even though good understandability of the communal language is preserved across generations, the initial language changes inevitably after several generations. Apart from UR , the inevitable language change can be observed by comparing the shared lexical knowledge among individuals of differ-

<p>Generation 0:</p> <p>UR= 1.0;</p> <p>Common lexical rules = 12</p> <p>Rule 1 (1.00):"fox"<->/28 14 /</p> <p>Rule 2 (1.00):"wolf"<->/27 /</p> <p>Rule 3 (1.00):"deer"<->/25 12 /</p> <p>Rule 4 (1.00):"goat"<->/15 /</p> <p>Rule 5 (1.00):"run<#>"<->/22 /</p> <p>Rule 6 (1.00):"hop<#>"<->/19 18 /</p> <p>Rule 7 (1.00):"cry<#>"<->/9 /</p> <p>Rule 8 (1.00):"jump<#>"<->/26 27 /</p> <p>Rule 9 (1.00):"fight<#>"<->/4 /</p> <p>Rule 10(1.00):"chase<#>"<->/28 14 /</p> <p>Rule 11(1.00):"stalk<#>"<->/13 /</p> <p>Rule 12(1.00):"attack<#>"<->/5 17 /</p>	<p>Generation 85:</p> <p>UR= 0.879; URcon= 0.821; URini= 0.358;</p> <p>Common lexical rules = 12</p> <p>Rule 1 (0.93):"fox"<->/28 /</p> <p>Rule 2 (0.93):"wolf"<->/27 /</p> <p>Rule 3 (0.93):"deer"<->/12 /</p> <p>Rule 4 (0.93):"goat"<->/15 /</p> <p>Rule 5 (0.92):"run<#>"<->/22 /</p> <p>Rule 6 (0.93):"hop<#>"<->/18 /</p> <p>Rule 7 (0.93):"cry<#>"<->/9 /</p> <p>Rule 8 (0.93):"jump<#>"<->/26 /</p> <p>Rule 9 (0.93):"fight<#>"<->/19 /</p> <p>Rule 10(0.90):"stalk<#>"<->/24 /</p> <p>Rule 11(0.93):"chase<#>"<->/13 /</p> <p>Rule 12(0.92):"attack<#>"<->/17 /</p>
---	--

Figure 8 The shared lexical knowledge among individuals in Generation 0 (left) and Generation 85 (right) in a language change simulation with $PCrate = 0.3$, $ACrate = 0.4$, $CCrate = 0.3$. For a shared lexical rule, the value within brackets is the average strength of this rule in all individuals, and the numbers within // are utterance syllables.

ent generations. Figure 8 shows such knowledge in two generations in a language change simulation. After several generations of transmission, some shared lexical rules have their utterances changed, which reduces UR_{ini} , but the high UR and UR_{con} indicate that the understandability of the communal language is high both within and across consecutive generations.

In V or O transmission, because of the learning bottleneck (Kirby, 2000) and unreliable cues, children tend to develop linguistic knowledge differently from adults. In H transmission, some of this knowledge may diffuse in the population via conventionalization, thus introducing changes in the communal language. Meanwhile, even if V or O transmission is sufficient, during H transmission, because of the implicit bottleneck (in H transmission, speakers may not express all meanings to listeners, who later on have to create new expressions to encode other meanings when talking to others; Vogt, 2005), some infrequently used knowledge may be gradually forgotten. When the meanings previously encoded by this knowledge are expressed, new knowledge has to emerge, some of which may diffuse in the population via future H, V, or O transmission, thus also introducing changes in the communal language. Considering these, cultural transmission contributes to the gradual change of the communal language. The cultural evolution of language can be viewed as a tinkering (Jacob, 1977) process, based on individual learning mechanisms and available infor-

mation during cultural transmission, and different forms of transmission collectively lead to a *dynamic equilibrium* of language evolution: individuals from consecutive generations can understand each other well (UR_{con} is high), but language is changing inevitably in the long run (UR_{ini} is not high).

