Title
The Complementary Importance of Static Structure and Temporal Dynamics in Teamwork Communication

Journal
Human Communication Research, 44(4)

ISSN
0360-3989

Authors
Hilbert, Martin
James, Ryan G
Gil-Lopez, Teresa
et al.

Publication Date
2018-10-01

DOI
10.1093/hcr/hqy008

Peer reviewed

The Complementary Importance of Static Structure and Temporal Dynamics in Teamwork Communication

Martin Hilbert*, Ryan G. James[+,#], Teresa Gil-Lopez*, Ke Jiang*, Yining Zhou*

University of California, Davis

*Department of Communication; [+]Complexity Sciences Center; [#]Physics Department

Accepted in Human Communication Research (May 2018)

Hilbert, M., James, R. G., Gil-Lopez, T., Jiang, K., & Zhou, Y. (2018). The Complementary Importance of Static Structure and Temporal Dynamics in Teamwork Communication. *Human Communication Research*. https://doi.org/10.1093/hcr/hqy008

ABSTRACT

Communicative exchanges consist of a certain degree of both static and dynamic structure that can be used for prediction. Temporal dynamics are often neglected in communication studies. We use Shannon's mathematical theory of communication to examine 1,594,690 contributions from 206,184 contributors to 38 open online collaborations. We find that about three fourths of the total predictability of turn taking stem from participation frequencies ('static variance'), while one fourth originates from the temporal sequence ('dynamic process'). Most dynamic structure is contained within consecutive dyads. We find a trade-off in the importance of static and dynamic structure, which we explain with a combination of both theoretical and empirical factors. We also show that the stationarity of the communication process plays a significant role in this trade-off. These findings have implications both for theorizing and methodologically measuring communication as a dynamic process, as well as for the practical design of online collaboration systems.

Acknowledgements: We are indebted to Jim Crutchfield for the continuous exposition to the depth and beauty of information theory and its extensions, as well as to Robert Bell, Peter Monge, Sarah Marzen, Grace Benefield, Harrison Hughes, and reviewers of IC2S2 for helpful comments.

The Complementary Importance of Static Structure and Temporal Dynamics in Teamwork Communication

Communication patterns of turn taking of team communication are complex, adaptive, and emergent phenomena. The basic signature of complex adaptive systems is their situating between structure and randomness: some aspects of them are predictable, while others are inherently random (Bialek et al., 2001; Crutchfield and Feldman, 2003; Crutchfield, 1994; Grassberger, 1986; Kolmogorov, 1959). We use an approach from the literature of complex systems, which has grown out of Shannon's (1948) mathematical theory of communication, to quantify the amount and kind of predictability in turn taking of teamwork communication. The goal is to deepen our understanding of its origins, constituents, and involved trade-offs.

Static Structure and Dynamic Patterns

The literature differentiates between variance- and process-based explanations of communication patterns (Barnett, Chang, Fink, and Richards, 1991; Monge et al., 1984; Poole, 2007; Poole et al., 2000). The former emphasizes relations among different variables during a given time window and the latter among variables at different points in time. Today, the vast majority of the literature analyzes variables or networks derived from a static snapshots or a time-collapsed sequence of communication processes (Keegan, Lev, & Arazy, 2016; Leenders, Contractor, & DeChurch, 2016; Monge et al., 1984; Pilny, Schecter, Poole, & Contractor, 2016). This answers the 'who speaks how often' question, but neglects the 'when' question. For example, we can calculate the share of contributions of users in Wikipedia and find that 10% of users make more than 90% of the contributions (Ortega, Gonzalez-Barahona, and Robles, 2008). This provides predictability: it is quite likely that the next edit will come from a top power-user. In this case, the entire editing history is seen as one static event.

The process-based approach recognizes that "communication is a process and should be explained as such" (Poole, 2007, p. 181). For example, we might notice that the contribution of a sporadic ad-hoc user might make it more or less likely that a power-user gets involved. Here we use dynamical patterns for predictions, conditioning on temporal sequences. The predictability of the next communicative turn depends on the particular historical context that immediately precedes and currently frames the present exchange.

