MODALITY SWITCH IN HUMAN LANGUAGE EVOLUTION*

ROLAND MÜHLENBERND
DANKMAR ENKE
MATTHIAS VILLING

Department of Linguistics, Universität Tübingen
roland.muehlenbernd@uni-tuebingen.de
dankmar.enke@uni-tuebingen.de
matthias.villing@googlemail.com

NATALJA GAVRILOV

Department of Animal Physiology, Universität Tübingen
natalja.gavrilov@student.uni-tuebingen.de

JONAS DAVID NICK

Department of Computer Science, Eidgenössische Technische Hochschule Zürich
jnick@student.ethz.ch

Gestural communication plays a major role in social interactions of our closest relatives, the great apes, and therefore quite certainly in social interactions of our own ancestors. In contrast, vocal communication plays a superior role in today's human language use. What kind of factors prompted a potential shift from gestural to predominantly vocal communication? We use computer simulations of an agent-based model of language evolution to examine these potential factors. Our simulation results show that specific interplays of different factors like brain size or group size lead to slower, faster or no emergence of predominately vocal language use.

1. Introduction

What is language? What turned language into what it is? How is it organized? And why is complex language peculiar to humans? Considering fundamental scientific insights into the origin of life, the answers to these questions are comparatively unclear. However, these global issues have experienced a renaissance in recent years given considerable advances and discoveries in cognitive science. Possibly the most important point of concurrence among researchers

*This paper is based on a project within the interdisciplinary Forum Scientiarum Fellows Program for the study of *language and cognition* at the University of Tübingen. Thanks to Miriam Haidle, Gerhard Jäger, and Hans-Peter Mallot for their helpful advice.

is that language evolution research must be cross-disciplinary in order to cope with the complexity of language evolution and provide sufficient constraints on theorizing to make it a legitimate scientific enquiry (cf. Christiansen & Kirby-Christiansen & Kirby2003, Christiansen & KirbyChristiansen & Kirby2003). A strongly debated issue concerns whether human language originated in gestures (cf. ArbibArbib2002, ArbibArbib2002; CorballisCorballis2003, CorballisCorballis2003) or emerged from vocal interactions (cf. DunbarDunbar2003a, Dunbar-Dunbar2003a; Owren, Amoss, & RendallOwren et al..2011, Owren, Amoss, & RendallOwren et al..2011). By taking a look at our closest relatives, the great apes, evidence suggests that while their vocal communication is quite fixed and cognitively hard-wired, their abilities of gestural communication are highly flexible and play a major role in social interactions (cf. Tomasello & ZuberbühlerTomasello & Zuberbühler2002, Tomasello & ZuberbühlerTomasello & Zuberbühler2002). In contrast, vocal communication plays a predominant role in today's human language. Consequently, we should ask i) how gestural and vocal communication interacted in the course of language evolution and ii) what were the potential factors prompting the shift to primarily vocal communication.

2. Biological and Cultural Influences

Along with the question of *how*, many scientists have given increasing credibility to another, equally important question: *when* did our human language begin? A possible, but indirect, reference to the time of origin of human language can be found in the development of the structure and size of our brain. It is widely assumed that a large brain size was essential for the emergence of language. It has also been acknowledged that the human brain has rapidly increased in size over the past two million years (Aiello & DunbarAiello & Dunbar1993, Aiello & DunbarAiello & Dunbar1993). It remains still unclear, however, why the brain has continued to grow significantly. Several theories try to explain this: the *social brain theory* (DunbarDunbar2003b, DunbarDunbar2003b) argues that this growth resulted from an increased number of communication partners. Thus, more processing capacity is needed in order to act in a larger group. This theory is primarily supported by the fact that the increase of group size and the increase of the brain volume occurred in a parallel fashion (Aiello & DunbarAiello & Dunbar1993, Aiello & DunbarAiello & Dunbar1993). Therefore, we integrate both factors - brain and group size - into our model.

