[redacted]

How do humans acquire universal linguistic features? Are the same principles and processes that enable the acquisition of universal linguistic features found in other cognitive domains such as causal reasoning?

Introduction

In the following essay I analyse the principles and mechanisms that underlie the acquisition of linguistic universal features and the development of causal reasoning. First, I illustrate how language acquisition is constrained by domain general principles such as inferential learning. Inductive biases present in inferential learning are then magnified through the process of cultural transmission, which enables the emergence of universal linguistic features. Second, I show that inferential learning also underlies the acquisition of causal reasoning. However, causal reasoning is governed by additional domain specific constraints that operate within so called intuitive theories/schemas. In the last part I argue that causal reasoning and language comprehension play different functions, therefore the mechanisms that allow for their complete development also differ. Language is primarily used as a means of navigating in changing social interactions whereas causal reasoning enables accurate prediction of the constant features in the natural world.

Language as shaped by domain general constraints

The question of the evolution of universal linguistic principles is a matter of a heated debate between several different approaches. Adaptationists hold that universal grammar is a result of the brain mechanisms that are specific to language acquisition and have evolved over a long period of natural selection (e.g. Pinker and Bloom, 1990). Non-adaptationists claim that specialized brain mechanisms have not developed through adaptation but rather through some other route (e.g. Gould, 1993). For the purposes of this essay I will take a third view. I will argue that the language has evolved through the process of cultural evolution to be easy to comprehend and understand. Language is easy to comprehend for us because it has adapted to our brains not the other way around (Christiansen and Chater, 2008). Namely, language was shaped by the combination of domain general mechanisms, which underlie the development of all cognitive domains. These mechanisms are perceptual-motor factors, cognitive limitations on learning and processing, constraints from thought and pragmatic constraints (Chater and Christiansen, 2008).

Cultural evolution emphasizes weak inductive biases

Human language is primarily used as a means of communicating and coordinating in complex, ever changing societies. Thus, emergent features of language do not result merely from individual's innate learning principles but are also largely shaped by social practices and hence cultural transmission. More specifically, cultural transmission results in the emergence of linguistic universals through the process of iterated learning (Kirby, Dowman, and Griffiths, 2007). The latter can be described as a process in which individual's linguistic behaviour results from observing the linguistic behaviour of other language learners who acquired their behaviour in the same process (Kirby, Griffith and Smith, 2014).

In the following, I present one way of studying inferential learning and cultural transmission that underlie language acquisition. Bayesian inference is a procedure for modelling the way an individual combines learning biases with the incoming data when evaluating hypotheses about a language (Kalish, Griffith and Lewandowsky, 2007). Weak internal inductive biases act as innate constraints on learning and memory in a way that they make one language easier to learn than another (Kirby et al., 2007). Learner's biases form a probability distribution, which represents a *prior* distribution over languages. A prior distribution is composed of learner's degrees of belief in each hypothesis before receiving linguistic data. After seeing the data the learner calculates the *posterior* probability of each language. Then, a learner's task is to select a language based on the posterior probability distribution. One possibility of selecting a hypothesis is that learners choose arbitrarily between hypotheses that have maximum posterior probability (Kirby et al., 2007). Then the learner generates data for the next learners in the transmission chain.

Interestingly, Kirby et al. (2007) have shown that other factors besides learner's biases may influence the distribution of languages. The amount of data learners obtain, for example, regulates how much the prior preference for particular language is emphasized. When learners obtain limited linguistic input their prior preferences are magnified in the final distribution of languages. This shows that weak biases can have a very strong effect. In other words, in a series of iterated learning steps weak inductive biases are emphasized through the process of cultural transmission. This enables weak biases to be transformed into strong linguistic features without the presence of strong innate constraints (Kirby et al., 2007).

Other cognitive domains

Above, I presented one way in which humans deal with the indeterminacy of the incoming linguistic data through the employment of domain general inferential learning. Inferential learning also plays a crucial role in the development of other cognitive domains, such as reasoning about causal relationships, learning about unobserved properties and categorization (Tenenbaum, Griffiths and Kemp, 2006). However, unlike language comprehension, causal reasoning is also governed by domain specific principles. In the following I first, briefly explain what these principles are and second, explore one possible explanation of functional differences between domain general and domain specific learning principles.

Kemp et al. (2010) present a computational model to examine principles of causal learning in humans. When children start causally structuring the world they employ domain general principles such as inferential learning, which is also the principle underlying the acquisition of linguistic universals. Additionally, causal structuring is made easier because inferential tasks are specific for families of related problems (Kemp et al., 2010). After children have solved several inferential problems from the same family, they can induce a schema that facilitates learning of new causal structures within one domain in the future. Schema or intuitive theory contains a set of domain specific abstract principles, structural constraints and concepts that describe the structure of domain specific problems (Carey, 1985). The role of the schema is to integrate the incoming data and subsequently discover specific principles that hold for one domain. This enables the succeeding sparsely observed features to be generalized based on the hypotheses spaces constrained by the schema (Tenenbaum et al., 2006).

C-Induction and N-induction

Next, I will address the following emerging question. Why is it that the language acquisition is constrained by weak domain general mechanisms emphasized by the cultural transmission, whereas causal reasoning, once developed, is constrained by strong domain specific principles?

Chater and Christiansen (2010) argue that human development involves solving two different but interrelated kinds of problems: ones that are concerned with understanding the *natural* world and others that are connected to the *social coordination*. Solving both kinds of challenges depends on inductive learning, however Chater and Christiansen (2010) differentiate between *two* types of induction, natural (N-induction) and cultural (C-induction), which have different functions.

