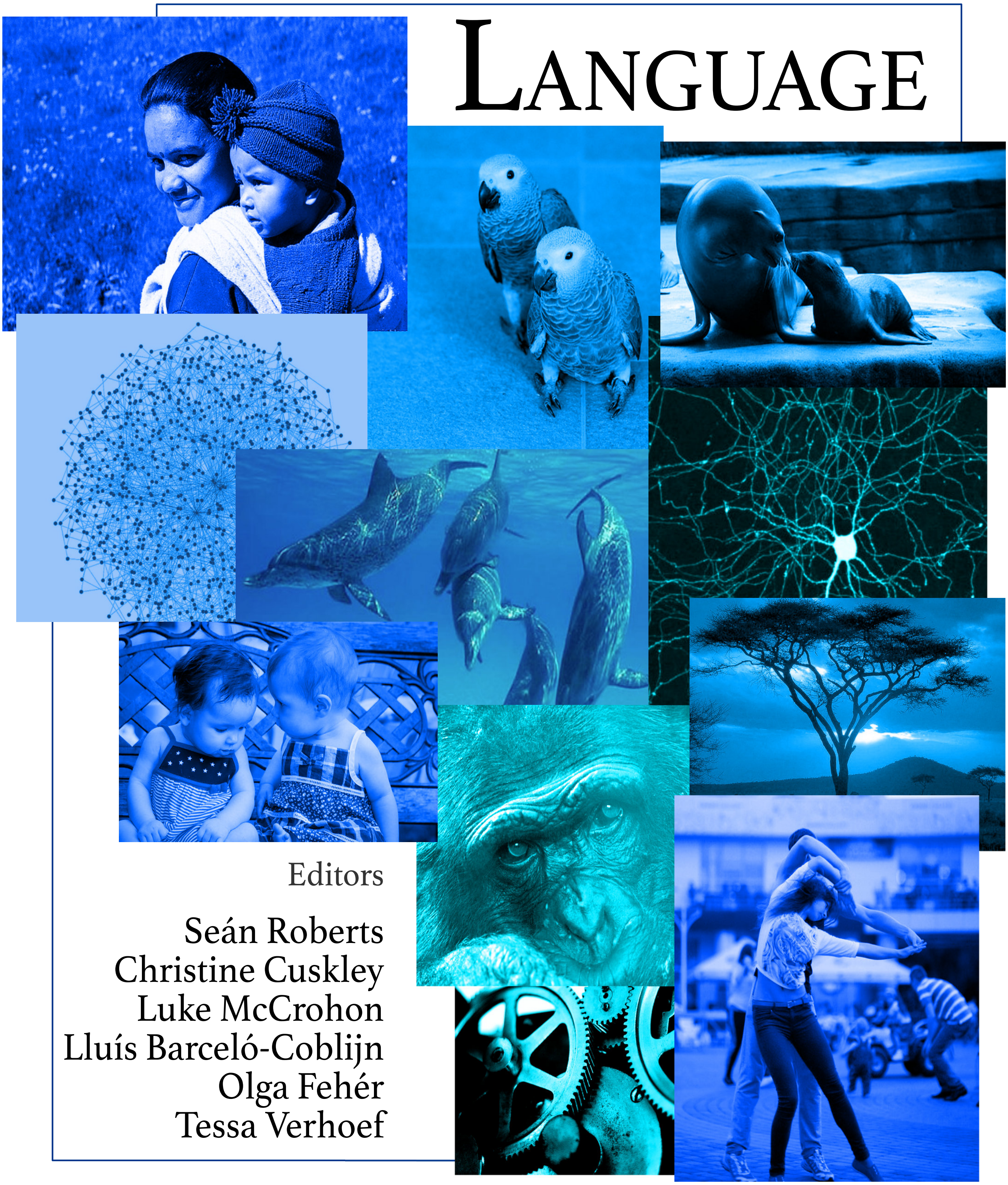
The Evolution of Language

Proceedings of the 11th International Conference
New Orleans, LA
March 21-24 2016



Editors

Seán Roberts
Christine Cuskley
Luke McCrohon
Lluís Barceló-Coblijn
Olga Fehér
Tessa Verhoef

The Evolution of Language: Proceedings of the 11th International Conference. New Orleans, USA.

