Title
Interaction dynamics affect the emergence of compositional structure in cultural transmission of space-time mappings

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Authors
Verhoef, Tessa
Walker, Esther
Marghetis, Tyler

Publication Date
2022

Peer reviewed

Interaction dynamics affect the emergence of compositional structure in cultural transmission of space-time mappings

Tessa Verhoef[1] (t.verhoef@liacs.leidenuniv.nl)
Creative Intelligence Lab, Leiden University, Niels Bohrweg 1
2333 CA Leiden, The Netherlands

Esther Walker (ejwalker17@gmail.com)

Tyler Marghetis[1] (tyler.marghetis@gmail.com)
Department of Cognitive & Information Sciences, University of California, Merced

Abstract

People talk about time using the language of space. The future is "ahead." Endless events are "long." Cross-linguistically, these conventions exhibit both universality and striking diversity. These mappings in language, therefore, might originate from a combination of shared cognitive biases and sociocultural processes. To investigate the mechanisms involved in the emergence of space-time mappings—and linguistic metaphor more broadly—we conducted an experiment in which participants had to communicate about abstract temporal concepts using entirely spatial signals. The spatial signals developed by one pair of participants were then transmitted to the next pair, creating chains of multiple generations. Together, these processes of interaction and transmission sometimes generated fully systematic, compositional systems—although sometimes also generated systems that lacked structure entirely. The deciding factor may have been how people responded to errors — with incremental adjustments or radical reconfiguration. Systematic metaphors, therefore, may emerge from a heterogeneous mix of mechanisms.

Keywords: cultural evolution; language emergence; space-time cognition; cognitive biases; social coordination

Introduction

Where is the future? For speakers of English, it's "ahead." If you've spent time in the US Military, you might say it's "to the right" (Hendricks, Bergen, & Marghetis, 2018). And if you're Aymara, the future and past are reversed, with the future construed as behind the speaker (Núñez & Sweetser, 2006). Around the world, people communicate about time using space—but with considerable variation in the details (Núñez & Cooperrider, 2013).

These observations raise questions about the emergence of idiosyncratic conventions for communicating about time within larger cross-linguistic regularities. Here, we conduct an experiment combining dyadic communication and iterated transmission to address this question.

Space and time in language

In cultures around the world, the language of space is used to describe temporal *duration*, temporal *sequence*, and temporal *deixis* (i.e., past, present, future). Within a language, the description of time using spatial words is both systematic (e.g., past and future are described using contrastive spatial terms) and productive (i.e., speakers can create new expressions). Scholars have thus argued that speakers possess an underlying *conceptual metaphor* that reflects the conventions of their language or culture (Lakoff & Johnson, 1980; Núñez & Cooperrider, 2013; Boroditsky, 2001; Pitt & Casasanto, 2020). We refer here to the linguistic conventions as "space-time mappings."

Across languages, temporal duration is reliably described using some dimension of magnitude. In English and Indonesian, duration is described as *length*. Days are "short" in the winter, "long" in the summer. In Greek and Spanish, duration is described as *quantity or amount* (e.g., Greek "poli ora" [much time]). While the spatial dimension varies, the orientation of the mapping from space to time appears to be stable cross-linguistically, with no attested examples in which *more* duration is described as *less* spatial magnitude.

Spatial descriptions of temporal sequence and deixis, on the other hand, exhibit more cross-linguistic variation. In English we describe temporal sequences using words like "before" and "after" that are now primarily temporal but which, historically, had primarily spatial meanings. The mapping is more explicit for temporal *deixis* (i.e., past, present, future), where we might say, "looking *forward* to the future," or "thinking *back* to the past." But the spatial axis and its orientation vary cross-linguistically: east to west (Boroditsky & Gaby, 2010), uphill (Núñez, Cooperrider, Doan, & Wassmann, 2012), rightward (Hendricks et al., 2018), or downward (Boroditsky, 2001). This variation is extensive (Núñez & Cooperrider, 2013).

Cross-linguistic space-time mappings thus exhibit both diversity and near-universality. This suggests that space-time mappings in language might originate from processes operating at both individual and sociocultural levels (Núñez & Cooperrider, 2013; Verhoef, Walker, & Marghetis, 2016). Which processes might be at work?