Furthermore, more than 2.5 million years ago, the lifestyle of the hominids underwent a fundamental change (Harcourt-Smith & AielloHarcourt-Smith & Aiello2004, Harcourt-Smith & AielloHarcourt-Smith & Aiello2004). Archaeological evidence shows that at this time, small groups developed into even larger groups, natural habitats changed, and first artistic work was made. These findings support either the hypothesis that the ability to learn a language requires both specific cognitive and physical skills or, inversely, that the production of complex tools and artworks requires special abilities. The findings, however, allow conclusions on a minimum degree of cognitive abilities which became available for the human language in its present form (HaidleHaidle2010, HaidleHaidle2010). In addition to the aforementioned biological influences, in our view, the use of tools in particular has had a considerable impact on this development. Knowledge and skills of this kind may have been barely discovered ab ovo within a lifespan, but rather they had to be passed down to subsequent generations. Therefore concepts of learning and teaching were needed. Interestingly, in a situation like this, vocal communication has a significant advantage over gestural communication: while gesturing - when using tools or producing them - is only possible by interrupting the activity, however, you can always simultaneously communicate in a vocal fashion. This also represents a plausible reason for the shift of gestures to facial expressions (CorballisCorballis2010, CorballisCorballis2010).

3. Simulation Model

Hereinafter, we will introduce an empirically motivated model that i.) integrates factors that play an important role in many theories of language evolution, ii.) simulates the reciprocal and dynamic character of language evolution and iii.) allows for the systematic investigation of essential variables that influenced the way human language evolved.

3.1. *Motivation*

Traditional theories about the origin and development of human language often fail to consider empirical facts. This is mainly because these theories are highly abstract and not well compatible with most of these facts. Additionally, gaining information about the language use of our ancestors from such empirical facts can only be indirect, since speech and gestures do not leave any records. Furthermore, the process of the evolution of human language is non-replicable, so it is almost impossible to gain direct insights from e.g. field studies or experiments.[a]

At this point, computer simulations can provide a remedy: a computer program performs simulation runs that can be seen as experiments, with the goal to generate predictions on the basis of hypotheses about causal dependencies and processes of simulated phenomena. In this way, it is finally possible to evaluate the accuracy of the predictions made by different theories.

Computer simulations have been shown to be a powerful tool in research issues concerning the evolution of human language (Cangelosi & ParisiCangelosi & Parisi2002, Cangelosi & ParisiCangelosi & Parisi2002). The starting point is often the so-called agent-based model: an implementation of a model of genetic

[a]Recently, studies of language evolution that conduct *laboratory experiments* became quite popular (cf. Scott-Phillips & KirbyScott-Phillips & Kirby2010, Scott-Phillips & KirbyScott-Phillips & Kirby2010).

Table 1. Essential situations in our ancestors' life.

Situation	Relevance	Involved members	Hours per day	CP	Utility
Tool making	Imitation / learning	All	$0 \rightarrow 1$	1	0.8 / 0.2
Defense	Alarm calls / signs	Women, men	0.5	1	0.9 / 0.1
Hierarchy	Internal fights / debates	Women, men	$0.25 \rightarrow 0.5$	0.2	0.8 / 0.2
Frustration	Problem overcoming	All	0.5	1	0.5 / 0.5
Hunting	Strategic communication	Men	0.5	0.1	0.8 / 0.2
Gathering	Informing / helping	All	4	0.1	0.7 / 0.3
Social play	Other social interactions	All	$2.7 \rightarrow 4$	1	0.6 / 0.4
Sleeping	Communication breaks	All	4	0	-

dispositions, cognitive abilities or communication behavior of language participants. With such a model, it is possible to draw a potential process of language evolution in dependence of evolutionary dynamics and in comparison with assured facts and evidence.