The function of N-induction is to accurately predict occurrences in the natural world, where the relationships between the events in the natural world are pretty stable. Therefore, in the early stages of causally structuring the world learners rely on inductive biases only until more reliable domain theories that govern their understanding of the natural statistical regularities are constructed (Tenenbaum et al., 2006). In other words, the hypothetical space of causal reasoning is delineated by individual's theory based induction that goes beyond the learner's direct interaction with the world and is thus independent of the social agreement (Tenenbaum et al., 2006).

In contrast to the natural world, human behaviour and social interactions are much less predictable and stable. Thus the function of the C-induction is to enable successful navigation in the social world that depends on learning from others and agreeing on different socially important matters, such as linguistic structure (Chater and Christiansen, 2010). This implies that strong domain specific constraints would in fact hinder immense linguistic flexibility that is sensitive to subtle changes in the social agreement. Following from that, the range of possible universal linguistic features is constrained by the continuous social negotiation through a series of generations where it is not crucial *which* phonological, syntactic or semantic regularities children favour when presented with linguistic stimulus. What is more important is that they converge on the *same* linguistic regularities on populational and, furthermore, on generational level (Chater and Christiansen, 2010).

All in all, in C-induction sharing the same arbitrary bias confers an advantage on those who share it, irrespective of what that bias actually is, whereas in N-type cases, there is some independent criteria (namely a fit with what the external world is actually like) that inferences from that bias have to match.

Conclusion

To summarize, I have shown that linguistic universals are developed through the interaction of domain general learning constraints, such as weak inductive biases that underlie inferential learning. These weak biases are then emphasized in the process of cultural transmission, which results in the emergence of strong linguistic universal features. The same principles and mechanisms of inferential learning are also involved in the acquisition of other cognitive domains, such as causal reasoning. However, in contrast to language, causal reasoning develops due to operation of both, domain general and domain specific constraints. One possible explanation of this difference could be that language and causal reasoning address very different types of tasks.

References

Carey, S. (2009). *The origin of concepts.* Oxford: Oxford University Press.

Chater, N. and Christiansen, M. H. (2010). Language acquisition meets language evolution. *Cognitive Science*, *34*, 1131-1157.

Christiansen, M. H. and Chater, N. (2008). Language as shaped by the brain. *Behavioural and Brain Sciences*, *21*, 489-558.

Gould, S. J. (1993). *Eight little piggies: Reflections in natural history*. New York, NY: Norton Paperback.

Griffiths, T. L., Kalish, M. L. and Lewandowsky, S. (2008). Theoretical and empirical evidence for the impact of inductive biases on cultural evolution. *Philosophical Transactions of the Royal Society of London, 263*, 3503-3514.

Kalish, M. L., Griffiths, T. and Lewandowsky, S. (2007). Iterated learning: Intergenerational knowledge transmission reveals inductive biases. *Psychonomic Bulletin & Review*, *14*(2), 288-294.

Kemp, C., Goodman, N. D. and Tenenbaum, J. B. (2010). Learning to learn causal models. *Cogntive Science, 34*(7), 1-59.

Kirby, S., Dowman, M. and Griffiths, T. L. (2007). Innateness and culture in the evolution of language. *Proceedings of the National Academy of Sciences, 104*(12), 5241-5245.

Kirby, S., Griffiths, T. and Smith, K. (2014). Theoretical and empirical evidence for the impact of inductive biases on cultural evolution. *Current Opinion in Neurobiology, 28*, 108-114.

Pinker, S. and Bloom, P. (1990). Natural language and natural selection. *Brain and Behavioural Sciences, 13*, 707–727.

Tenenbaum, J. B., Griffiths, T. L., Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *TRENDS in Cognitive Sciences*, *10*(7), 309-318.

[redacted]

GRADEMARK REPORT

FINAL GRADE

78/100

GENERAL COMMENTS

Instructor

I though this essay was very good - you have clearly read a bunch of relevant work, including some stuff that is relatively distant from the set readings, and have thought hard about how to fit these ideas together into a coherent story (which I find plausible). The essay is generally very clear, and the structure is excellent and always completely transparent what you are doing. However, the essay is not without its flaws. You are never 100% clear on the role of innate vs acquired constraints in causal reasoning - sometimes it sounds like you are saying there are both, but initially you seemed to be emphasising the role of the latter. Secondly, on the N- vs C-induction stuff, it's not clear why arbitrary, shared, strong biases wouldn't be ideal for C-induction - you switch from talking about specificity and strength of innate biases to their arbitrariness, and the link isn't really explained. All in all though, a very well-worked essay.

PAGE 1

Comment 1

Excellent summary

Comment 2

Both rather informally defined - is comprehend the same as 'learn'? - but generally good here, clear.

PAGE 2

Comment 3

This section is a very clear and somewhat detailed summary of this work, well done. My only concern at this point is that you are running out of space for your central point.

PAGE 3

Comment 4

This is nice, but I feel that you are lacking a couple of concrete examples - that would take more space, which you would need to save from the section on weak biases, bayesian iterated learning etc.

Comment 5

Although from your previous section it sounds like these strong biases are / might be learned?

Comment 6

OK, this suggests that it's both innate and acquired constraints but it's a little frustrating that it's vague still.

PAGE 4

Comment 7

Although this doesn't speak to the *strength* of those biases - wouldn't a shared, arbitrary, strong bias be great for C-induction problems?

PAGE 5