4 Discussions, Conclusions, and Future Work

This article proposes an acquisition framework to study the roles of H, V, and O transmissions on language origin and change. The simulation results show that H transmission helps to maintain the understandability of the communal language within generations, and V and O transmissions help to preserve the understandability of the communal language across generations. These findings are consistent with those of Gong et al. (2010). In addition, we reveal an integrated role of O transmission that combines the roles of H and V transmissions. This role was not discussed in previous studies. Furthermore, as shown by the results of this model and another one in the Appendix, in a multiagent cultural environment, both O and H transmissions are more necessary than V transmission to preserve the understandability of the communal language across consecutive generations.

These conclusions can inspire reconsideration of the claim that language is solely determined by prior

learning biases (Chomsky, 1972; Dediu, 2009; Griffiths & Kalish, 2007). As shown in the Appendix, these conclusions are less dependent on the parameters that control the cultural (e.g., the population size, the number of children introduced in each generation, and the critical period hypothesis) and language-related settings (e.g., the size of the semantic space), and the adopted language evolution model. Therefore, these conclusions can be borrowed directly to study other cognitive, social, political, or economic phenomena in which cultural transmission is the major medium of information exchange. Moreover, by manipulating the probabilities of various forms of transmission, we implement some extreme conditions (e.g., the cases with zero V transmission) in which certain forms of transmission are inadequate or totally absent. These conditions, usually difficult to access in empirical studies, help to better understand the roles of particular forms of transmission. From this point of view, the simulation approach can assist the empirical approach.

The proposed acquisition framework adopts a *global* perspective that defines parameters to respectively manipulate the ratios of different forms of transmission in the total number of cultural transmissions. Apart from this *global* perspective, there is a *local* perspective that focuses on the number of different adults a child can interact with: If this number is 0, the transmission that involves this child is horizontal, only among children; if this number is 1, the transmission is vertical, with the child's parent; and if this number is more than 1, the transmission is either vertical or oblique, with the parent or other adults. This local perspective fails to evaluate the roles of different forms of transmission in language evolution, since V and O transmissions remain entangled, similarly to the setting in Gong et al. (2010).

In addition, language evolution in our model follows an *analytic* style (Wray, 2002): A compositional language emerges out of a holistic signaling system via segmentation of holistic utterances into words and combing rules. This style was inspired by the *gestalt* style (Prizant, 1983) in first or second language acquisition (e.g., Fillmore, 1979). However, based on the *synthetic* style (Peters, 1977) in first language acquisition, some scholars (e.g., Calvin & Bickerton, 2000; Jackendoff, 2002) suggest that language may evolve in a synthetic manner, starting from isolated words, followed by word combination and syntactic hierarchy. Considering the independence of the acquisition

framework, the conclusions on the roles of different forms of transmission in language evolution will still hold in models that follow the synthetic style.

Furthermore, based on the results of two language evolution models, we draw similar conclusions that H and O transmissions are more necessary than V transmission for language evolution in a multiagent environment. This seems to imply that the predicate–argument structure and relevant syntactic processing in the lexicon–syntax coevolution model are useless to explore the roles of cultural transmission in language evolution, and V transmission is not necessary in this model. However, as shown in Figure 7b, the region with the highest UR_{ini} has a high $PCrate$, which indicates that V transmission is still necessary for preserving linguistic understandability across generations. Compared with the model in the Appendix, this necessity concerns the syntactic processing in the lexicon–syntax coevolution model. Recalling the model details, acquisition of syntactic rules and categories is more difficult than that of lexical rules, since the former requires not only common lexical items but also similar usage of these items in exchanged sentences, while the latter only requires some recurrent patterns in exchanged sentences. This is also shown in real children's language acquisition. Noting this, compared with learning from multiple adults whose linguistic knowledge, especially syntactic knowledge, may not be identical, consistently learning from a particular adult in V transmission is relatively easier for a child to acquire that adult's syntactic knowledge. This may explain why V transmission is necessary to maintain identical knowledge across many generations. In this sense, the involvement of syntactic processing actually reveals some hidden role of a certain form of cultural transmission in language evolution, and a more language-like model is necessary to comprehensively understand the relative roles of different forms of cultural transmission in language evolution.