Mechanisms of cultural evolution of language

Lab-based experiments have helped isolate some of the relevant mechanisms that play a role in the emergence of novel linguistic systems (Scott-Phillips & Kirby, 2010). Communication systems can emerge rapidly from repeated interactions between pairs of participants engaged in a communication task (Garrod, Fay, Lee, Oberlander, & MacLeod, 2007; Fay, Garrod, & Roberts, 2008; Fusaroli & Tylén, 2012). In these experiments, conventions emerge from social coordina-

[1]These authors contributed equally.

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society.*

tion and shared communication history.

In Verhoef et al. (2016), we demonstrated that these mechanisms can lead to the emergence of space-time mappings. Pairs of participants communicated repeatedly about time concepts using a spatial signaling device (Fig. 1a). Conventions emerged in which space was used systematically to represent time. For instance, participants reliably used greater spatial length to indicate greater temporal duration, and nobody systematically used the opposite strategy (Fig. 1b), in line with the patterns found cross-culturally in natural languages. We attributed this to the presence of a strongly shared cognitive bias to associate length and duration. A more variable convention involved the use of spatial location to indicate temporal deixis or sequence. While all pairs tended to settle on a consistent mapping (e.g., future-up, past-down; Fig. 1c), there was variation in its orientation (i.e., some pairs settled on future-down, past-up; Fig. 1d). This pattern therefore resembled the cross-linguistic variation in the spatialization of temporal deixis and sequence. We argued that these conventions depended on social coordination and shared history within the task.

Another potential mechanism is transmission to new language learners (Kirby, Cornish, & Smith, 2008; Verhoef, 2012; Kirby, Griffiths, & Smith, 2014). Transmission of a language causes signals to be filtered through the cognitive constraints of learners; signals that are easily learned are more likely to be reproduced and passed on. Repeated transmission can thus make signals more systematic. Indeed, while all interacting pairs in Verhoef et al. (2016) developed some space-time mappings, these were never sufficiently elaborated or systematic to achieve perfect communication. Moreover, widespread compositionality, in which subparts of the signals clearly referred to specific parts of the meaning space, did not emerge.

It has been suggested that both transmission and interaction are needed to develop more robust systematicity, resembling patterns in human language (Kirby, Tamariz, Cornish,

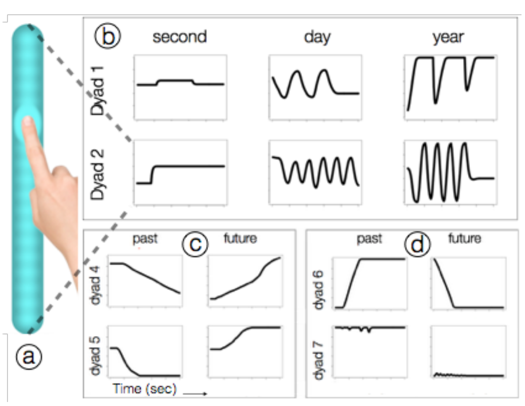

Figure 1: (a) Signaling device, (b) Consistent use of spatial length to communicate relative duration, (c) and (d) Contrasting mappings for past/future. From Verhoef et al. (2016)

& Smith, 2015). Past work has investigated the relative contributions of transmission and interaction by manipulating the presence of each process. Results have been mixed. Kirby et al. (2015), for instance, found that compositional structure did *not* emerge in a novel language when dyads interacted repeatedly and transmission was absent, but it did appear when both interaction and transmission were present. Other studies have found that interaction on its own *can* produce systems with some structure, but that the most structured systems emerged when interaction was combined with transmission (Theisen-White, Kirby, & Oberlander, 2011; Saldana, Kirby, Truswell, & Smith, 2019). Others have found that structure can emerge successfully without any transmission at all (Winters, Kirby, & Smith, 2018; Raviv, Meyer, & Lev-Ari, 2019; Nölle, Staib, Fusaroli, & Tylén, 2018). Finally, in some contexts, transmission can actually result in *less* structure than interaction (Garrod, Fay, Rogers, Walker, & Swoboda, 2010). The role of transmission, therefore, may vary.