Cover image credits: https://www.flickr.com/photos/sjcockell/4684828794 : Simon Cockell

https://www.flickr.com/photos/69221340@N04/8652893210 : Papooga

https://www.flickr.com/photos/jdebberly/2850385433 : Jay Ebberly

https://www.flickr.com/photos/pustovit/14336071644 : Vladimir Pustovit

https://www.flickr.com/photos/kittysfotos/10037699755 : Kitty Terwolbeck

https://www.flickr.com/photos/ajmat/3730694413 : AJ.Mat

https://www.flickr.com/photos/neilspicys/2348972925 : NeilsPhotography

https://www.flickr.com/photos/kapten/1085243998 : Björn Söderqvist

https://www.flickr.com/photos/thomasclaveirole/463202335 : Thomas Claveirole

https://www.flickr.com/photos/deanwissing/3565255880 : Dean Wissing

https://www.flickr.com/photos/mikeblogs/3101400087 : Mike Seyfang

Preface

This volume collects the refereed papers and abstracts of the 11th International Conference on the Evolution of Language (EVOLANG XI), held in New Orleans on 21st-24th March, 2016. Submissions to the conference were solicited in two forms, papers and abstracts, and this is reflected in the structure of this volume.
The biennial EVOLANG conference is characterised by an invigorating, multi-disciplinary approach to the origins and evolution of human language, and brings together researchers from many fields including anthropology, archaeology, artificial life, biology, cognitive science, computer science, ethology, genetics, linguistics, neuroscience, palaeontology, primatology, psychology and statistical physics. The multi-disciplinary nature of the field makes the refereeing process for EVOLANG very challenging, and we are indebted to our panel of reviewers for their very conscientious and valuable efforts.
For the first time, the proceedings of EvoLang XI are primarily available online in an open access format. Please visit `http://evolang.org/neworleans/` for up-to-date papers, workshop papers and supplementary materials. Thanks are due to the following people:

The EVOLANG committee:

- Andreas Baumann (University of Vienna),
- Rudolf Botha (Stellenbosch University),
- Christine Cuskley (Institute for Scientific Interchange),
- Erica Cartmill (University of California, Los Angeles) [First Contact/Co-chair],
- Jean-Louis Dessalles (Ecole Nationale Suprieure des Tèlècommunications, Paris),
- Ramon Ferrer i Cancho (Polytechnic University of Barcelona),
- Tecumseh Fitch (University of Vienna),

- Jim Hurford (University of Edinburgh),
- Simon Kirby (University of Edinburgh) [Co-chair],
- Chris Knight (University of East London),
- Heidi Lyn (University of Southern Mississippi),
- Luke McCrohon (University of Tokyo) [Webmaster],
- Kazuo Okanoya (University of Tokyo),
- Thom Scott-Phillips (Durham University),
- Andrea Ravignani (University of Vienna, University of Edinburgh),
- Nikolaus Ritt (University of Vienna),
- Kenny Smith (University of Edinburgh) [Treasurer],
- Maggie Tallerman (Newcastle University, UK),
- Natalie Uomini (University of Liverpool).

The local organising committee:
University of Southern Mississippi: Heidi Lyn (Chair), Megan Broadway, Mystera Samuelson, Jennie Christopher, and Beatrice Chenkin.
University of South Alabama: Stephanie Jett.
Tulane University: Olanike-Ola Orie.
Louisiana State University: XiaoChing Li

The workshop convenors: Christine Cuskley, Vittorio Loreto, Francesca Tria, Vito D.P. Servedio, Andrea Ravignani, Erin E. Hannon, Seán Roberts, Gregory J. Mills, Madza Farias-Virgens, Dillon Niederhut, Terrence Deacon, Tecumseh Fitch, Bart de Boer and Takeshi Nishimura.

The plenary speakers: Sharon Thompson-Schill, Thom Scott-Phillips, Ljiljana Progovac, Richard Moore, Erich Jarvis, Vincent Janik, Evelina Fedorenko, Dean Falk, Joan Bybee.

Finally, and most importantly, the authors of all the contributions collected here.

Seán Roberts, Christine Cuskley, Luke McCrohon,
Lluís Barceló-Coblijn, Olga Feher and Tessa Verhoef.
Scientific Committee,
February 2016

Reviewers

Alan Nielsen
Albert Naccache
Aleksandrs Berdicevskis
Alex Carstensen
Alfredo Ardila
Andrea Ravignani
Andreas Baumann
Andreea Geambasu
Andrew Feeney
Andrew Wedel
Anne Marie Di Sciullo
Anne van der Kant
Anneliese Kuhle
Antoni Gomila
Antonio Benítez-Burraco
Aritz Irurtzun
Ashley Micklos
Barbora Skarabela
Bart de Boer
Bill Thompson
Bodo Winter
Brady Clark
Camilla Power
Carel ten Cate
Carl Vogel
Carlos Santana
Carmen Saldana
Catriona Silvey
Cedric Boeckx
Chris Knight
Christian Abry
Christina Behme
Christine Caldwell
Christine Cuskley
Connie de Vos
Cory M Cuthbertson
Damian Blasi
Dan Dediu
Daniel Dor
Daniel Lenzen
Dankmar Enke
David Leavens
David Lightfoot
Dillon Niederhut
Elizabeth Irvine
Erica Cartmill
Estefania Santacreu-Vasut
Federica Cavicchio
Freek Van de Velde
Fusa Katada
Gareth Roberts
Gary Lupyan
Gerhard Jaeger
Greg Bryant
Gregory Mills
Hannah Little
Harald Hammarström
Heidi Lyn
Hope Morgan