Finally, there are some promising future directions for the current study. For example, as discussed above, we can further examine the role of V transmission based on models involving syntactic knowledge and syntactic processing. We can also evaluate the changing rate of linguistic knowledge across generations under different cultural settings. Moreover, in actual language acquisition, children may interact with their grandparents. Such V transmission is common in many cultures such as China or Russia, but disre-

garded in most previous studies. Simulating such transmission and studying its roles can enrich our understanding of the role of cultural transmission in language evolution.

Appendix: Robustness of the Results

In this appendix, we test the robustness of the simulation results by manipulating some parameters of this model and adopting another language evolution model. The effects of the parameters that control the learning mechanisms, such as the size of individual buffer for linguistic instances, the size of rule list for linguistic knowledge, the creation rate of holistic rules, and the reliability of cue, were discussed in detail in Gong (2009). In this section, we discuss the parameters such as the size of semantic space, the number of children in each generation, and the size of adult population. For each parameter, we consider two variations from the value used in this article, and collect the *UR* values in all 54 cases in the language origin simulations under these variations. In each case, 20 simulations are conducted, and average *peak-UR* is calculated.

Figure A.1 shows the results of two sets of simulations in which the semantic space contains 216 (formed by six agent, six single- and six double-argument predicate constituents) and 512 (formed by eight

agent, eight single- and eight double-argument predicate constituents) integrated meanings respectively. Because of the increase in the size of the semantic space, the number of transmissions in each generation is set to 800 and 1600 respectively, so that individuals can have enough samples of each integrated meaning. Compared with Figure 6, despite different *UR* values in the same cases, the general tendencies of *UR* in different regions and along different movements are similar, based on which the same conclusion on the roles of different forms of transmission in language evolution can be drawn, that is, the size of the semantic space cannot greatly affect the conclusion. Similarly, if the number of syllables in the signaling space is not too small, compared with the number of distinct semantic constituents (otherwise, some semantic constituents have to be mapped to more than one syllable to distinguish each other, but the chances for such recurrent patterns to occur are much lower), the size of signaling space cannot greatly affect the conclusion.

Figure A.2 shows the results of two sets of simulations in which the number of children introduced in each generation is 2 and 10 respectively. The number of transmissions in each generation is also changed to 80 and 400 respectively. Compared with Figure 6, the similar tendencies of *UR* indicate that the different settings in the generation turnover cannot greatly affect the conclusion.