In natural languages, both interaction and transmission are present. The cultural evolution of space-time mappings, therefore, has occurred in the context of both mechanisms. Here, we thus add transmission to the set-up of Verhoef et al. (2016) to test whether these two mechanisms in combination can produce fully compositional systems and error-free communication. To investigate the mechanisms' relative contributions, we look *within each chain* at how interaction and transmission shape the structure of space-time mappings.

The current study

In this study, pairs of participants had to communicate about abstract temporal concepts (e.g., next year, year before, yesterday). Their only signals, however, were vertical spatial movements. The spatial signals developed by a pair of participants were then transmitted to the next pair, who could use these signals as the basis for their own attempts at communication. This process thus allows for influences of individual biases (e.g., initial interpretations, shared expectations), social coordination (i.e., negotiation between pairs over the course of an interaction session), and iterated transmission (i.e., when the signals from one pair are shared with the next).

Methods

Following the design of Verhoef et al. (2016), pairs of participants ("dyads") had to communicate about temporal concepts using only spatial movement. Here we add a process of *iterated transmission* in which signals developed by one dyad are used as initial training for the next. This generated six transmission chains of eight interacting pairs. Undergraduate students ($N = 96$) at the University of California, San Diego (UCSD) participated in return for partial course credit.

Signals and meanings

Participants produced communicative signals using a touch screen (see Fig. 1a), which recorded and replayed sequences of vertical movement lasting exactly 5 seconds. In replay, a bubble moved continuously, reproducing the movement of the participant's index finger. The set of meanings were the same 18 meanings used in Verhoef et al. (2016):

- Duration: *'second', 'day', 'year'*
- Sequence: *'day before', 'day after', 'year before', 'year after', 'before', 'after'*
- Deictic: *'now', 'yesterday', 'today', 'tomorrow', 'last year', 'this year', 'next year', 'past, 'future'*

Note that these items can be distinguished broadly along three dimensions: temporal length (e.g. 'day' vs. 'year'), temporal direction (e.g. 'future' vs. 'past', 'before' vs. 'after'), and category (duration, sequence, deictic).

Procedure

Participants received instructions together and were then placed in separate rooms. They could only communicate through the touch screen interface. An initial training was followed by four rounds of interaction.

Training Each participant first completed a training phase. They were told, correctly, that their partner was being trained on the same system. In the training, a signal was played (i.e., movement on the vertical bar) while its intended meaning was displayed alongside. The meaning then disappeared and the participant was asked to reproduce the signal on the touch screen. After their attempted reproduction was recorded, they were asked to select the signal's intended meaning from the full list of 18 options. Each meaning was shown twice in training, for a total of 36 training items.

The first dyad in each chain was trained on one of three "seed" systems produced by interacting dyads in Verhoef et al. (2016). These systems were not fully systematic or unambiguous, but all exhibited some form of space-time mapping. For instance, each seed system used greater length to indicate greater duration (Fig. 1b). Seeds varied in their use of vertical location to differentiate past from future (two used up = future; one used up = past; Fig. 1c and d). Each of the three initial languages seeded two transmission chains. Training systems in subsequent generations were taken from the final interaction round of the previous dyad, meaning that 50% of the signals came from each participant.

Interaction After training, participants alternated attempts to communicate a target meaning to their partner. Trials began with a target meaning displayed on the signaller's screen. The signaller recorded a signal for this meaning using the vertical bar. This signal was then replayed on the recipient's screen, who then guessed the intended meaning out of a list of all 18 options. Both players received feedback after each trial (i.e., target meaning and guessed meaning). Each dyad completed four rounds of eighteen trials.

Results

During training, participants successfully reproduced the training signals and recalled their meanings. To calculate distance between spatial signals, we used Dynamic Time Warping (DTW) (Sakoe & Chiba, 1978), a score based on the effort required to align two temporal sequences. Participants produced signals that were very close to the training signals (DTW distance, $M = 0.03$, $SD = 0.01$). They also accurately recalled each signal's meaning ($M = 0.96$, $SD = 0.04$), although this was tested immediately after they had been shown the correct meaning, so this was mostly an attention check. Participants thus paid attention during the training and had no difficulty reproducing the signals.

