[Review]

Origins of Human Communication


KEN-ICHI IDEHARA
Shiga University*

Keywords: language acquisition, language development, cooperative communication, common ground, joint attention

1. Introduction

Until recently, the origin of human language was not extensively studied as a topic of linguistics. With advances in related sciences such as biology, psychology, and cognitive science, however, this “taboo” area has come to be actively explored in recent years. In the view of some geneticists and generativists, the discovery of FOXP2, which some researchers regard as a language gene, might help unravel the mystery of the origin of language as well as strengthen their fields’ theoretical foundations. Admittedly, this language-specific research may contribute to the elucidation of how “linguistic competence” arose, but this approach will not tell us why language was needed by our ancestors in certain circumstances. Nor will it answer the question of how human communication developed over time.

The book under review, Origins of Human Communication, is one of the few that place the question of the emergence of human language in the broader context of communication, gestures, social interactions, and so forth. As suggested by the word “origins” in the title, it is not unnatural to think that human languages emerged as a result of a process involving a multitude of human abilities and skills. Michael Tomasello, the author of this book, is one of the few persons who can integrate such cognitive factors and construct a unified theory of the birth of human language. He has written two books that deal mainly with language—Tomasello (1999) focuses on the sociocultural development of human beings and Tomasello

* I would like to thank two anonymous reviewers for valuable comments and suggestions. Needless to say, all remaining errors are my own.
(2003) pays more attention to ontogenetic language acquisition. Thus, in a sense, his new book may well be considered a culmination of his long term investigation into human communication.

This book consists of seven chapters. Chapter 1 outlines the purpose of the book: the “identification of the species-unique features of human communication and their ontogenetic and phylogenetic roots” (p. 11). Chapters 2 and 3 compare the cognitive skills of primates with those of human beings and examine why only humans have the highly sophisticated communication tool of language. Tomasello then discusses how the cognitive skills and motivations required for human communication appear ontogenetically (Chapter 4) and phylogenetically (Chapter 5). Chapter 6 discusses how and why grammar came into existence. Finally, Chapter 7 summarizes the book’s argument.

2. Overview

Chapter 1, “A Focus on Infrastructure,” is short but very important, in that we can see Tomasello’s view of language. He hypothesizes that, in the beginning, human language originated from such gestures as pointing and pantomiming, a theory supported by his enormous research on animal and human psychology. Though such gestures are seemingly too limited to carry abundant information, they can serve rather effectively when participants have shared knowledge, which Tomasello calls common (conceptual) ground. To form common ground, cognitive skills such as intention-reading, joint attention, and shared experience are required. Thus, one of Tomasello’s main claims is that these psychological infrastructures are needed for human communication.

Another element Tomasello emphasizes is the motivation to point, that is, to (reciprocally) inform others of some helpful information. Considering pointing is not found in the animal kingdom, this cooperative behavior is seen to be unique to human beings. Therefore, it is probable that this feature is a prerequisite to the birth of language.

These cognitive skills are discussed in more detail in Chapters 2 and 3, but it should be noted here that this idea suggests that pragmatics is far from a “wastebasket,” but rather one of the most fundamental areas of linguistics. Tomasello’s point of view conforms to recent research in not only cognitive linguistics or interactional sociolinguistics but also other fields of academic study, e.g. behavioral economics, which I will discuss in the next section.
Chapter 2, “Primate Intentional Communication,” starts by describing the vocal communication of nonhuman primates. Given that “primate vocalizations would seem to be mainly individualistic expressions of emotion, not recipient-directed acts” (p. 19) and not far removed from other mammalian modes of expression, Tomasello says the run-up to human language must not be in oral but visual channels, especially gestures.

Great apes have two kinds of gestural communication: intention-movements and attention-getters. Intention-movements enable recipients to understand what communicators will do just by performing the first step (e.g. by arm-raising to initiate play and touching back (by infants to mothers) to request being carried). (This seems to me to be the distant origin of metonymy.) Gestures are learned ontogenetically and are less emotionally charged than vocal signals, but are still basically dyadic and unidirectional. Attention-getters attract the attention of the recipient (e.g. by slapping the ground). These gestures are also not triadic in most cases (i.e. the communicator rarely makes the recipient attend to something else), but it has been observed that chimpanzees raised in some human contexts learn to point. Still this can be considered an extension of their attention-getting gestures, used only to request something, not to give peers some information or to share an interest with others, i.e. performing cooperatively. Although non-human primates can point imperatively, it is considerably difficult for them to understand the intentions of pointing by others. Thus, non-human primate pointing has more restrictions compared with human pointing.