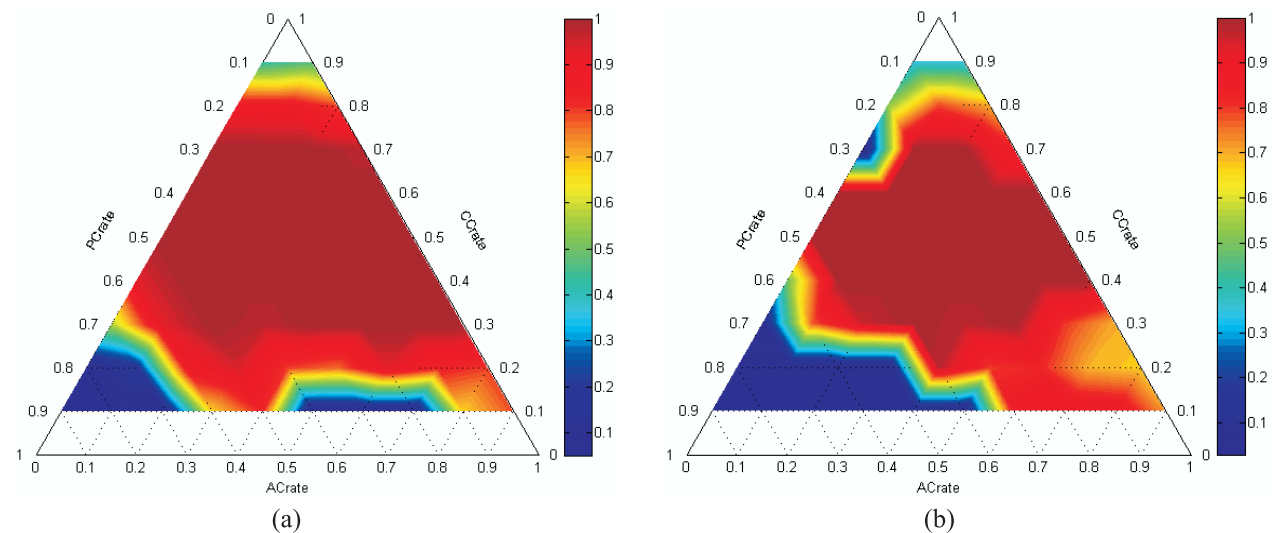


Figure A.1 The average *peak-UR* of the language origin simulations under different sizes of the semantic space: (a) 216 integrated meanings; (b) 512 integrated meanings.

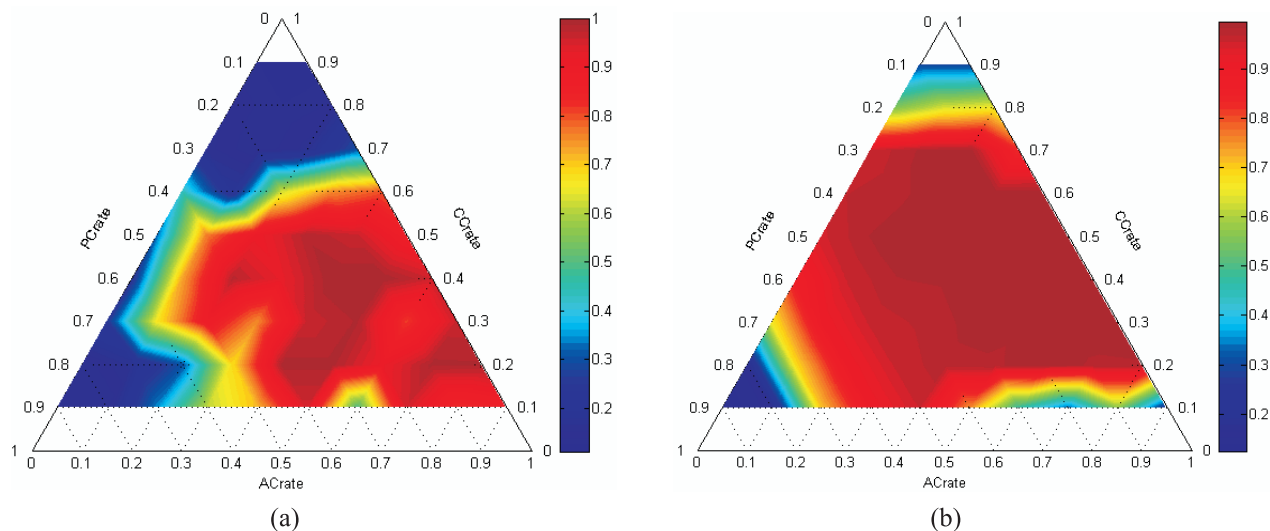


Figure A.2 The average *peak-UR* of the language origin simulations under different numbers of children introduced in each generation: (a) 2 children; (b) 10 children.

Figure A.3 shows the results of two sets of simulations in which the population size is set to 20 and 50 respectively. Because of the increase in the population size, the number of transmissions in each generation is set to 400 and 1000 respectively, and the total number of generations is set to 200 and 500 respectively. In each generation, half of the population is replaced by new children. Compared with Figure 6, the conclusion

on the roles of different forms of transmission in language evolution still holds under different population sizes.