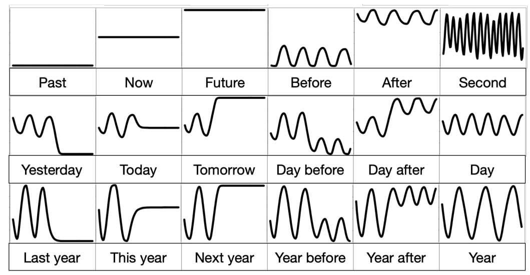

Figure 2: Some transmission chains generated fully compositional systems.

Emergence of compositional structure

In the communication task in Verhoef et al. (2016), participants produced signals with stable space-time mappings, and these signals' compositional structure increased over rounds of interaction. However, these systems never became fully systematic and unambiguous. In the current experiment, by contrast, repeated interaction and transmission produced near-perfect communicative success in some transmission chains. This perfect communicative accuracy reflected the gradual appearance of fully systematized, compositional space-time mappings. Fig. 2 illustrates one system that emerged at the end of one transmission chain. Note the signals' systematic space-time mappings and compositionality. For instance, temporal deixis (past/future) is expressed through location, with future mapped to higher locations. Temporal duration, on the other hand, is expressed by spatial extent, with day-meanings using smaller spaces than year-meanings. These basic conventions, moreover, were combined to create more complex signals, with duration indicated at the start of the signal and sequence or deixis indicated at the end. Communicative accuracy among participants using this system was perfect (100%).

Structured systems were associated with communicative success

We analyzed communicative success across rounds and generations using a linear mixed effects model of trial-by-trial accuracy, with fixed effects of round and generation and a random effect of chain. Accuracy increased significantly over the four rounds within each generation ($b = 0.067 \pm 0.016$ *SE*, $p < .001$), but not over generations of iterated transmission ($b = 0.012 \pm 0.008$, $p = .16$).

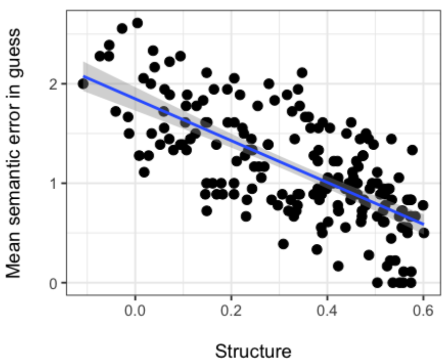

Figure 3: More structured languages are associated with semantically less distant guessed words in interaction.

To quantify a system's compositional structure, we calculated the correlation between pairwise meaning distance and pairwise signal distance, following Kirby et al. (2008). The intuition is that compositional systems, since they express more complex meanings by combining simpler signals, will use more similar signals to express more similar meanings. Distances between signals were calculated with Dynamic Time Warping. Distances between meanings ("Semantic Distance") was calculated by summing differences in temporal order (past/before < now < future/after), temporal extent (e.g. now < today < year after), and semantic class (duration, sequence, or deictic). Structure was calculated as Pearson correlation between pairwise signal distance between all possible pairs of signals and pairwise semantic distance between associated meanings. This was calculated for each round of each dyad in each chain.

Communication was easier with more structured systems (Fig. 3). Among dyads using more structured systems, guesses were closer to the target meaning. This was confirmed by a linear mixed effects model of semantic error in the guessed meaning as a function of structure and generation; variance across chains was modeled with a random intercept and a random slope for generation. Semantic error decreased significantly as structure increased ($b = -1.064 \pm 0.359$ *SE*, $p < .01$), but there was no change over generations ($b = 0.059 \pm 0.038$ *SE*, $p = 0.14$). There was a significant interaction between structure and generation, with a stronger association between structure and decreased error in later generations ($b = -0.171 \pm 0.074$ *SE*, $p < .01$), driven by the few chains in which later generations developed near-perfect accuracy and fully systematized space-time mappings.