Irene Pepperberg
Jérôme Michaud
James Kirby
James Steele
James Winters
Jasmeen Kanwal
Jean-Louis Dessalles
Jennifer Culbertson
Jeremy Collins
Joan Bybee
Joana Rosselló
Jon Carr
Jordi Arranz
Juan Carlos Moreno Cabrera
Kang-Suk Byun
Katie Slocombe
Kazuo Okanoya
Keelin Murray
Kenny Smith
Kensy Cooperrider
Kenta Suzuki
Kevin Stadler
Klaas Seinhorst
Kobin Kendrick
Koji Fujita
Ljiljana Progovac
Lluís Barceló-Coblijn
Luke Fleming
Luke McCrohon
Madza Farias-Virgens
Maggie Tallerman
Marcus Perlman
Marie Coppola
Marie Montant
Marieke Schouwstra
Marieke Woensdregt
Marion Laporte
Mark Atkinson
Mark Bartlett
Mark Dingemanse
Mark Ellison
Maryia Fedzechkina
Matt Cartmill
Matt Spike
Matthew Lou-Magnuson
Maurício Martins
Mauricio Martins
Michael Arbib
Michael Franke
Michael Pleyer
Michael Spranger
Michelle Spierings
Mieko Ogura
Molly Flaherty
Molly Lewis
Monica Tamariz
Natalie Uomini
Nicolas Fay
Nikolas Gisborne
Nikolaus Ritt
Olga Feher
Olga Vasileva
Olivier Morin
Paul Vogt
Peeter Tinits
Phillip Alday
Phillip M. Alday
Przemysław Żywiczyński
Ramon Ferrer I Cancho
Remi Van Trijp
Richard Blythe
Richard Moore
Rick Janssen
Rie Asano
Robert Berwick
Robert Truswell
Roland Mühlenbernd
Rory Turnbull

Rose Stamp	Sergio Balari	Tecumseh Fitch
Roslyn Frank	Simon D. Levy	Ted Bayne
Ruth Sonnweber	Simon Kirby	Teresa Bejarano
Ryan Lepic	Simon Levy	Tessa Verhoef
Sławomir Wacewicz	Simon Townsend	Thom Scott-Phillips
Sabine van der Ham	Sonia Ragir	Till Bergmann
Sabrina Engesser	Stefan Hartmann	Timo Ahlers
Samarth Swarup	Takashi Hashimoto	Tom Sawallis
Savithry Namboodiripad	Takeshi Nishimura	Vanessa Ferdinand
Seán Roberts	Tao Gong	Willem Zuidema

Contents

Papers

DEICTIC TOOLS CAN LIMIT THE EMERGENCE OF REFERENTIAL SYMBOL SYSTEMS

ELIZABETH IRVINE
School of English, Communication and Philosophy, Cardiff University
Cardiff, UK
irvinee@cardiff.ac.uk

SEÁN G. ROBERTS
Language and Cognition department, Max Planck Institute for Psycholinguistics
Nijmegen, Netherlands
sean.roberts@mpi.nl

Previous experiments and models show that the pressure to communicate can lead to the emergence of symbols in specific tasks. The experiment presented here suggests that the ability to use deictic gestures can reduce the pressure for symbols to emerge in co-operative tasks. In the *gesture-only* condition, pairs built a structure together in *Minecraft*, and could only communicate using a small range of gestures. In the *gesture-plus* condition, pairs could also use sound to develop a symbol system if they wished. All pairs were taught a pointing convention. None of the pairs we tested developed a symbol system, and performance was no different across the two conditions. We therefore suggest that deictic gestures, and non-referential means of organising activity sequences, are often sufficient for communication. This suggests that the emergence of linguistic symbols in early hominids may have been late and patchy with symbols only emerging in contexts where they could significantly improve task success or efficiency. Given the communicative power of pointing however, these contexts may be fewer than usually supposed. An approach for identifying these situations is outlined.

1. Introduction

Gesture, and pointing in particular, is often heralded as a crucial step in the evolution of language (Arbib, Liebal, & Pika, 2008; Tomasello, Carpenter, & Liszkowski, 2007; Tomasello, 2010; Liszkowski, 2010). Pointing by itself is versatile enough to communicate about activities, objects, people, events, places and more, and can be made sophisticated when combined with theory of mind, advanced pragmatic inference, and abilities to engage in displaced reference. Used with attention-getters, pointing is more powerful still. One might wonder then, under what conditions symbols are necessary to complete complex collaborative tasks, given a pre-existing deictic system.