Apart from these parameters, we also compare the *UR* values in simulations where adults can update their knowledge in V or O transmission. Instead of all 54 cases, we use one case as an example (Case 22: $PCrate = 0.2$, $ACrate = 0.3$, $CCrate = 0.5$), in which

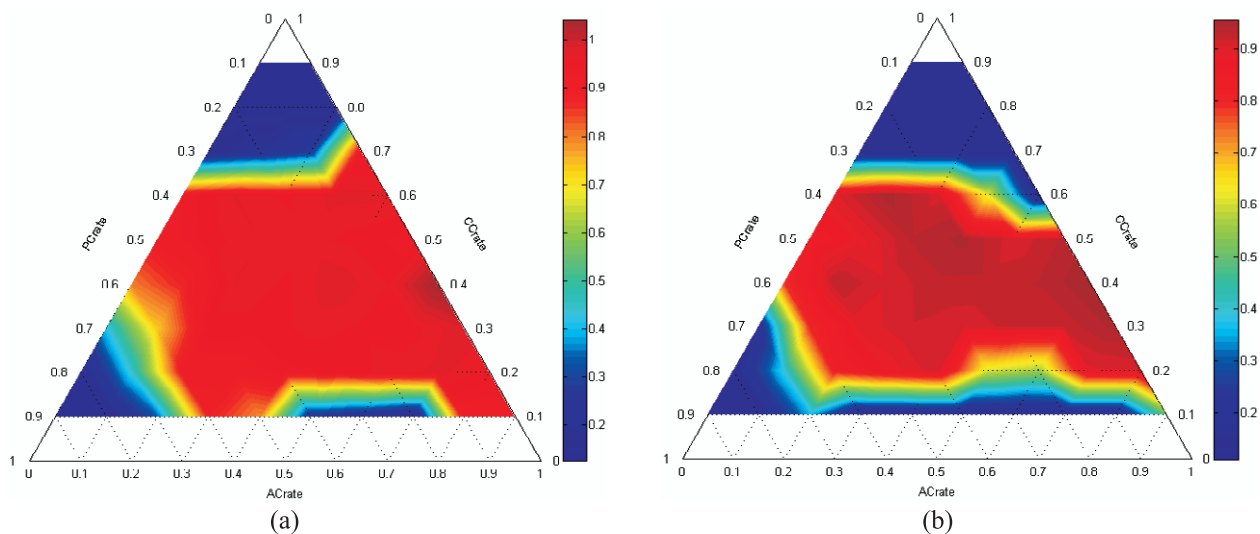


Figure A.3 The average *peak-UR* of the language origin simulations under different population sizes: (a) 20-agent population; (b) 50-agent population.

the *UR* value is high, as shown in Figure 6. Figure A.4 shows the average and standard error of *peak-UR* in simulations of this case where adults have different probabilities to update their linguistic knowledge in V or O transmission. The similar *peak-UR* values under different updating probabilities suggest that whether adults update their idiolects or not cannot greatly affect the roles of different forms of transmission in language evolution, that is, the critical period hypothesis cannot greatly affect the conclusion of this study.

Finally, we evaluate the independence of the acquisition framework based on a different language model, that is, the category game model (Puglisi, Baronchelli, & Loreto, 2008). Instead of coevolution of lexical items and word orders, this *unstructured* model (Wagner et al., 2003) simulates the coevolution of linguistic categories and their word labels. In this model, individuals have no syntactic knowledge or mechanisms to process syntactic information. Instead, they can form perceptual categories to discriminate N (?2) stimuli (the minimum distance between any two stimuli is larger than d_{min}) from a continuous perceptual channel ([0.0 1.0]) and update these categories and their word labels based on the result of discrimination. During a category game, if the listener fails to discriminate the topic from other stimuli, the speaker will point out the topic. This is not a direct check of individuals' knowledge, because individuals in general tend to develop different categorization patterns

using different boundaries to divide the perceptual channel. Indirect meaning transfer could be the only similarity between these two models. After some category games, some adjacent perceptual categories tend to use identical word labels and join into a linguistic category, and the boundaries of linguistic categories among individuals gradually become similar. The details and applications of this model can be found in Puglisi et al. (2008), Gong, Puglisi, Loreto, and Wang (2008), and Baronchelli, Gong, Puglisi, and Loreto (2010). In this model, *UR* or UR_{con} can be defined similarly as the percentage of successful category games (in which the listener successfully discriminates the topic based on his own knowledge and the speaker's word that describes the topic) among individuals of the same or consecutive generations. As a model that explores the origin of categorization patterns, it is meaningless to predefine a set of categories and observe their evolution. Therefore, language change simulations are not applicable. Based on the category game model, we reproduce *peak-UR* in all 54 cases in Figure A.5.