3.2. *The Model*

Our model involves a computer simulation about a group of individuals that have early hominid properties and live in an environment whose characteristics are determined by empirical findings. A multitude of mutual influential and nondeterministic factors have an impact on such a group's life. Thus we reconstruct the life of generations of individuals by nondeterministic computer simulations, called virtual experiments. The basic idea involves the assumption that during the process of the development of human language, individuals communicated essentially in two modalities: gestural and vocal.

The individuals live in groups that consist of children, women, men, and elders. Each subgroup is involved in the group's activities in different ways. During a simulation run, individuals age over time (child $\rightarrow$ man/woman $\rightarrow$ elder), finally die and are replaced by new-born children. The demographical structure of the group is controlled by two control variables: group size and old-youth-ratio. Studies of Aiello and DunbarAiello and Dunbar1993 (Aiello and DunbarAiello and Dunbar1993) provide essential data of our ancestors' group sizes and its change over time. Data of old-youth-ratios are given by Caspari and LeeCaspari and Lee2004 (Caspari and LeeCaspari and Lee2004), based on degrees of wear of jaw bones.

Based on data from primate studies, archeological findings and field research with primitive people are indicative of 8 situations that were essential in our ancestors' daily life. Table 1 shows all 8 situations and for each situation i) its cultural and linguistic relevance ii) which group members are involved, iii) the share per day (excluding night hours), iv) the probability of communication (CP) and v) the utility value for vocal (first value) and gestural (second value) communication.

Note that the hours per day for the situations tool making, hierarchy and social play changes over time. Since individuals very likely needed a larger brain size for tool-making, the appropriate time increases from 0 to 1 hour per day depending on the growth of brain size. Time for social play and hierarchical fight increases with the growth of group size in our simulations (HaidleHaidle2010, HaidleHaidle2010).

The utility values are based on the assumption of how successful each of the modalities is in a given situation.[b] Furthermore, each individual has a fitness value that is determined by the probability of choosing a specific modality in a particular situation and its utility value. Thus more successful individuals (higher utility) have a higher fitness value and a higher number of offspring. This corresponds to the assumption that in the times of early men selection was essentially based on intelligence (DunbarDunbar2003b, DunbarDunbar2003b). New-born offspring are similar to the parents in communication behavior skills (by abstracting from the question if the passing on is of cultural or biological nature). Such a similarity of parents and children is expressed by a so-called *tradition value* that increases with the brain size of our ancestors.

All in all, the dynamics nature of our model includes three independent variables whose values are given by anthropological data: brain size, group size and old-youth-ratio. Furthermore, the probabilities of three situations change over time: tool making (in dependence of brain size), social play and hierarchal fights (in dependence of group size). Further dependent variables are the tradition value (in dependence of brain size), the fitness value of an individual (depends on all variables) and the share of vocal/gestural communication (depends on situations and fitness values). Finally, the utility values are fixed (see table 1) and have a direct influence on the fitness of each individual. The resulting causal network is depicted in Figure 1.

3.3. *Simulations and Results*

Our simulations start with individuals that obtain a niveau of great apes in terms of brain size and/or group size and ends when individuals obtain a niveau of Homo sapiens. During a simulation run, individuals of the group are faced with different communication situations, where they can choose between vocal or gestural communication. Since the starting point of our simulation reconsiders human ancestors that were quite similar to the great apes, we do not assume an initial tendency biased for vocal communication. We expect that during the simulation the modality will shift to a final stage of mainly vocal communication. And indeed, our experiments showed that all simulation runs ended almost exclusively with

[b]E.g. tool making and gathering involves handwork so that the hands are not available for gestural communication. Vocal communication is especially efficient for alarm calls and hierarchy fights since it more easily attracts attention of other members. In addition, for strategic situations like hunting, vocal communication supports the involvement of bigger groups, with the small handicap of probably attracting the attention of the prey. Finally, vocal communication has efficient application in the social play of Homo sapiens and thus gains slightly more utility than gestural communication.