Experiments and models using a director-matcher tasks suggest that referential symbol systems routinely emerge (Steels, Belpaeme, et al., 2005; Verhoef, Roberts, & Dingemanse, 2015; Skyrms, 2010). In these tasks, a director requests an object with a signal and a matcher must guess which object they want. They are then given feedback indicating the correct answer. Stable symbols for objects usually emerge over time through interaction. However, the availability of communicative tools in many of these experiments are managed is such that the only way to communicate (and the only way to reliably complete the task) is with referential symbols. It in this case, it is not surprising that they emerge.

In addition, setting up a novel symbol system within the context of a task is costly, as it takes time and energy that could be spent completing the task. In this case, if a symbol system does not offer a sufficient pay-off in terms of increasing task efficiency that can off-set this cost, then players may be biased against setting up such a symbol system. In director-matcher tasks, the costs and pay-offs are fairly straightforward: symbols are the only way to reliably complete the matching task (high pay-off), but they take time to stabilise so many early trials are unsuccessful (high cost). In situations with more freedom, the costs and pay-offs can be more complex, and not favour the emergence of symbols in such a directed way.

Accordingly, in the current experiment we tested whether novel symbol systems emerged, and whether they aided task performance, in a co-operative building task set in an embodied, 3D environment where it was possible to point and gesture. The task was complex enough such that having a symbol system would allow participants to complete the task more efficiently, but symbols were not the only means of communicating. Under these conditions, and in contrast to the experiments above, we found that no symbol systems developed. This suggests that deictic tools, and features of specific tasks (complexity, degree of collaboration required, etc) can affect the likelihood of the emergence of a symbol system. This type of experiment may then also help address debates about the extent to which linguistic abilities can be inferred from complex material artefacts and evidence of cooperation (Davidson & Noble, 1992; Cuthbertson & McCrohon, 2012).

2. Experiment

The experiment tested whether partners would develop a symbol system while performing a reasonably complex co-operative task in Minecraft (Mojang, 2015), a computer game set in a three dimensional virtual world. The experiment had two conditions which manipulated the ways in which partners could communicate. In the *gesture-only* condition partners were taught a pointing convention (gaze at something and jump repeatedly), and were told that they could only communicate using their avatar in the game (so they could use pointing and other gestures). In the *gesture-plus* condition partners were also taught the pointing convention, but were additionally allowed to use a second communication channel: knocking on the table (we do not assume that auditory communication is special; this was just an easy way to implement an additional ostensive channel). This provided a

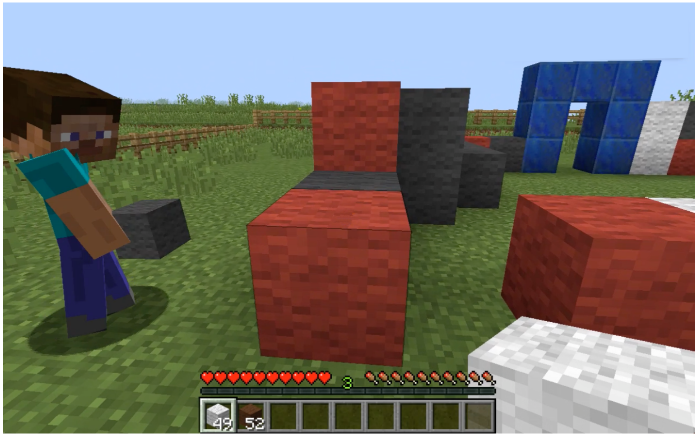


Figure 1. A screenshot from the experiment. The view is from the the eyes of a virtual avatar. In the foreground are some of the placed blocks of the structure and one of the doorways. On the left is the other player's avatar, whose head orientation indicates they are looking at the red block in the centre foreground.

discrete signal space which was suited to labelling the discrete objects in the task.

The main hypothesis was that participants in both conditions (*gesture-only, gesture-plus*) would perform the task equally well, measured in terms of time taken to complete task, and number of errors made. In addition, we predicted that participants in the *gesture-plus* condition would not use their extra communication channel to develop a referential symbol system (though they might use it in other ways). These predictions stem from the ideas that 1) pointing and other deictic gestures are sufficient for many communicative needs and 2) it is costly to set up novel symbol systems, so they will only emerge if there is a direct pressure to do so.

2.1. *Method*

Participants built a construction together in Minecraft. Each participant can move around in and manipulate the world via a humanoid virtual avatar and can see their partner's avatar, including their gaze direction. The world is made up of a regular matrix of crate-sized blocks of different kinds (rock, dirt, stone etc.) that can be placed or removed (like life-sized lego, see figure 1).