Compared with Figure 6, there are two apparent differences. First, based on the category game model, the *UR* values near the top angle, in which H transmission is excessive, are much higher than those based on the lexicon–syntax coevolution model. Because there

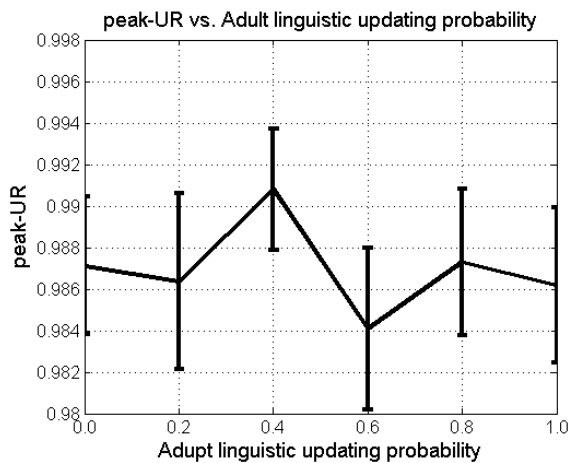


Figure A.4 The average and standard error (indicated by the error bar) of *peak-UR* in the language origin simulations of Case 22 ($PCrate = 0.3$, $ACrate = 0.2$, $CCrate = 0.5$) under different adult linguistic updating probabilities.

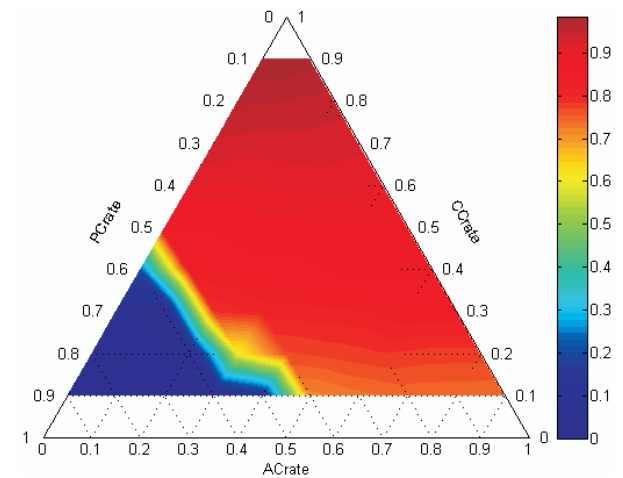


Figure A.5 The average *peak-UR* of the language origin simulations based on the category game model. $N = 2$ and $d_{min} = 0.01$. There are 50 individuals in the first generation, 25 of which reproduce in each generation. There are 100 generations and 5×10^6 transmissions in each generation.

is no rule forgetting in the category game model, communications involving many different individuals can greatly help conventionalization of the categorical knowledge among individuals. Since both H and O transmissions provide opportunities for individuals to interact with each other, *UR* values become high if either of these two forms of transmission is excessive. Second, based on the category game model, the highest *UR* values occur in regions close to the top angle. This reflects the fact that O transmission is less efficient than H transmission to achieve communal knowledge within a population.

Despite of these differences, we can draw a similar conclusion, based on Figure A.5, that both H and O transmissions are more necessary than V transmission to trigger a similar categorization pattern and high linguistic understandability in a multiagent cultural environment. This evidence indicates that the conclusion on the roles of different forms of transmission in language evolution is less dependent on the adopted language evolution models. It seems that given a suitable communication scenario involving indirect meaning, transfer of linguistic understandability within and across generations of individuals is mainly determined by different combinations of the three forms of transmission.

Acknowledgments

Some preliminary results of this study were reported in the 8th International Conference on Language Evolution (Evolang8). The author acknowledges support from the Alexander von Humboldt Foundation and the British Academy of Sciences, and thanks Simon Kirby from the University of Edinburgh and Bernard Comrie from the Max Planck Institute for Evolutionary Anthropology for their valuable suggestions. The article has also benefited a lot from the comments of three anonymous reviewers.