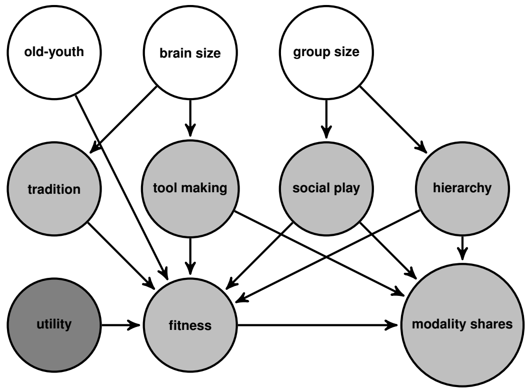

Figure 1. The causal network with its independent variables (white nodes), the fixed utility values (dark gray node) and its dependent variables (gray nodes). The resulting value is the share of modality (share of vocal and gestural communication among the group, both sum up to one).

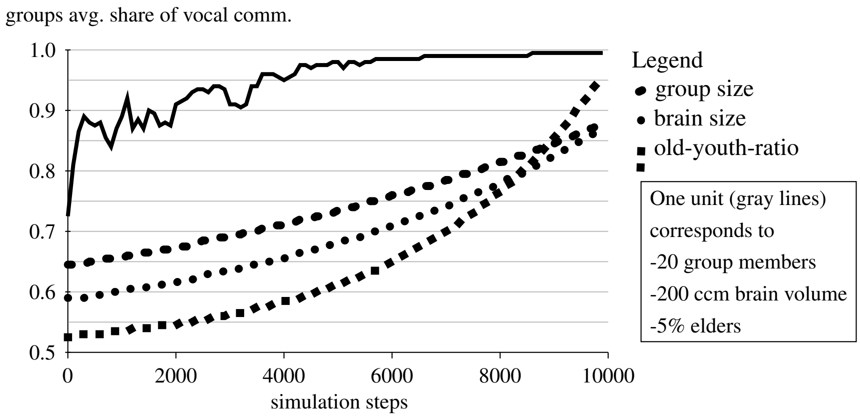

Figure 2. The course of the share of vocal communication over 10000 simulation steps (solid line). The independent variables *group size*, *brain size* and *old-youth-ratio* (in form of the share of elders) increase from initial values of great apes to final values of Homo sapiens.

vocal communication for the case that the three free variables increase according to evidences. Figure 2 depicts the course of the group's average share of vocal communication of a exemplary simulation run for such a case.

To analyze the plausibility and impact of ostensibly very relevant biological and sociocultural factors, we will conduct simulation runs for different settings: with or without the interference of the factor(s) in question: increasing brain size, increasing group size, and increasing old-youth-ration. As we already mentioned

Table 2. 6 experiments with different settings for independent factors and the resulting share of vocal communication averaged over 20 simulation runs each.

Experiment No.	A1	A2	A3	A4	A5	A6
Old-youth-ratio	increasing	increasing	increasing	constant	constant	constant
Group size	increasing	constant	constant	increasing	constant	constant
Brain size	increasing	increasing	constant	increasing	increasing	constant
share of VC	.99	.96	.54	.99	.97	.54

and depicted in Figure 2, with the interference of all three factors, a level of finally solely vocal communication is reached. If the absence of a specific factor does not lead to the expected outcome, we can assume that this factor made a substantial contribution to the evolution of human language.

The appropriate experiments of Experiment series A and its final results are depicted in Table 2. We made 6 different experiments with 20 runs each. Each experiment corresponds to a different setting and the final result depicts the final share of vocal communication, averaged over all simulation runs. The results show the following things: first, the old-youth-ration does not contribute significantly to the shift to vocal communication, since it does not affect the final values (no significant difference in results of A1 vs. A4, A2 vs. A5 and A3 vs. A6). Furthermore, the increase of group size has a small effect on the final outcome (see A1 vs. A2 and A4 vs. A5). However, as expected, the increase of brain size is the crucial factor for the increase of vocal communication. Without it, the share of vocal communication decreases to almost 50%; with it it reaches at least 95%.