First, in a training exercise, participants were taught how to move and look around the 3D environment. Then they were taught how to place and remove blocks, and given a plan of a small practice structure to build independently. The next task was to learn a pointing convention, which consisted of gazing directly

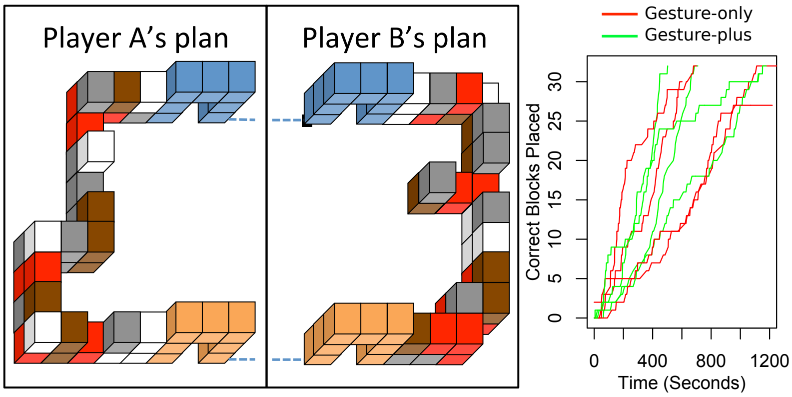


Figure 2. Left: The plan of the structure which player A and player B were given. The two doorways at the top and the bottom were already placed in the world. Right: The rate for correctly placing communicated blocks. Lines are coloured by condition.

at a block and then jumping. Participants practiced 'pointing' to a block for their partner. Experimenters checked participants' progress and answered questions.

The main task involved building a structure from a plan. Each player had only half of the plan for the whole structure (figure 2). Each side of the plan included the same number of blocks, but the plan was not symmetrical. The plan included coloured doorways which were already placed in the world so that the participants had obvious anchor points to use. One of each type of block was also placed to one side of the doorways (figure 3), though participants' attention was not directed to these. Each side of the plan was composed of four different coloured blocks (white, brown, grey and red - made of Minecraft 'dyed wool' which is easy to destroy). One player was given only white and brown blocks, while the other was given only grey and red blocks.

Participants were told that they could not place block types belonging to their partner, but they could destroy any blocks. Participants were then told that they would have to communicate with their partner to get them to place their coloured blocks in the appropriate places. Participants were told to complete the task in 20 minutes (but were given as much time as they needed).

The key manipulation was in the ways that participants could communicate. Participants in the *gesture-only* condition could only communicate via the actions of their avatar in the game. In the *gesture-plus* condition, participants were additionally allowed to knock on the desk with their knuckles in order to communicate. After the experiment, participants were given a questionnaire on their strategies and previous experience with Minecraft. A symbol system was considered to have

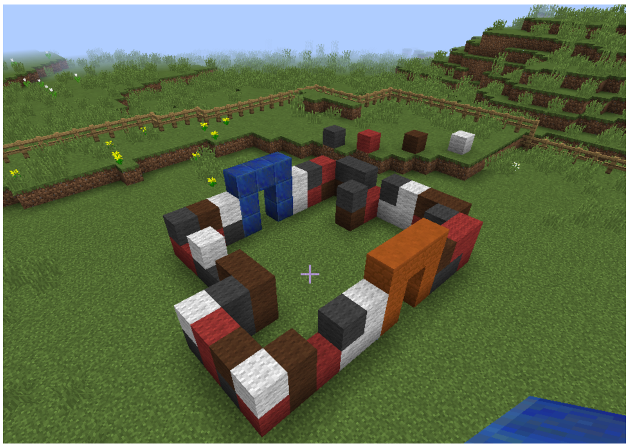


Figure 3. A completed structure viewed from the air, with some of the surrounding environment.

developed if participants reported that they could refer to at least one block type using a non-deictic method (either auditory or gestural). Participants played on two MacBook laptops which were connected by Wi-Fi. A CraftBukkit Minecraft server with customised plugins was used to manage the world and record player activity. Video and audio were also recorded for later coding.

2.2. *Results*

Results for 4 pairs in each condition were analysed. Participants' performance was rated according to the accuracy with which they matched the plan, the amount of time they took and the rate of successfully communicated blocks (blocks correctly placed by a player on their partner's side of the plan, for which communication was necessary).