Recent studies have shown that (mainly human-raised) great apes have the cognitive ability to understand that others have intentions and goals (2.4.1). Nevertheless, the reason why primates cannot use language is, Tomasello insists, that they do not have cooperative motives, that is, they cannot even imagine altruistic behavior. Emancipation from individualism gives birth to language.

Chapter 3, “Human Cooperative Communication,” discusses the “psychological infrastructure of shared intentionality within which humans use their natural gestures” (p. 60) and introduces a model of human communication, the Cooperation Model.

Human gestures for communication are divided into two types: pointing and pantomiming. Pointing is used mainly deictically and is similar to ape attention-getters. On the other hand, pantomiming is performed iconically and is similar to ape intention-movements. However, in both cases, there are critical differences between human and ape gestures. While ape attention-getters only serve to make a recipient attend to the source of noises or
touches (i.e. the communicator), human pointing can direct the recipient’s interest toward external targets, which means, as I have mentioned earlier, it functions triadically. Likewise, whereas apes can predict the communicator’s next action by their intention-movements, humans can read the communicator’s intention by their iconic gestures. What causes these big differences between seemingly similar gestures? According to Tomasello, the answer is the Cooperation Model, which humans have and apes lack.

The Cooperation Model comprises two factors. One is the cognitive skill for creating common (conceptual) ground, which is also called the joint attentional frame and “is necessary for the recipient to determine both what the communicator is directing attention to (his referential intention) and why he is doing it (his social intention)” (p. 75). The more information we share, the less expression we need, so much information can be conveyed even with the poorer communication tool of pointing and pantomiming. The other factor is cooperative social motivation. Although other primates, as noted above, communicate with each other simply to try to manipulate individual goals, humans communicate not only to request help or information from others, but to inform others of things that are helpful to *hearers*, or to share feelings or attitudes about something with others. These mutual assumptions of helpfulness allow humans to reason and act collaboratively. This is considered the first step toward language.

Chapter 4, “Ontogenetic Origins,” investigates the emergence of the above-mentioned gestures and cognitive skills in human children’s early development. Of great importance is that, around 12 to 14 months of age, infants can produce and understand the pointing gesture like adults (though not exactly in the same way). This indicates that the infrastructure of cooperative communication operates in the use of this gesture initially around the first birthday. But why can’t they point earlier? Actually, at a few months of age, infants can already hold up their index finger, get adults to do what they want them to (e.g. by crying), and even share emotions with others. These activities, however, are not associated with an understanding on the infant’s part of how they work on intentionally. The informing motive awaits the first birthday, when infants understand others as intentional agents. The emergence of intention-reading enables a child to point for communication. Human infants begin to show such cognitive and social skills as shared intentionality as early as nine months of age, so this phenomenon is called the “nine-month revolution.”

For early pantomiming, skills such as imitation, simulation, symbolic representation, and pretense are required, whereby young children enact *absent*
actions with the understanding of communicative intention (e.g. young children pretend to drink from an empty cup not only for communicative purposes but also for playful purposes, which is one simple way of representing events not in the “here and now” context). In this sense, these gestures are very similar to human language. In fact, they compete with language for the same function. It appears that iconic gestures and language then get compartmentalized gradually during early development, pantomiming being more imagistic and language being more propositional. This is the general process by which human children learn language.

Chapter 5, “Phylogenetic Origins,” gives deep consideration to Tomasello’s evolutionary theory of language. The key concept here again is “cooperation.” Chimpanzees cannot form joint attention and behave collaboratively with others, whereas human infants can. Where do the skills and the motives for cooperation come from? Tomasello proposes (1) mutualism, (2) indirect reciprocity, and (3) cultural group selection. Regarding (1), the individuals who became more helpful and tolerant and less competitive would probably have had a survival advantage for collaborative activities, which would have led to communicative requests for help and compliance with them. Furthermore, those who also helped others outside of mutual collaboration, e.g. by giving information useful for others ((2)), would benefit in the long run, because that activity would bring them the good will of others, which would raise the probability of getting help from others in return before long. Given (1) and (2), it is logical that attempts to increase common ground and solidarity with peers, made by sharing emotions and attitudes with others, would create cultural groups ((3)).