In the early 2000, less than 10% of articles in 40 communication journals dealt explicitly with temporal dynamics (Poole, 2007). The current neglect of temporal dynamics in the communication literature exists despite a focus on processes in early works (Berlo, 1960; Schramm, 1955), and despite a big push in the 1970s and 1980s to use dynamical systems theory, which included the study of group decision-making (Ellis and Fisher, 1975; Fisher, 1970; Krain, 1973; Poole and Roth, 1989), relational control in relationships (Ellis, 1979; Fisher and Drecksel, 1983; Hawes and Foley, 1973), mass communication (Watt and VanLear, 1996), and talk and silence sequences in conversations (Cappella, 1979, 1980; Cappella and Planalp, 1981). A main historical account for why this work was discontinued was that “gathering everyday conversations… is nearly impossible…. unless one carries a tape recorder around all day (a cumbersome and hardly practical endeavor)” (Fisher and Drecksel, 1983, p. 68). Additional culprits are the “adoption of approaches from other fields such as psychology that do not emphasize process as much as communication” (Poole, 2007, p. 181), “the perceived scope of effort required from the researcher” (Monge et al., 1984, p. 28), and that dynamics were “simply impractical to compute” (Attneave, 1959, p. 22) before today’s computing power.

An Opportunity

Today’s digital reality provides the opportunity to deepen our understanding of communication as a process with fresh data and unconventional methods. The digital footprint allows us to obtain sequential communication patterns easily and computational power enables unorthodox techniques that only provide meaningful statistics with such ‘big’ data. We work with 38 open computer-supported collaborations on crowdsourcing platforms, including Wikipedia, GitHub, and OpenStreetMap, with a median of 23,456 conversational turns per collaboration project.

Building the Vocabulary of Turn-taking

How can we capture dynamic patterns? We need to create the vocabulary of the dynamic that builds a seamless bridge from a static snapshot to a dynamic process. We start by defining the shortest (static) unit. There are many possible ways to classify discrete components of a dynamical communication exchange. Goffman (1981; p. 5) used “Replies and Responses”. Fisher (1970) developed two dozen interaction categories, and Jurgens and Lu (2012) use thousands of different types. For reasons of simplicity and tractability, we simply coded for conversational turns of participants, following the examples of Butts (2008), Gibson (2003, 2005), and Keegan et al. (2016). In order to calculate the static frequency of the distribution, we simply count how often different communicators like A and B contribute.

In order to introduce dynamics, we analyze blocks of sequences over time. If we only have two communicators, A and B, we have exactly four possible blocks of length two: AA, AB, BA, BB. We can analyze how often each of these blocks occurs. We might find that some of them never occur, while others are quite frequent. Knowing this increases our predictability of temporal patterns of two consecutive turns. For sequences of three consecutive turns, we have eight possible turn motifs, and for four consecutive turns, sixteen, etc.. The result is a dictionary whose vocabulary are the existing temporal motifs.

Creating the frequency statistics of these motifs provides for temporal predictability. If AB and BA occur more frequently than AA and BB we can predict temporal reciprocity in turn taking. In the more colorful words of more recent literature, such “sequential structural signatures” (Leenders et al., 2016, p. 98) provide the “meaningful unit of analysis” (Keegan et al., 2016, p. 1070) that results in the temporal “rhythms that give it structure” (Begole, Tang, Smith, and Yankelovich, 2002, p. 334). Existing literature often chooses the length of such motifs ex-ante, such as two- (Fisher & Drecksel, 1983), four- (Gibson, 2003, 2005), or five consecutive turns (Keegan et al., 2016). Since this seems a bit arbitrary, we use a systematic and data-driven approach and test for the importance of all possible sequences of length 1, 2, 3, 4, etc. This leads to our first concrete research question:

> Research Question 1 (RQ1): What is the contribution of communication motifs of different length to the total predictability of the overall communication pattern?

