de Boer, Bart (2011) Infant-directed speech and language evolution,
Oxford Handbook of Language Evolution, Oxford University Press, pp. 322–327

Infant-directed speech and language evolution

Infant-directed speech is here defined as a set of speech registers that caretakers use to address infants. There are at least three different kinds of infant-directed speech: one that is used to get the infant's attention, one that is used to soothe the infant, and one that is used to address the infant with linguistically meaningful utterances. All have different characteristics, but there are also similarities, and they could be considered as points on a continuum. Nevertheless, it is unfortunate that in general no difference is made between the different kinds of infant-directed speech in the literature. The different kinds are used in different circumstances, and from an evolutionary point of view, they would make different contributions to the infant's fitness.

All kinds of infant-directed speech are characterized by slower speech rate and larger intonation contours (e.g. Fernald *et al.*, 1989; e.g. Kuhl *et al.*, 1997). Attention-getting infant-directed speech is characterized by higher volume and extreme intonation excursions, but it does not necessarily consist of meaningful utterances. Basically, caretakers use any speech-like means to get the infant's attention. Soothing infant-directed speech is characterized by lower volume, sometimes even whispered speech, and very flowing intonation contours. Lullabies could be considered part of this kind of infant-directed speech.

Although it can be argued that all kinds of infant-directed speech can serve adaptive purposes, for this article, the meaningful kind of infant-directed speech is most interesting. It is characterized by short, syntactically simple utterances that are generally about the direct context, exaggerated intonation contours, lower tempo and clearer articulation.

Caretakers use infant-directed speech automatically when addressing infants, even without being aware of doing so. They also automatically adapt the complexity of their speech to the level of linguistic competence of the infant. In addition, infants prefer to listen to infant-directed speech over adult-directed speech (Fernald, 1985). Infant-directed speech also appears to be nearly universal cross-culturally (Ferguson, 1964; Fernald *et al.*, 1989) and a similar register is attested in signed languages (Masataka, 1992). Whether it really is universal cross-culturally is still an open question. It has been

argued (e. g. Ochs, 1983) that there are cultures in which no special speech registers for infants are said to exist, or where speech in the presence of infants has very different properties from ordinary infant-directed speech, but these cases appear to be exceptional. There are even claims that in some cultures infants are not addressed by adults. However, in these cultures infants do receive linguistic input from older children, as is clear from the original sources (Schieffelin, 1985).

Furthermore, it has been shown by computational experiments that it is easier to acquire vowel categories on the basis of infant-directed speech than on the basis of adult-directed speech (de Boer & Kuhl, 2003; Vallabha *et al.*, 2007). Other computational experiments have shown that infant-directed speech can help to preserve stability of vowel systems over time (de Boer, 2003, 2005). Experiments with infants have shown that infants learn to distinguish speech categories from a second language when they are exposed to caretaker-child interactions in that language, but not when they are exposed to (adult-directed) speech from a television (Kuhl *et al.*, 2003).

When a behavior such as infant-directed speech is found to be almost universal across cultures, when it is engaged in automatically and when it contributes to fitness, it becomes worthwhile to investigate how it could have evolved and how it could have influenced the evolution of language in general. Different theories about the adaptive purpose of infant-directed speech and its role in the evolution of language have been proposed. It has been proposed that infant-directed speech serves an important role in acquiring speech and language, and that it is therefore part of what must be explained by a theory of language evolution. It has also been proposed that honest signaling, which is a necessary prerequisite for language to evolve, can most easily be explained as emerging from mother-child interactions in combination with kin selection (Fitch, 2004). Finally, it has been proposed that language itself has emerged from mother-child interactions that have become more and more complex over time (Falk, 2004; Locke, 2006; Locke & Bogin, 2006). Here the focus will be on the adaptive role of infant-directed speech, rather than on its role in the emergence of language. Falk and Locke's theories are explained in separate chapters [[insert reference here]] while the role of kin selection will be discussed only briefly here [[is there perhaps a chapter on types of selection, or on kin selection

itself to which I could refer here?]]. For a critique of these theories, see (Tallermann, 2005).