In a second Experiment B, we investigated the temporal interaction between the increase of brain size and group size. According to the Social-Brain-Theory (Dunbar, 1998), both factors are dependent on each other and therefore happened in the proximity of time. The experiment showed that if the increase of brain size and group size is set wider apart than a specific value, the share of vocal communication does not increase, but rather decreases to 50%.

4. Conclusion

At the beginning of our study we asked the question: "how could modern human language evolve by taking the communication of our closest related species, the great apes, as a starting point?" By taking the two most efficient communication modalities (vocal and gestural) into consideration, we considered the hypothesis that the human language evolved continuously and by a dynamic interaction of gestural and vocal communication. Furthermore, we assumed that biological and cultural influences played an important role in the evolution of human language as well. And we took three variables into account for which change over time was well-documented by evidence: brain size, group size and old-youth-ratio.These are the independent variables in our model, which we analyzed in their contribution to the emergence of mainly vocal communication as the modern human's language system.

Our simulation results reveal that, first, an increase of the share of vocal communication can only occur if the brain size increases. Note that brain size influences the tradition value, which involves important features of modern human language, such as imitation and learning. Second, without an increase of group size, the increase of the share of vocal communication is slightly lower. This supports the assumption that group enlargements supported or demanded more efficient vocal communication. Third, the increase of old-youth-ratio had no effect on the evolutionary course. Finally, for an increase of the share of vocal communication to evolve, the increase of brain and group size must have happened in the proximity of time.This result supports, to some degree, the social brain theory.

References

Aiello, L., & Dunbar, R. (1993). Neocortex size, group size, and the evolution of language. *Current Anthropology*, *34*(2), 184–193.

Arbib, M. (2002). Grounding the mirror system hypothesis for the evolution of the language-ready brain. In *Simulating the evolution of language* (pp. 229–254). Berlin: Springer.

Cangelosi, M., & Parisi, D. (2002). Computer simulation: A new scientific approach to the study of language evolution. In *Simulating the evolution of language* (pp. 3–28). Berlin: Springer.

Caspari, R., & Lee, S.-H. (2004). Older age becomes common late in human evolution. *PNAS*, *101*(30), 10895–10900.

Christiansen, M., & Kirby, S. (2003). *Language evolution.* Oxford University Press.

Corballis, M. (2003). From hand to mouth: The gestural origins of language. In *Language evolution* (pp. 210–218). Oxford University Press.

Corballis, M. (2010). Mirror neurons and the evolution of language. *Brain and Language*, *112*, 25–35.

Dunbar, R. (2003a). The origin and subsequent evolution of language. In *Language evolution* (pp. 219–234). Oxford: Oxford University Press.

Dunbar, R. (2003b). The social brain: Mind, language, and society. *Evolutionary Perspective Annual Review of Anthropology*, *32*, 163–181.

Haidle, M. N. (2010). Working memory capacity and the evolution of modern cognitive capacities - implications from animal and early human tool use. *Current Anthropology*, *51*(1), 149–166.

Harcourt-Smith, W. E. H., & Aiello, L. C. (2004). Fossils, feet and the evolution of human bipedal locomotion. *Journal of Anatomy*, *204*(5), 403–416.

Owren, M. J., Amoss, R. T., & Rendall, D. (2011). Two organizing principles of vocal production: Implications for nonhuman and human primates. *American Journal of Primatology*, *73*(6), 530–544.

Scott-Phillips, T. C., & Kirby, S. (2010). Language evolution in the laboratory.

Trends in Cognitive Sciences, *14*(9), 411–417.
Tomasello, M., & Zuberbühler, K. (2002). Primate vocal and gestural communication. In *The cognitive animal: Empirical and theoretical perspectives on animal cognition* (pp. 293–299). Cambridge, MA: The MIT Press.