All participants were essentially at ceiling in terms of accuracy. 5 trials had no errors, while the rest had 1, 2 and 10 errors respectively (less than 5% of possible errors in the bounding box of the plan), the latter two from the *gesture-only* condition. The majority of the last case can be attributed to duplicate insertions which amplified the errors (i.e. the errors are lower in terms of edit distance), and this may have more to do with map reading than difficulty communicating. All participants stated that they were satisfied that their final structure matched the plan. Pairs took between 8.3 minutes and 21.9 minutes. While the sample size is not big enough to pick up small differences, the two conditions did not differ significantly in time taken. Figure 2 shows the rate for correct placement of communicated blocks, demonstrating that the two conditions overlap. Participants varied in their previous experience with Minecraft, but this did not predict success at the task.

Finally, there was minimal use of the extra communication channel (knocking). One pair did not used the channel at all. Two pairs used knocks as attention-getters at crucial times in the game (though used only once by one pair, and twice by another). One pair made more regular use of knocking, but knocks were not used referentially: one partner used it as an attention getter and to mark progress (e.g. roughly 'that's correct, now move on'), and the other partner similarly used it to mark communicative success or lack of it (roughly 'I understand' and 'I don't understand'). There was no overlap in the 'meanings' of knocks across this pair, and they were only used when Player A was directing the placement of Player B's blocks. Interestingly, the pair who did not use the channel at all had the shortest completion time. These participants reported that "I didn't think we needed to [use sound]" and "we found other ways to communicate what we wanted to say".

As reported in questionnaires, at least two participants attempted to develop a symbol system for colour in the *gesture-only* condition. One tried to code for colour via number of jumps, and another tried to code for the colour of their partner's block via the colour of the block they were holding (white means red, brown means grey), but their partners either did not notice, or did not understand, and these symbols were not adopted.

In general however, it is hard to identify trends. There was massive variation in the way communicative conventions were established (or not). Sometimes conventions were shared across a pair, such that the signs used by the first 'director' (communicating where they wanted differently coloured blocks placed) were copied by the second 'director'. Sometimes these signs were streamlined by the second director (e.g. only one sign each for location and colour, or a smaller meaning space), and sometimes they were added to (more ways to convey the same meaning, and a bigger meaning space). Some partners had signs for 'correct' and 'wrong' but some did not.

One clear result is that the same gestures were often used to convey multiple meanings within a pair (even to one sign meaning both 'that's right' and 'no, stop!'). The meaning intended was usually obvious from context. Deictics aimed at coloured blocks, used to convey both colour and location were usually easily disambiguated via context, and via the universally adopted system of referring to colour first, and location second. See the supplementary materials for video examples of various phenomena above.

Also, the fastest pairs all adopted a common strategy of completing their own side of the building as far as they could by themselves, then working with a partner, completing one colour at a time. This was efficient both in terms of time, and in terms of communicative effort: if only one colour is being placed at a time, then the director only needs to communicate about location.

3. Discussion

Our hypotheses regarding the use of pointing in reasonably complex co-operative tasks were confirmed: no pair developed a symbol system, and providing an extra communication channel did not aid performance. Instead, pairs used both the pointing convention they had been taught, and developed other deictic and non-referential signs (13/16 participants, or 7/8 pairs, used other deictic gestures), which were often fairly fluid in their meanings.

Digging deeper into questionnaires and debriefings to see why no symbol systems were developed in either condition, something like a paradox emerges. The less skilled or practised a participant is at the task, the less easily they can move in a precise way, the less efficiently they perform the task (so e.g. fail to use the one-colour-at-a-time strategy), and so the more they need symbols to minimise communicative effort. But given their lack of control over precise movements, setting up a symbol system is hard and costly. As one participant noted: "not being able to very accurately control [my avatar]... the risks of miscommunication are far too great". Even if it had been possible to set up a simple symbol system, being physically unable to reliably repeat it would have hampered its use, and probably led to it being abandoned. Another participant explained that a symbol system "would have taken much longer to develop", highlighting the cost involved. Instead, a range of fairly direct deictic gestures (hitting and jumping on blocks) were more easily performed and interpreted, and their precise form (e.g. number of jumps) does not matter.

Conversely, the more skilled or practised a participant is, the more easily they can move precisely, the more efficiently they do the task (e.g. use the one-colour-at-a-time strategy), and so the less they need symbols. They are more physically capable of setting up a symbol system, but use game strategies that minimise the need to communicate. In this case, deictics and non-referential signs, combined with highly efficient task strategies, are entirely sufficient for their needs.

Combinations of deictic and non-referential gestures (e.g. affirmatives) may then function as a local fitness peak in the context of some tasks. Those at the top of the local fitness peak can use deictics and other non-referential gestures to successfully and efficiently complete these tasks, and would not gain much by paying the costs of setting up a symbol system (switching to the global fitness peak). Those at the bottom of the fitness landscape would benefit most by developing a symbol system, but doing so would require developing their deictic skills, which would trap them in the local peak of the first group. In neither case then is a symbol system likely to emerge within some kinds of tasks, where deictic and other non-referential gestures can form a kind of '**deictic fitness trap**'.