Accordingly, natural gestures no doubt became conventional, occasionally accompanied with vocalizations which were arbitrary at first, but then solidified and became more widespread, as in the process that leads to a dead metaphor. The predominance of vocalization over gestures may have happened for some added reason, such as to communicate at longer distances or in public spaces.

There is still a big problem left unsolved: the emergence of grammar. There are too many factors involved to summarize Tomasello’s explanations and arguments (Chapter 6) in the remaining space, so I will focus on grammar’s phylogenetic origin in light of the three motives. First, let us examine the requesting motive. Some language-trained great apes, which use this motive almost exclusively, can indicate two different aspects in a single re questive situation (e.g. “action + object” as in “EAT CHEESE”), which may well indicate “some first glimmerings on the road to human
syntax” (p. 252). But the differences of sign order do not change the whole meaning and, in most cases, gestures with more than one sign rarely occur together since multiple signs are not needed in “here and now” situations. Considering this, Tomasello presumes that the earliest human beings, Homo, the enhancement of the great apes’ natural gesture sequences, would have come about to partition a situation into events and participants. In all likelihood, at this point, events and participants would have been expressed with pantomiming and pointing, respectively. Chances are that these partitions are the origin of the parts of speech, i.e. verbs and nouns.

Next let us examine the informing motive. As Homo proceeded to inform one another of useful information outside of collaborative activities, the evolved Homo, i.e. Earlier sapiens, would have found it necessary by definition to express things not in the “here and now” context, because helpful information in this case is information the recipient does not know: Who did (or will do) what where and when? To accomplish these rather complicated tasks, they would have had to develop conventional syntactic devices to identify absent or unknown referents and events (e.g. sorts of tense markers), to indicate semantic roles (e.g. word order or case markings), and to clarify the illocutionary act (e.g. imperatives or declaratives). This development would have triggered a fundamental form of syntax.

Last, but not the least, is the sharing motive, which prompted Earlier sapiens into social identification and bonding. What would have helped to form solidarity with community members then would have been narratives, such as myths, folk tales, or parables. To convey such long and coherent stories as a whole, more elaborate syntactic devices (e.g. relative clauses or contrastive word order) would have been created, partly through grammaticalization or entrenchment. This process may well be considered similar to a pidgin developing into a creole, and so the communication devices Later sapiens would have used can be said to resemble the modern languages we Homo sapiens use.

The last chapter (Chapter 7), “From Ape Gestures to Human Language,” itemizes the points discussed thus far, followed by Tomasello’s concluding remarks. He makes many significant suggestions there, but I can say that his major insight is consolidated into the following comment: “Language, or better linguistic communication, is thus not any kind of object, formal or otherwise; rather it is a form of social action constituted by social conventions for achieving social ends, premised on at least some shared understandings and shared purposes among users.” (p. 343). This view of language has high affinity with the symbolic view of grammar cognitive
linguistics endorses and the idea of the “communicative competence” that interactional sociolinguistics adopts. In particular, the verb island hypothesis, the model of language acquisition advocated by Tomasello (though it is not discussed so much in this book), has much in common with Ronald Langacker’s usage-based model. Therefore, it is entirely fair to say that this book provides a psychological basis for these linguistic schools.

Finally, I would like to briefly discuss this book’s contributions to linguistics from the interdisciplinary point of view.

3. Comment

First, it is important to discuss economics for a moment. Neoclassical economics has been in the mainstream for a long time, on the grounds that it apparently follows “scientific” methods. One reason for this is that it removes uncertainties; that is, humans are assumed to be *Homo economicus*, a being that is completely rational, self-interested, and self-contained in economic activities. According to this assumption, humans will always make the utmost effort to get the lowest price on whatever they want, even a single apple, as efficiently as possible in order to maximize economic utility. Of course, in reality, we are not as perfect as that, but it has been widely thought in most modern disciplines that such assumptions are necessary for scientific research.

However, a whole new idea, behavioral economics, has slowly gained acceptance, and its advocates, Daniel Kahneman and Amos Tversky, won the Nobel Prize in 2002 for their work. According to behavioral economics, humans sometimes make logically wrong moves, not because we are foolish, but because it is more efficient in most cases to think by a rule of thumb than to think deductively (for example, we could not have a normal life if we spent hours carefully examining where to buy one apple). In other words, this school regards economic entities as real humans. By doing this, we can investigate a number of phenomena that could not be treated earlier, like the mechanism of the emergence of the economic bubble. (In neoclassical economics, this is not possible, because economic entities are assumed to behave perfectly.)