Changes in duration vs. direction mappings

In Verhoef et al. (2016), repeated *interaction* produced reliably similar length-duration mappings, while dyads' location-direction mappings (e.g., up-future) were more variable. We thus analyzed whether these two types of convention were differentially affected by *transmission*. To quantify the length-duration convention, meanings with a duration were grouped: Moment (1), Day (2) or Year (3). For each signal, we calculated the Vertical Space Used (VSU) as the standard deviation of the vertical locations throughout the signal (Fig. 4, left). Across chains and generations, 'year' items (blue) had the largest VSU, 'day' items (green) fell in the middle, and 'moment' items (red) had the the smallest VSU. We ran a linear mixed effects model to predict VSU, with fixed effects of temporal duration and generation, and a random effect of chain. As predicted, greater temporal duration was associated significantly with greater VSU ($b = .136 \pm .005$ *SE*, $p < 0.001$). VSU also increased over generations ($b = .013 \pm .002$ *SE*, $p < 0.001$), indicating that signals in general became slightly larger with repeated transmission. In addition, we found a significant interaction between temporal duration and generation ($b = -.004 \pm .001$ *SE*, $p < 0.001$), driven by a decreased difference in VSU between duration categories in later generations. In other words, minimal differences distinguishing durations became more minimal.

We quantified location-direction mappings by determining the focal location of the signal (point of maximal density), between -1 (bottom) and +1 (top). While these mappings were typically stable within a generation, they were more variable across generations than the length-duration mappings, with the location-direction mapping sometimes switching between generations (Fig. 4, right). This switch happened predominantly after transmission, during the initial negotiation of round. Complete reversals of the convention, however, happen only 4 times throughout the experiment, while most pairs retained the previous generation's convention.

Effects of interaction and transmission on structure

Structure increased systematically within each generation over the course of the four interaction rounds. By the final round, every dyad managed to settle on a system with more structure than expected by chance. This is shown in Fig. 5 (left), where structure is plotted as the z score for the measured correlation based on 1,000 random pairings of signals and meaning. The dotted line is the upper limit of the 95% confidence interval for random data; values above this have significantly more structure than expected by chance.

Structure changed with iterated transmission, although not in any systematic direction (Fig. 5, middle). It increased over generations for some chains, but clearly not for all. In some chains (e.g., 3 and 4), structure increased over the course of the transmission chain and always remained above chance. Other chains experienced steep drops in structure, sometimes followed by a recovery over the remaining generations.

These patterns were confirmed by a linear mixed effects model of system structure, with fixed effects of round and generation, random intercepts for chains and random slopes for generation. Structure increased significantly over the multiple interaction rounds within each generation ($b = 0.043 \pm 0.021$ *SE*, $p = .042$). Structure did not increase significantly over generations within each chain ($b = -0.01 \pm 0.017$ *SE*, $p = 0.58$). There was no significant interaction between round and generation. Therefore, structure increased system-