Quantifying Structure and Process

The statistics of the blocks of concatenated conversational turns allows us to quantify the roles of static structure and temporal process. The distribution among the motifs of length 1 provide the static frequencies: how often does each communicator contribute? From a static perspective, which assumes that the process is independent and identically distributed (i.i.d.), what makes communication predictable is the skewness of the distribution of individual contributions. If we have an upper limit of total predictability of the pattern, we can calculate how much of the total predictability stems from the skewedness of the collapsed time series, and how much from the vocabulary of concatenated conversational turns in time. Formal theorems from information theory allow us to ask such question:

> Research Question 2 (RQ2): How much of the total predictability of large-scale communication structures originates from the static distribution of communication frequency and how much from the dynamic temporal sequence?

Changing Dynamics of Temporal Patterns

One of the main opportunities of today's massive datasets is that it allow us to delve deeper into profound conceptual differences, such as the two complementary notions of generalizability. From a static perspective, phenomena are detected and generalizable across a sufficiently large number of similar cases. If two teams are sufficiently similar, we can predict one from the other. From a dynamic perspective, phenomena are detected and generalizable across a sufficiently large number of repetitions across other periods. It two periods are sufficiently similar, we can predict one from the other. This implies two kinds of scientific generalizability. The generalization among static cases requires the *ceteris paribus* assumption: the cases need to be sufficiently similar. The generalization among temporal events requires the stationarity assumption: the periods need to be sufficiently similar. Stationarity demands that general summary statistics do not change throughout the time series: the dynamic persists over time.

What complicates things is that social dynamics are almost always changing while unfolding. Their non-stationarity originates endogenously, through life cycles (Ellis and Fisher, 1975; Fisher, 1970; Poole and Roth, 1989) or the ambition of social agents to improve the process (Hilbert, 2014; Madsen, Flyverbom, Hilbert, and Ruppert, 2016), or exogenously through external shocks (DeDeo, 2016; Poole et al., 2000). Given the useful properties of stationarity and its rare appearance, it is often swept under the methodological rug and no statistical tests are executed, not even when published under the scrutiny of the world's most highly ranked Journals, (Song, Qu, Blumm, and Barabási, 2010). We confront the topic head-on and ask:

> Research Question 3 (RQ3): How do static structure and temporal dynamics relate to the level of stationarity of the communication process?

Method: the Mathematical Theory of Communication

The notion of using motif dictionaries for the identification of structure in communication was formally generalized by Shannon's (1948) mathematical theory of communication. He mainly explored statistical patterns of block codes to calculate the "Prediction and Entropy of Printed English" (Shannon, 1951, p. 50). He called the blocks "n-gram" (Shannon, 1948, p. 387), which is a term that found its way into modern content analysis and computational linguistics. Today, Shannon's comprehensive intellectual framework is known as "information theory" (for a general introduction, see Gleick, 2011, and Pierce, 1980; for a more technical treatment, see Cover and Thomas, 2006, and MacKay, 2003).

We use this framework to calculate the predictability of communication patterns in open teamwork. Just like individual letters make up more or less frequently appearing words (while certain sequences of letters never appear), we look for empirical evidence for certain 'words' that emerge from the 'letters' representing the contributions of different communicators. Such adoptions of information theory to distinguish among structural and random components of (nonlinear) dynamics is quite common in statistical mechanics, neuroscience, and dynamic systems theory in general (Bialek, Nemenman, and Tishby, 2001; Crutchfield and Feldman, 2003; Crutchfield, 1994; Crutchfield and Packard, 1983; Grassberger, 1986; Kolmogorov; 1959). Recently, this analytical approach has been applied to study the predictability of human mobility, showing that next location of individuals is up to 93% predictable (Lu, Bengtsson, & Holme, 2012; Smith, Wieser, Goulding, & Barrack, 2014; Song et al., 2010).

Measures: Basic Quantities of Information Theory

We will work with the most fundamental measure of information theory: entropy. We then calculate its rate and apply it to a mathematical theorem to obtain the overall predictive limit of a dynamic (Cover and Thomas, 2006, Chapter 2).

Entropy. Entropy is a measure of uncertainty and randomness, and therefore also for structure and predictability. Entropy is at its maximum if we are faced with a uniform distribution of possible events (maximal uncertainty) and at its minimum if there is only one possible choice (no uncertainty and perfect predictability). In our case, it is a summary measure of the uniformity or skewedness of the communication components. Our approach to estimating the entropy of our time series is somewhat refined from the black-box approach used in recent studies of human mobility (e.g. Song et al., 2010). Since we want to understand the accumulative importance of motifs with different length, we use a sliding window with increasing lengths (Crutchfield and Feldman, 2003), more in line with early information theory applications (Attneave, 1959, p. 25; Miller and Frick, 1949).