References

- Acerbi, A., & Parisi, D. (2006). Cultural transmission between and within generations. *Journal of Artificial Societies and Social Systems*, 9 (<http://jasss.soc.suney.ac.uk/9/1/9.html>).
- Axelrod, R. (1997). The dissemination of culture: a model with local convergence and global polarization. *Journal of Conflict Resolution*, 41(2), 203–226.
- Baronchelli, A., Felici, M., Loreto, V., Caglioti, E., & Steels, L. (2006). Sharp transition towards shared vocabularies in multi-agent systems. *Journal of Statistical Mechanism*, P06014.
- Baronchelli, A., Gong, T., Puglisi, A., & Loreto, V. (2010). Modeling the emergence of universality in color naming patterns. *Proceedings of the National Academy of Sciences of the USA*, 107(6), 2403–2407.
- Batali, J. (1998). Computational simulation 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, UK: Cambridge University Press.
- Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. Chicago, IL: Chicago University Press.
- Burling, R. (2005). *The talking ape: How language evolved*. Oxford: Oxford University Press.
- Calvin, W. H., & Bickerton, D. (2000). *Lingua ex machina: Reconciling Darwin and Chomsky with the human brain*. Cambridge, MA: MIT Press.
- Cavalli-Sforza, L. L., & Feldman, M. (1981). *Cultural transmission and evolution: A quantitative approach*. Princeton, NJ: Princeton University Press.
- Chater, N., & Christiansen, M. (2009). Language acquisition meets language evolution. *Cognitive Science*. Published Online: Jun. 22, 2009. DOI: 10.1111/j.1551-6709.2009.01049.x.
- Chomsky, N. (1972). *Language and mind*. New York: Harcourt Brace Jovanovich, Inc.
- Christiansen, M. H., & Ellefson, M. R. (2002). Linguistic adaptation without linguistic constraints: The role of sequential learning in language evolution. In: A. Wray (Ed.), *The transition to language* (pp. 335–358). Oxford: Oxford University Press.
- Christiansen, M. H., & Kirby, S. (2003). Language evolution: Consensus and controversies. *Trends in Cognitive Sciences*, 7(7), 300–307.
- Cornish, H. (2010). Investigating how cultural transmission leads to the appearance of design without a designer in human communication systems. *Interaction Studies*, 11(1), 112–137.
- Dediu, D. (2009). Genetic biasing through cultural transmission: Do simple Bayesian models of language evolution generalize? *Journal of Theoretical Biology*, 259(3), 552–561.
- 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.
- 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 the USA*, 99(1), 529–534.
- Griffiths, T. L., & Kalish, M. L. (2007). Language evolution by iterated learning with Bayesian agents. *Cognitive Science*, 31(3), 441–480.
- Gong, T. (2009). *Computational simulation in evolutionary linguistics: A study on language emergence* (Frontiers in Linguistics, Monograph IV). Taipei: Institute of Linguistics, Academia Sinica.
- 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., Puglisi, A., Loreto, V., & Wang, W. S.-Y. (2008). Conventionalization of linguistic knowledge under simple communicative constraints. *Biological Theory: Integrating Development, Evolution, and Cognition*, 3(2), 154–163.
- Hurford, J. R., & Kirby, S. (1999). Coevolution of language size and the critical period. In: D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 39–63). Mahwah, NJ: Lawrence Erlbaum.
- Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford & New York: Oxford University Press.
- Jacob, F. (1977). Evolution and tinkering. *Science*, 196, 1161–1166.
- 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.
- Kirby, S. (1998). *Language evolution without natural selection: From vocabulary to syntax in a population of learners*. Technical report, University of Edinburgh.
- Kirby, S. (2000). Syntax without natural selection: How compositionality emerges from vocabulary in a population of learners. In: C. Knight (Ed.), *The evolutionary emergence of language: Social function and the origins of linguistic form* (pp. 303–323). Cambridge, UK: Cambridge University Press.
- Labov, W. (2001). *Principles of linguistic change: Social factors*. Oxford, UK: Basil Blackwell.