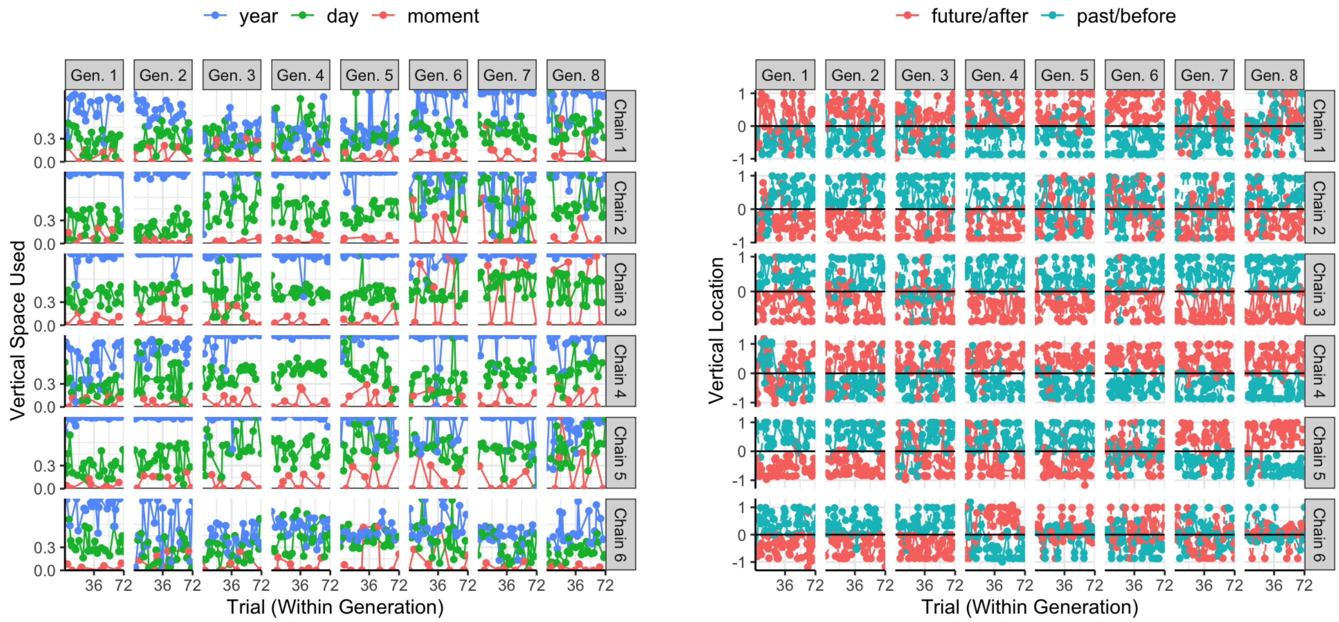

Figure 4: Conventions for temporal direction (left) and duration (right) over generations (columns) and across chains (rows). Each dot represents a single item. (LEFT) Longer durations were reliably represented by more vertical space used (y-axis). (moment = red, day = green, year = blue). (RIGHT) Future/after (red) was typically represented by either higher or lower locations, with past/after (blue) represented by the opposite location (y-axis = location of signal's maximum density).

atically with repeated interaction among dyads but did not reliably increase with transmission between dyads.

This pattern of results was clarified by looking at how signals changed between rounds and across generations (Fig. 5, right). We calculated the change in structure due to interaction as the difference in structure between successive rounds within a generation (i.e., round 1 vs. 2, etc.; lower panel). The change in structure due to transmission was defined as the difference in structure between the final round of one generation and the first round of the next (upper panel). Round-to-round change in structure was distributed unimodally, with a peak slightly above 0, indicating that interaction produced systems that were increasingly structured. Change in structure due to transmission, on the other hand, was distributed bimodally, with a large peak centered at 0, and a smaller peak *below* 0, indicating that structure was mostly maintained in transmission from one generation to the next, but infrequently there was a precipitous drop in structure.

Repair after miscommunication

The analyses above show that signals became increasingly structured and discriminable within generations, as dyads interacted repeatedly, but they did not always maintain this structure as they were transmitted between generations. Why? Recall that the training signals and their meanings were reproduced with high fidelity during each dyad's initial training phase, so this loss of structure did not seem to occur during initial exposure to the previous generation's system.

Instead, loss of structure may have been precipitated by renegotiations after incorrect guesses. Participants made larger changes in their signals if their partner's guess was more semantically distant from the target meaning (Fig. 6). In other words, highly incorrect guesses could prompt very large changes in the communicative system. This was confirmed by a linear mixed effects model of how much a participant changed their signal for a meaning within a generation, with a fixed effect of the semantic distance between their partner's guess and the intended meaning. We included random intercepts for participants. Signals were indeed changed more after incorrect guesses that were further from the correct meaning ($b = 0.023 \pm 0.002$ *SE*, $p < .001$).

Discussion

We investigated the emergence of space-time mappings in an experiment that combined repeated interaction and iterated transmission. This combination of interaction and transmission produced, in some cases, communicative systems that were fully compositional and used with 100% communicative success. Our previous work found that this did not happen with repeated interaction alone (Verhoef et al., 2016). Similar to what was found in Verhoef et al. (2016), however, we saw that duration mappings were stable across generations and similar between independent chains, suggesting there may be a strong shared bias to associate greater duration with greater length. We also confirmed a more flexible bias to map temporal deixis to spatial direction. In Verhoef et al. (2016), differ-