Intriguingly, there was also evidence of the emergence of ways to organise action sequences. All but 3 participants used signals that helped to segment the task structure ('yes, now move on' or 'no, that's wrong'), even when these were not necessary. One pair also developed something like a repair sequence, entirely mediated through pointing (see the supplementary materials). However, they are clearly helpful when participants must self-organise to complete sequences of sub-tasks. Previous models and experiments rule out the need for sequence organisation by assigning roles to participants (e.g. in director-matcher tasks), but this experiment suggests that signals to aid with sequence organisation emerge before symbols.

This can be used to suggest that symbol systems may have been late and patchy in emerging, and tied to features of specific tasks, though preceded by and existing along with deictic and non-referential communication conventions, and backed by well-developed theory of mind and abilities for pragmatic inference. More experimental research is needed on exactly what sort of tasks, relevant to hominid lifeways, would have absolutely required symbolic communication, and which would not. Classic signal guessing games are a poor framework in this regard, since if pointing was available, it would often make the development of symbols redundant. Embodied, cooperative and (somewhat) ecologically valid tasks such as the one in this experiment force researchers to specify which concrete tasks or social goals demand the use of symbol systems.

An alternative approach comes from (Sterelny, 2014, 2015), who identifies major new pressures on communication that would likely have acted on fission-fusion hominid groups present between 500 and 100kya (moving into sapiens and Neanderthals). These hominids would have needed to track economic exchanges covering a large range of resources over extended periods of time (e.g. how much is a leg of deer from last week worth now in tubers), and track others' reputation so as to plan future exchanges.

In this case a combination of pressures may lead to the emergence of symbols: 1) the need to refer to the distant elsewhere and elsewhen including 2) a wide range of resources that must be differentiated, among 3) individuals who do not share a huge amount of common ground (because of the fission-fusion structure). In contrast to tasks carried out in the here-and-now, in which deictics embedded in routines would likely have sufficed for communication, symbols would make communication far more efficient in these economic exchanges, and come with a start up cost worth paying. Identifying such social and communicative constraints on hominid lifeways is likely to be a productive way of framing specific hypotheses about symbol emergence, situated as they are within the activities and social contexts in which hominids had to function.

Acknowledgements

SR is supported by an ERC Advanced Grant No. 269484 INTERACT to Stephen Levinson. We would like to thank the Philosophy Department and the AHR Centre of Excellence for the Dynamics of Language at ANU for hosting the authors while these experiments were carried out.

References

Arbib, M. A., Liebal, K., & Pika, S. (2008). Primate vocalization, gesture, and the evolution of human language. *Current Anthropology*, *49*(6), 1053–1076.

Cuthbertson, C., & McCrohon, L. (2012). Sea crossings are an unreliable indicator of language ability in hominids. In T. C. Scott-Phillips, M. Tamariz, E. A. Cartmill, & J. R. Hurford (Eds.), *The Evolution of Language: Proceedings of the 9th International Conference (EVOLANG9)* (pp. 88–95). World Scientific.

Davidson, I., & Noble, W. (1992). Why the first colonisation of the australian region is the earliest evidence of modern human behaviour. *Archaeology in Oceania*, *27*(3), 135–142.

Liszkowski, U. (2010). Deictic and other gestures in infancy. *Acción psicológica*, *7*(2), 21–33.

Mojang. (2015). Minecraft 1.8.8. *Retrieved from* `https://minecraft.net` *10/08/2015.*

Skyrms, B. (2010). *Signals: Evolution, learning, and information.* Oxford University Press.

Steels, L., Belpaeme, T., et al.. (2005). Coordinating perceptually grounded categories through language: A case study for colour. *Behavioral and brain sciences*, *28*(4), 469–488.

Sterelny, K. (2014). A paleolithic reciprocation crisis: symbols, signals, and norms. *Biological Theory*, *9*(1), 65–77.

Sterelny, K. (2015). Deacon's challenge: From calls to words. *Topoi*, 1–12.

Tomasello, M. (2010). *Origins of human communication.* MIT press.

Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child development*, *78*(3), 705–722.