The reason I have written at some length on economics is that Tomasello uses the example of the ultimatum game in this book (p. 187), a game commonly referred to in behavioral economics as a case example that dramatizes that humans will act not solely in self-interest but rather collaboratively in their community. He does not start the discussion with the perfect humans,
either. The point is that, in several fields, the scientific studies based on warts-and-all humans are beginning to flourish, using the insights of cognitive psychology. Therefore, there is no doubt that this book will play an important role in interdisciplinary studies with the same viewpoint.

Needless to say, the contributions this book makes to linguistics are outstanding. It would cover too broad a range of topics to note them all here, so I will discuss the significance of this book as it relates to one large wave that has recently arisen, especially in the Japanese academic community of linguistics. With a few exceptions, traditional linguistics has attempted a bird’s-eye analysis of the semantic and syntactic structures of human language. Even the term “subjectification,” which Langacker advocates, cannot be said to be free from this spell, especially in the way of description (self, the speaker, is explicitly displayed in his diagram). This may be influenced by the nature of European languages, which prefer to construe an event from the bird’s-eye view, resulting in transitive constructions. In contrast, languages like Japanese are likely to express just what the speaker sees, which leads to intransitive constructions. Ikegami (2003, 2005) has been quick to find this tendency in contrastive studies, and he names the difference “objective construal” and “subjective construal,” respectively. In the same vein, but from the ecological psychological point of view, Honda (2005) analyzes the subjective aspects of not only Japanese but also English, using such ideas as “ecological self” or “affordance.” In particular, his analysis of English middle constructions is highly persuasive, in that this seemingly anomalous construction can be given a natural explanation in parallel with other normal constructions, such as copulative perception verb constructions, as well as other exceptional ones, such as prepositional-subject constructions. Furthermore, from a pragmatic viewpoint, Ide (2006) emphasizes the importance of the logic of “ba (context),” the viewing field of the speaker embedded in the speech event involved. The common point of these studies is that they think of language as an outcome resulting from human interaction in real speech activities, not interaction by highly-idealized human beings. This presupposition is, by all means, true also of Tomasello. Therefore, this book may provide the ontogenetic and phylogenetic basis for this new trend.

Finally, I would like to comment on Tomasello’s statement: “It is not that the evolution of some kind of innate syntactic template such as universal grammar is impossible, it is just that currently there is no evidence for it empirically, no precise formulation of it theoretically, and no need for it at all—if the nature of language is properly understood.” (p. 313).
is safe to say that this view follows the scientific method as seen through “Occam’s razor,” that is, all things being equal, the best hypothesis is the one that makes the fewest assumptions. It must be admitted that the picture Tomasello depicts is (as he points out himself) still a rough sketch, but the same holds for physics. Even the most advanced model cannot explain the birth of the universe as unquestionably as creationism, which is invincible in that it can explain all phenomena by saying “God made it the way it is.” Of course, I do not mean from this comparison that generative theory is a religion. I willingly admit that it follows the scientific method of hypothesis, replication, and disconfirmation. And I also admit that, as of now, we cannot say “all things being equal,” comparing Tomasello’s hypothesis with UG. But it would be a rush (and not be scientific) to think that UG is a given and conclude that the theories that do not rely on it are highly unlikely, considering rapidly evolving fields like brain science and complex systems. It is quite possible that mirror neurons and fractal (or recursive) structure have something to do with language. Although no one knows if this interdisciplinary and bottom-up approach, as presented in Tomasello’s book, can solve Plato’s problem, we cannot completely deny the possibility that UG might face the same fate as did ether, the hypothetical matter which was thought to exist in space. I think it sound science that, in the study of language acquisition, the top-down oriented approach and the bottom-up oriented approach should develop like “two wheels of one cart,” competing with each other. By definition, these two approaches are unlikely to merge and unite, but both will give intriguing insight into the origins of language. Tomasello’s research can be expected to lead the bottom-up approach.

REFERENCES


Ikegami, Yoshihiko (2005) “Gengo ni okeru “Syukansei” to “Syukansei” no Gen-gotekisihyoo (Locutionary Subjectivity and its Linguistic Indices (2)),” *Ninti-
gengogakuronkoo (Studies in Cognitive Linguistics), No. 4, 1–60, Hituzi Syobo, Tokyo.

[received December 19 2009, revised and accepted April 20 2010]

Department of Social Systems
Shiga University
1–1–1 Bamba, Hikone-shi
Shiga 522–8522
e-mail: idehara@biwako.shiga-u.ac.jp