The adaptive value of ID-speech

The properties of infant-directed speech, and more specifically, the kind that is used for linguistically meaningful utterances, facilitate acquisition of speech. Slower speed and more exaggerated intonation can help to divide the speech stream into its building blocks syntactically, lexically and phonemically. Exaggerated articulation helps to separate the different categories of speech sounds. Using only simple, contextualized utterances also helps to facilitate acquisition of lexical and grammatical knowledge. It can be argued that an infant is not alone in the task of acquiring a language. The behavior of caregivers helps in this task, and should therefore be taken into account when investigating the evolution of language.

A fact that is often overlooked when discussing the evolution of language is that the behavior of the caregivers themselves needs an evolutionary explanation. Producing simplified language is not necessarily as simple as it might seem. It requires (subconscious) knowledge of the language and an ability to adapt one's utterances to the level of knowledge of the interlocutor.

An ability to adapt one's utterances to the level of knowledge of the interlocutor is something that is also needed in adult-to-adult conversation. In the case of pragmatics this is well-known, but it is also found in other aspects of language, such as is evident when talking to foreigners or to the hard-of-hearing. Being able to produce infant-directed speech does therefore not necessarily require a lot of extra adaptations than are needed for producing adult-directed speech. However, the fact that caregivers tend to use infant-directed speech automatically and almost without being aware of it indicates that some specialized mechanisms have evolved for adult speech to infants.

Given that linguistic input to infants is adapted to their linguistic abilities, the task of language acquisition becomes much easier for infants. Issues about the difficulties of acquisition that are sometimes raised on the basis of examples from adult-directed written language are most likely not relevant to acquisition of spoken language by infants. Complex syntactical constructions that would be impossible to acquire if presented to the

infant directly are easier to acquire if complex examples are presented only when the less complex ones have been mastered. It might as a result turn out that the cognitive machinery that is necessary for acquiring language can be much simpler than is sometimes assumed. Consequently there would have to be fewer biologically evolved and genetically coded adaptations in the brain for acquiring language. Potentially, co-evolution of infants' learning behavior and caregivers' teaching behavior is a more plausible evolutionary scenario than the isolated evolution of infant language acquisition would be (but this is still a topic for research). Part of this hypothesis is testable by investigating whether infant-directed speech really is more learnable than adult-directed speech (e. g. de Boer & Kuhl, 2003; Kuhl et al., 2003; Vallabha et al., 2007).

There are also indications (e. g. de Boer, 2003, 2005) that in a system where caregivers use infant-directed speech, more complex linguistic systems remain stable over a longer time. Even though an infant would in principle be able to learn such a complex system without infant-directed speech, one would expect so many learning errors, and therefore variation in the population, that after a few generations the system simplifies. This is an instance of the fact that the class of learnable human languages is not the same as the class of human languages that can be stably transmitted over multiple generations - (Brighton *et al.*, 2005). In the presence of infant-directed speech however, learning errors and variation diminish, and therefore more complex systems can remain stable. The presence of more complex, culturally transmitted linguistic systems then makes biological, genetic adaptations to the existing complexity useful. In this way, the cultural presence of more complex linguistic systems can drive biological evolution of the language faculty. Thus infant-directed speech can help to drive biological evolution.