Verhoef, T., Roberts, S. G., & Dingemanse, M. (2015). Emergence of systematic iconicity: Transmission, interaction and analogy. In D. Noelle, et al. (Ed.), *The 37th annual meeting of the Cognitive Science Society (CogSci 2015).*

Abstracts

NONLINEAR BIASES IN ARTICULATION CONSTRAIN THE DESIGN SPACE OF LANGUAGE

RICK JANSSEN[1], BODO WINTER[2], SEÁN ROBERTS[1],
SCOTT R. MOISIK[1], DAN DEDIU[1]

[1]*Max Planck Institute for Psycholinguistics,*
Nijmegen, the Netherlands
rick.janssen@mpi.nl, sean.roberts@mpi.nl, scott.moisik@mpi.nl, dan.dediu@mpi.nl

[2]*Cognitive & Information Sciences,University of California, Merced,*
California, United States of America
bodo@bodowinter.com

In Iterated Learning (IL) experiments, a participant's learned output serves as the next participant's learning input (Kirby et al., 2014). IL can be used to model cultural transmission and has indicated that weak biases can be amplified through repeated cultural transmission (Kirby et al., 2007). So, for example, structural language properties can emerge over time because languages come to reflect the cognitive constraints in the individuals that learn and produce the language. Similarly, we propose that languages may also reflect certain *anatomical* biases. Do sound systems adapt to the affordances of the articulation space induced by the vocal tract?

The human vocal tract has inherent nonlinearities which might derive from acoustics and aerodynamics (cf. quantal theory, see Stevens, 1989) or biomechanics (cf. Gick & Moisik, 2015). For instance, moving the tongue anteriorly along the hard palate to produce a fricative does not result in large changes in acoustics in most cases, but for a small range there is an abrupt change from a perceived palato-alveolar [ʃ] to alveolar [s] sound (Perkell, 2012). Nonlinearities such as these might bias all human speakers to converge on a very limited set of phonetic categories, and might even be a basis for combinatoriality or phonemic 'universals'.

While IL typically uses discrete symbols, Verhoef et al. (2014) have used slide whistles to produce a continuous signal. We conducted an IL experiment with human subjects who communicated using a software-implemented slide whistle for which the degree of nonlinearity is controlled. A single parameter (α) changes the mapping from slide whistle position (the 'articulator') to the acoustics. With α=0, the position of the slide whistle maps Bark-linearly to the acoustics. As α approaches 1, the mapping gets more double-sigmoidal, creating three plateaus where large ranges of positions map to similar frequencies. In more abstract terms, α represents the strength of a nonlinear (anatomical) bias in the vocal tract.

Six chains (138 participants) of dyads were tested, each chain with a different, fixed α. Participants had to communicate four meanings (pictographs showing different animals) by producing a continuous signal using the slide-whistle in a 'director-matcher' game, alternating roles (cf. Garrod et al., 2007).

Results show that for high αs, subjects quickly converged on the plateaus. This quick convergence is indicative of a strong bias, repelling subjects away from unstable regions already within-subject. Furthermore, high αs lead to the emergence of signals that oscillate between two (out of three) plateaus. Because the sigmoidal spaces are spatially constrained, participants increasingly used the sequential/temporal dimension with higher αs (i.e., more nonlinear mappings). As a result of this, the average duration of signals with high α was ~100ms longer than with low α. These oscillations could be an expression of a basis for phonemic combinatoriality.

We have shown that it is possible to manipulate the magnitude of an articulator-induced non-linear bias in a slide whistle IL framework. The results show that language might indeed come to reflect the nonlinear mapping from the articulators to acoustics. In particular, the signaling systems in our study quickly converged (within-subject) on the use of stable regions. While these conclusions were drawn from experiments using slide whistles with a relatively strong bias, weaker biases could possibly be amplified over time by repeated cultural transmission, and likely lead to similar outcomes.

Future studies could investigate anatomical biasing with more realistic models of the articulators, and address the interaction with other factors (socio-linguistics, environment, etc.). Our model is a deliberate abstraction from reality in order to tightly control experimental conditions. In reality of course, anatomical biases have to be thought of as one factor that shapes human language, but in complex ways.

References

Garrod S., Fay N., Lee J., Oberlander J., MacLeod T. (2007) Foundations of representation: where might graphical symbol systems come from? *Cognitive Science*, 31, 961-987.

Gick, B., & Moisik, S. R. (2015). Quantal biomechanics in an embodied phonetics. *JASA*, 137(4), 2301–2301.

Kirby, S., Dowman, M., & Griffiths, T. L. (2007). Innateness and culture in the evolution of language. *PNAS*, *104*, 5241-5245.

Kirby, S., Griffiths, T.L., & Smith, K. (2014). Iterated learning and the evolution of language. *Current Opinion in Neurobiology*, *28,* 108-114.

Perkell, J. S. (2012). Movement goals and feedback and feedforward control mechanisms in speech production. *J. Neuroling.*, 25(5), 382–407.

Stevens, K. N. (1989). On the quantal nature of speech. *J. Phon.*, *17,* 3-46.

Verhoef, T., Kirby, S., & de Boer, B. (2014). Emergence of combinatorial structure and economy through iterated learning with continuous acoustic signals. *Journal of Phonetics*, *43*, 57-68.

Index