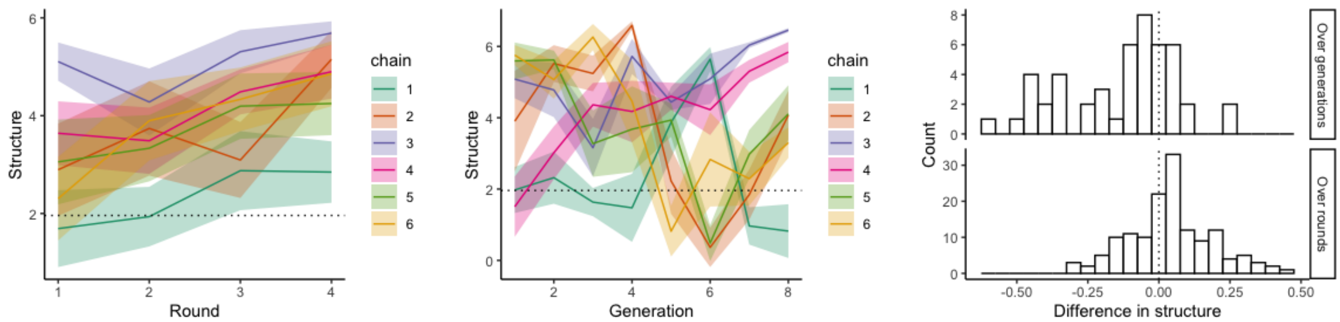

Figure 5: Structure increased reliably with interaction but not transmission. (LEFT and MIDDLE) Across all six chains (colors), mean structure (± standard errors) increased over rounds within a generation (left) but not across generations (middle). Dotted lines indicate the upper limit of the 95% confidence interval for random data. (RIGHT) Interaction over rounds was associated with slight increases in structure (bottom), while transmission across generations was associated with drops in structure (top).

ent dyads settled on opposing strategies, and here the strategy would sometimes switch radically within transmission chains. Notably, while some prior work suggested that structure should gradually increase within transmission (Kirby et al., 2008; Verhoef, 2012; Kirby et al., 2014), here we found that signals became increasingly structured with interaction but often became *less* structured with transmission. This happened predominantly due to re-negotiations after incorrect guesses. Participants changed their signals more if their partner's guess was further from the intended meaning.

While previous studies have proposed that both transmission and interaction are necessary to drive languages to become structured and usable, there is active debate over the precise role of each of these mechanisms. Here, we zoomed in on the precise role of each of these mechanisms within chains that combined transmission and interaction. We saw, overall, that a combination of interaction and transmission could in some cases lead to fully compositional, unambiguous systems. A fine-grained analysis of changes in structure, however, found that structure increased with interaction (i.e, over rounds) but either stayed stable or sharply decreased with transmission (i.e., over generations). This suggests that the gradual increase of structure happens most robustly through interaction, but when it is sufficiently maintained during transmission, more widespread structure and perfect communication can emerge.

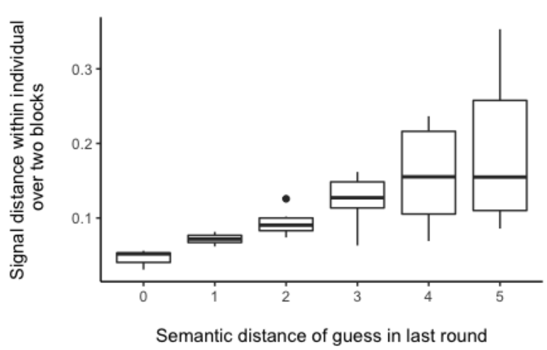

Figure 6: Worse guesses (x-axis) were followed by larger changes in the associated signal (y-axis).

Not all chains led to fully systematized languages. Signals were often ambiguous. Natural languages are similar, with considerable ambiguity in the way they refer to time. Consider: "Wednesday's meeting has been moved forward two days." Interpretations are typically split, with some thinking the meeting is now on Monday, others on Friday (Boroditsky, 2000). But at least one language has found a solution: the variety of English spoken in the US military (Hendricks et al., 2018). In this variety, rescheduling an event earlier is described as moving it "to the left," while rescheduling an event later is described as moving it "to the right." The ambiguity has been removed. So, while it's possible to establish a fully unambiguous spatial language for time, it is difficult — the exception rather than the rule.

Producing a fully structured, compositional, unambiguous system may thus require a perfect storm of pressures and processes. We speculate that repeated interaction may encourage the careful elaboration or *exploitation* of the current set of conventions. Transmission across generations, by contrast, may encourage a more radical reconfiguration or *exploration* of alternative conventions. Trading off between interaction and transmission may thus produce a kind of *simulated annealing*, with repeated cycles of exploration and exploitation. The right balance of these two may be necessary to discover a set of fully unambiguous space-time mappings.