This is of course something that is very difficult if not impossible to investigate directly in real human language. The results that are referred to here have been achieved with agent-based computer models, based on the iterated learning paradigm (K Smith *et al.*, 2003). In this paradigm, small computer programs (agents) model essential aspects of human behavior. There is one parent agent and one child agent. After a number of linguistic interactions, the parent agent is removed, the child agent becomes the parent, and a new (empty) child agent is generated. In the case of (de Boer, 2003, 2005) the linguistic interactions consist of the parent producing vowels, and the child learning

them. The very first parent is initialized with a given 5- or 7- vowel system. The system is run for a number of generations and it is then checked how well the structure of the original vowel system is preserved. In one condition, the infant agent has an innate mechanism to compensate for reduced speech, in the other it does not have such a mechanism, but the adult agent produces exaggerated examples of the vowels (comparable to the ones in real infant-directed speech). It is found that the five-vowel system remains equally stable in both conditions, but that for the 7-vowel systems both some kind of innate compensation and infant-directed speech are necessary. Although this project barely scratches the surface of the relevance of infant-directed speech to language evolution, it nevertheless provides an example of how computer models could be constructed and investigated. Experiments with (adult) human subjects are also conceivable. It is possible to create an experiment along the lines of the ones reviewed in (Kenny Smith *et al.*, 2008) in which in one condition no teacher-learner adaptation is done, while in another condition teachers must adapt to the level of the learners. However, as far as I am aware, no such experiments have been done, yet.

Kin selection and ID-speech

A rather different perspective on caretaker-child interactions and the evolution of speech is taken by Fitch (Fitch, 2004). He observes that language is a system that requires cooperation because most of the information that is exchanged is honest. He also observes that speech is a relatively cheap means of communication: the metabolic cost of producing a speech signal is low. He argues that such honest signaling systems with cheap signals are evolutionarily problematic. Signaling information to someone increases the receiver's fitness, but not necessarily that of the sender. The temptation to give false information would then be big, but if signals are unreliable receivers would evolve to ignore them. Honest signaling systems with cheap signals assume some degree of altruism between sender and receiver. Altruistic systems can evolve in two situations: if the altruism is reciprocal (that is, an individual can expect to benefit in return for its altruistic behavior) or if an individual behaves altruistically towards its kin. According to Fitch reciprocal altruism is rare and problematic in biology, whereas altruism towards kin

is common. He therefore proposes kin selection as an evolutionary mechanism through which language can have started.

Fitch has proposed that mother-child interaction is at the basis of language, and has called this hypothesis the mother tongue hypothesis. A mother shares half of her genes with her offspring, and therefore any gain in fitness of the offspring (that is not deleterious to the mother's own fitness) helps to spread the mother's genes. It is therefore in the mother's advantage to communicate to her offspring as much as possible of her knowledge and experience about survival. For this reason a communication system could have developed between mother and child, growing ever more complex as the ability to transfer more and more complex information would increase the fitness of the offspring. In a later evolutionary stage, such communication systems could then be used to communicate with less-closely related kin and with non-kin.

Using the hypothesis that kin selection played a role in starting language evolution Fitch derives a number of properties of language and compares these to the properties of language that would be expected if sexual selection (as proposed by other researchers) played a role. Kin selection correctly predicts that language is acquired young and that women are better at language than men. Sexual selection would predict an onset of language in puberty and also predicts that men would be better at production than women. Kin selection can also help to explain why language is learnt more accurately than is necessary for transfer of purely semantic information, and that the ability to learn language this accurately is lost before the age at which humans generally leave the place where they grew up. In this way linguistic accents can serve as reliable markers of kinship: speakers with the same accent are likely to be more closely related than speakers with different accents. This could also explain why people tend to be more favorably disposed towards people with the same accent than towards people with a different accent.

The theory of kin selection therefore gives a central role to mother-child interactions in the evolution of language. It should be noted, however that Fitch only proposes kin selection as a selection mechanism, and does not provide a detailed scenario of how complex language would evolve, and why it would not have evolved in other species.

Furthermore, as Tallerman (2005) has pointed out, Fitch does not provide an explanation of how language would spread beyond the use in immediate kin interactions.