- Lenaerts, T., Jansen, B., Tuyls, K., & De Vylder, B. (2005). The evolutionary language game: An orthogonal approach. *Journal of Theoretical Biology*, 235(4), 566–582.
- Lenneberg, E. (1967). *Biological foundations of language*. New York: John Wiley & Sons, Inc.
- Lewis, D. (1969). *Convention: A philosophical study*. Cambridge, MA: Harvard University Press.
- Mellow, J. D. (2008). The emergence of complex syntax: A longitudinal case study of the ESL development of dependency resolution. *Lingua*, 118(4), 499–521.
- Mesoudi, A., & Whiten, A. (2008). The multiple roles of cultural transmission experiments in understanding human cultural evolution. *Philosophical Transactions of the Royal Society B*, 363, 3489–3501.
- 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, China: Shanghai Education Press.
- Mufwene, S. S. (2001). *The ecology of language evolution*. Cambridge, UK: Cambridge University Press.
- Nowak, M. A., & Komarova, N. L. (2001). Towards an evolutionary theory of language. *Trends in Cognitive Sciences*, 5(7), 288–295.
- Nowak, M. A., Plotkin, & J. B., Krakauer, D. (1999). The evolutionary language game. *Journal of Theoretical Biology*, 200(2), 147–162.
- Penfield, W., & Roberts, L. (1959). *Speech and brain-mechanisms*. Princeton, NJ: Princeton University Press.
- Peters, A. M. (1977). Language learning strategies: Does the whole equal the sum of the parts? *Language*, 53(3), 560–573.
- Pinker, S. (1994). *The language instinct*. New York: Morrow.
- Prizant, B. M. (1983). Language acquisition and communicative behavior in autism: Toward an understanding of the “whole” of it. *Journal of Speech and Hearing Disorders*, 48(3), 296–307.
- Puglisi, A., Baronchelli, A., & Loreto, V. (2008). Cultural route to the emergence of linguistics categories. *Proceedings of the National Academy of Sciences of the USA*, 105(23), 7936–7940.
- Singleton, D., & Lengyel, Z. (Eds.) (1995). *The age factor in second language acquisition*. Clevedon, UK: Multilingual Matters.
- Skyrms, B. (2009). Evolution of signaling systems with multiple senders and receivers. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364, 771–779.
- Smith, A. D. M. (2003). Semantic generalization and the inference of meaning. In: W. Banzharf, T. Christaller, P. Dittrich, J. T. Kim, & J. Ziegler (Eds.), *Proceedings of the 7th European conference on artificial life* (pp. 499–506). Berlin: Springer-Verlag.
- Smith, K., Brighton, H., & Kirby, S. (2003). Complex systems in language evolution: The cultural emergence of compositional structure. *Advances in Complex Systems*, 6(4), 537–558.
- Smith, K., & Hurford, J. (2003). Language evolution in populations: Extending the Iterated Learning Model. In: W. Banzharf, T. Christaller, P. Dittrich, J. T. Kim, & J. Ziegler (Eds.), *Proceedings of the 7th European conference on artificial life* (pp. 507–516). Berlin: Springer-Verlag.
- Smith, K., & Kirby, S. (2008). Cultural evolution: Implications for understanding the human language faculty and its evo-

- lution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1509), 3591–3603.
- Steels, L. (2005). The emergence and evolution of linguistic structure: from lexical to grammatical communication systems. *Connection Science*, 17(3–4), 213–230.
- 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.
- Tomasello, M. (2003). *Constructing a Language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- Vogt, P. (2005). On the acquisition and evolution of compositional languages: sparse input and productive creativity of children. *Adaptive Behavior*, 13(4), 325–346.
- 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.
- Wray, A. (2002). *Formulaic language and the lexicon*. New York: Cambridge University Press.

About the Author



Tao Gong obtained his PhD from the Chinese University of Hong Kong in 2007. He is now an Alexander von Humboldt Postdoctoral Fellow at the Department of Linguistics, Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany. By means of computational modeling, he carries out research on evolutionary linguistics, including semantics–syntax coevolution, categorization, syntactic development, cultural transmission and social networks.