As shown in Figure 1, we slide a window of a specific length over the entire time series, one component at a time. We will start with a window of length 1, and increase the window length to 2, 3, 4, 5, etc. We denote a sequence S of L consecutive variables by S^L. Unique blocks of length L give us our different motifs. We then identify the probability distribution that different motifs occur in the overall time series, $P(s^L)$, and calculate its entropy $H(L)$, which is defined as:

$$H(L) = -\sum_{s^L \in A^L} P(s^L) * log_2 P(s^L) \qquad (1)$$

The sum in equation (1) runs over all possible blocks of L consecutive symbols (Figure 1a); A^L is the set that consists of all motifs of length L (concatenated words from single contributions represented with an alphabet A), and since we use the logarithm of base 2, entropy is measured in bits. $H(L)$ is never negative and at its maximum if all different motifs s^L are equally likely.

Entropy rate. We then take the discrete differences of entropies of different motif length L. This represents the uncertainty captured by motifs that are one symbol longer.

$$\Delta H(L) = H(L) - H(L-1) \qquad (2)$$

The unit of $\Delta H(L)$ is bits/symbol, or in our case, bits per turn. It is never negative: $H(L) \geq H(L-1)$. $\Delta H(L)$ is also called the entropy gain, since it is the uncertainty we gain when working with longer motifs (Crutchfield and Feldman, 2003).

Following the logic of order-R Markov processes, $\Delta H(L)$ can also be written as a conditional entropy of the distribution of motifs with length L, conditioned on the distribution of motifs with length *L-1*: $\Delta H(L) = h_\mu(L) = H(S_L|S^{L-1})$ (Cover and Thomas, 2006). In this equivalent approach, we calculate the probability of the next turn conditioned on the previous motif. For motifs with a certain length L, $h_\mu(L)$ is the estimate of the remaining randomness (uncertainty) of the pattern if the structural (predictable) components of block length L are considered (Crutchfield and Feldman, 2003). In words, $\Delta H(L = 4) = h_\mu(L = 4)$ asks: given all occurring motifs of length 4, what is the average uncertainty of the next turn?

Entropy curves. By calculating the average uncertainty of the occurrence of motifs with increasing lengths, we expand our dictionary of possible motifs and aim at capturing ever more structure in the time series. The structure is the amount of information contained in the specification of the motifs in the arising dictionary of sub-sequences, while the irreducible randomness is contained in the remaining entropy.

Example. For reasons of pedagogical illustration, let us assume two communicators, A and B. Without anything else given, entropy is 1 bit (50% chance of each contributing). Let us assume a simple communication process that follows the deterministic pattern of ABBB: [ABBBABBBABBB…] (e.g. a question about location, and the answer in the form of the three coordinates of spatial dimensions). In this simple communication pattern, symbol A appears 25% of the time and B makes 75%. This is our static frequency. It results in $H(L=1) \approx 0.81\, bits$ (equation (1)) and therefore reduces uncertainty and provides useful predictability: B is more likely than A. Armed with this insight, our best prediction for the next turn would always be B. We would be right 75 % of the time. Now we also track all motifs of block length two, of which we find three possible sequences: AB (25% of the time), BB (50%), BA (25%). This dictionary of dyads results in $H(L=2) = 1.50\, bits$ (equation (1)). Therefore, according to equation (2), $h_{\mu}(L=2) = H(L=2) - H(L=1) = 1.50 - 0.81 \approx 0.69\, bits$. We find four motifs of block length three (ABB, BBB, BBA, BAB), all equally likely, which results in $H(L=3) = 2.0\, bits$, and also four uniformly distributed motifs of block length four (ABBB, BBBA, BBAB, BABB), which also implies $H(L=4) = 2.0\, bits$. This leads to the fact that $h_{\mu}(L=4) = 0$. There is no uncertainty remaining when considering motifs of block length 4. This is the necessary consequence of a deterministic process: after the period is reached, the process becomes completely predictable, as we capture all of the inherent structure. Figure 1 illustrates that the entropy rate (lower curve) is a discrete derivative of the entropy (upper curve).