Conclusion

Caregiver-child interactions and infant-directed speech can have played different roles in the evolution of language. Falk's and Locke's scenarios are presented elsewhere in this volume [[is there a way to give a more precise reference?]] As for kin selection, it could have facilitated the emergence of a cooperative system with cheap signals in the context of caretaker-child interaction. It is also possible that kin selection allowed this communication system to become more symbolic and more complex in nature. Once complex language was present, infant-directed speech could have helped to make language easier to learn, and for complex language to remain stable for longer periods of time. This again could have driven evolution for more adaptations to complex language. It must be said however, that in none of these scenarios there is a compelling argument of how human language became semantically complex and symbolic in nature, while other primates' communication systems did not. Although caretaker-child interaction might therefore have been an important factor in language evolution, other factors (and likely historical coincidence as well) must also have played important roles.

Referenties

Brighton, H., Kirby, S., & Smith, K. (2005). Cultural selection for learnability: Three principles underlying the view that language adapts to be learnable. In M. Tallermann (Ed.), *Language origins: Perspectives on evolution* (pp. 291–333). Oxford: Oxford University Press.

de Boer, B. (2003). Conditions for stable vowel systems in a population. In W. Banzhaf, T. Christaller, P. Dittrich, J. T. Kim & J. Ziegler (Eds.), *Advances in artificial life, lecture notes in computer science 2801* (pp. 415–424). Berlin: Springer.

de Boer, B. (2005). Infant directed speech and the evolution of language. In M. Tallermann (Ed.), *Evolutionary prerequisites for language* (pp. 100–121). Oxford: Oxford University Press.

de Boer, B., & Kuhl, P. K. (2003). Investigating the role of infant-directed speech with a computer model. *Acoustics Research Letters Online, 4*(4), 129–134.

Falk, D. (2004). Prelinguistic evolution in early hominins: Whence motherese? *Behavioral and Brain Sciences, 27*(4), 491-541.

Ferguson, C. A. (1964). Baby talk in six languages. *American Anthropologist, 66*(6, part 2), 103–114.

Fernald, A. (1985). Four month-old infants prefer to listen to motherese. *Infant Behavior and Development, 8*, 181–195.

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language, 16*, 477–501.

Fitch, W. T. (2004). Kin selection and "mother tongues": A negelected component in language evolution. In D. K. Oller & U. Griebel (Eds.), *Evolution of communication systems: A comparative approach* (pp. 275–296). Cambridge, MA: MIT Press.

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevikova, E. V., Rysinka, V. L., et al. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science, 277*, 684-686.

Kuhl, P. K., Tsao, F. M., & Liu, H. M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences, 100*(15), 9096-9101.

Locke, J. L. (2006). Crying, cooing, babbling, and the evolution of language. *Human Nature, 17*(2), 155-168.

Locke, J. L., & Bogin, B. (2006). Language and life history: A new perspective on the development and evolution of human language. *Behavioral and Brain Sciences, 29*(3), 259-325.

Masataka, N. (1992). Motherese in a signed language. *Infant Behavior and Development, 15*(4), 453-460.

Ochs, E. (1983). Cultural dimensions of language acquisition. In E. Ochs & B. B. Schieffelin (Eds.), *Acquiring conversational competence* (pp. 185-191). London: Routledge & Kegan Paul.

Schieffelin, B. B. (1985). The acquisition of kaluli. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition* (Vol. 1, pp. 525–593). Hillsdale (NJ): Erlbaum.

Smith, K., Kalish, M. L., Griffiths, T. L., & Lewandowsky, S. (2008). Introduction. Cultural transmission and the evolution of human behaviour. *Philosophical Transactions of the Royal Society of London, 363*(1509), 3469–3476.

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

Tallermann, M. (2005). Kin selection and linguistic complexity. In A. D. M. Smith, K. Smith & R. Ferrer i Cancho (Eds.), *The evolution of language* (pp. 307–314). Singapore: World Scientific.

Vallabha, G. K., McClelland, J. L., Pons, F., Werker, J. F., & Amano, S. (2007). Unsupervised learning of vowel categories from infant-directed speech. *Proceedings of the National Academy of Sciences, 104*(33), 13273-13278.