Conclusion

We conducted an experiment to help increase our understanding of how space-time mappings in language can evolve from multiple mechanisms: individual biases, social coordination and iterated transmission. The interplay between these mechanisms may explain the cross-linguistic commonalities as well as the variety found in space-time language.

References

Boroditsky, L. (2000). Metaphoric structuring: Understanding time through spatial metaphors. *Cognition*, *75*(1), 1–28.

Boroditsky, L. (2001). Does language shape thought?: Mandarin and english speakers' conceptions of time. *Cognitive psychology*, *43*(1), 1–22.

Boroditsky, L., & Gaby, A. (2010). Remembrances of times east: absolute spatial representations of time in an australian aboriginal community. *Psychological science*, *21*(11), 1635–1639.

Fay, N., Garrod, S., & Roberts, L. (2008). The fitness and functionality of culturally evolved communication systems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1509), 3553–3561.

Fusaroli, R., & Tylén, K. (2012). Carving language for social coordination: A dynamical approach. *Interaction studies*, *13*(1), 103–124.

Garrod, S., Fay, N., Lee, J., Oberlander, J., & MacLeod, T. (2007). Foundations of representation: Where might graphical symbol systems come from? *Cognitive science*, *31*(6), 961–987.

Garrod, S., Fay, N., Rogers, S., Walker, B., & Swoboda, N. (2010). Can iterated learning explain the emergence of graphical symbols? *Interaction Studies*, *11*(1), 33–50.

Hendricks, R. K., Bergen, B. K., & Marghetis, T. (2018). Do metaphors move from mind to mouth? evidence from a new system of linguistic metaphors for time. *Cognitive Science*, *42*(8), 2950–2975.

Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*, *105*(31), 10681–10686.

Kirby, S., Griffiths, T., & Smith, K. (2014). Iterated learning and the evolution of language. *Current opinion in neurobiology*, *28*, 108–114.

Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, *141*, 87–102.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. University of Chicago Press.

Nölle, J., Staib, M., Fusaroli, R., & Tylén, K. (2018). The emergence of systematicity: How environmental and communicative factors shape a novel communication system. *Cognition*, *181*, 93–104.

Núñez, R., & Cooperrider, K. (2013). The tangle of space and time in human cognition. *Trends in cognitive sciences*, *17*(5), 220–229.

Núñez, R., Cooperrider, K., Doan, D., & Wassmann, J. (2012). Contours of time: Topographic construals of past, present, and future in the yupno valley of papua new guinea. *cognition*, *124*(1), 25–35.

Núñez, R., & Sweetser, E. (2006). With the future behind them: Convergent evidence from aymara language and gesture in the crosslinguistic comparison of spatial construals of time. *Cognitive science*, *30*(3), 401–450.

Pitt, B., & Casasanto, D. (2020). The correlations in experience principle: How culture shapes concepts of time and number. *Journal of Experimental Psychology: General*, *149*(6), 1048.

Raviv, L., Meyer, A., & Lev-Ari, S. (2019). Compositional structure can emerge without generational transmission. *Cognition*, *182*, 151–164.

Sakoe, H., & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE transactions on acoustics, speech, and signal processing*, *26*(1), 43–49.

Saldana, C., Kirby, S., Truswell, R., & Smith, K. (2019). Compositional hierarchical structure evolves through cultural transmission: an experimental study. *Journal of Language Evolution*, *4*(2), 83–107.

Scott-Phillips, T. C., & Kirby, S. (2010). Language evolution in the laboratory. *Trends in cognitive sciences*, *14*(9), 411–417.

Theisen-White, C., Kirby, S., & Oberlander, J. (2011). Integrating the horizontal and vertical cultural transmission of novel communication systems. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 33, pp. 956–961).

Verhoef, T. (2012). The origins of duality of patterning in artificial whistled languages. *Language and cognition*, *4*(4), 357–380.

Verhoef, T., Walker, E., & Marghetis, T. (2016). Cognitive biases and social coordination in the emergence of temporal language. In *Proceedings of the 38th annual conference of the cognitive science society* (p. 2615-2620). Austin, TX: Cognitive Science Society.

Winters, J., Kirby, S., & Smith, K. (2018). Contextual predictability shapes signal autonomy. *Cognition*, *176*, 15–30.