Limit of Predictability

We now ask about the probability (Π) that the best predictive algorithm can correctly predict which communicator will make the next contribution. We are indifferent about the specificities of the predictive algorithm. It can consist of (non-)linear extrapolation, Bayesian inference, cutting-edge artificial intelligence with deep learning, or something else. Information theory tells us that regardless of the method, the limit of predictability must be subject to Fano's inequality (Fano, 1961). The inequality relates the probability of error of any prediction, P_e, to the entropy rate, i.e. $h_{\mu}(L) = H(S_L|S^{L-1})$, and provides a lower bound for error (Cover and Thomas, 2006):

$$P_e \geq \frac{H(S_L|S^{L-1}) - 1}{\log|A|} \qquad (3.1)$$

Where $|A|$ is the size of the alphabet, in our case, the number of different contributors among which we try to predict the next one. Since the error is always larger than the right-hand side of equation (3.1), no predictive algorithm can be better than our limit of predictability Π:

$$\Pi = 1 - P_e \qquad (3.2)$$

Separating Atemporal and Temporal Uncertainties

We calculate the limit of predictability for three kinds of entropies. First we estimate the predictability Π^A when merely knowing the number of different communicators. We know neither their frequencies nor any temporal pattern and therefore assume all communicators make the same amount of contributions in a random order (maximum entropy): $h_\mu^A(L = 0) = log_2|A|$.

Next, we estimate the predictability when knowing the distribution of contributions, but not their temporal pattern. Predictability Π^H is calculated by $h_\mu^H(L = 1) = H(L = 1)$. This measures the entropy of single symbols, and entropy is a measure of the skewness of a distribution. By assessing the mere distribution, it essentially estimates the predictability of a static aggregation of what is actually a dynamic conversation (akin to a static variance of contributions). The more skewed the distribution of contributions, the more predictable.

Last, we use our sliding-window approach on the empirically recorded time series and calculate the limit of predictability when considering both the distribution and the temporal order, Π, by calculating $h_\mu(L = max)$. Naturally, uncertainties will decrease with the consideration of more detailed patterns ($h_\mu \leq h_\mu^H \leq h_\mu^A$), while predictability will increase: $\Pi \geq \Pi^H \geq \Pi^A$.

Stationarity and predictability

While econometric time series analysis has developed a solid theoretical framework to tests for stationarity in scalar, or at least ordinal variables (Wooldridge, 2008), communicative turns consist of categorical/qualitative data. We follow an approach developed to test the stationarity of transition probabilities of Markov chains (Anderson and Goodman, 1957). It has previously been applied to communicative turn-taking (Cappella, 1980) and consists of a χ^2 test among the first-order distributions of contributions of different periods (that is, for word length 1). This previous research has cut the total period in rather qualitatively defined sub-periods and tested subsequent pairs. Being equipped with more computational power nowadays, we suggest cutting the total time series in as many sub-periods as allowed by the χ^2 test and test all possible combinations of pairs of sub-periods. The number of possible sub-periods g is defined by the length of the time series $|T|$, the alphabet size $|A|$, and the χ^2 test demand that (at least 80% of) the expected count in the resulting matrix is over 5 (Yates, Moore, and McCabe, 1998, p. 734):

$$g = \frac{|T|}{|A| * |A| * 5} \qquad (4)$$

Data: in Need of 'Big' Data

The length of analyzable communication motifs is limited by data availability. Given the exponential growth of the combinatorial possibilities of how symbols create motifs (or 'letters create blocks of words'), datasets of considerable size are required to achieve minimum statistical representativeness. Past research has limited the number of considered communicators distinguishing among some 3-4 participants (Gibson, 2003, 2005) and worked with communication motifs of maximal block length 3 (e.g. Fisher, 1970; Poole and Roth, 1989). For sequence length $L = 4$, when distinguishing among merely 7 different communicators (e.g. A, B, C, D, E, F, G), the guarantee of at least 10 appearances of each possible motif already requires a time series with more than 24,000 sequential records. The required length T of the time series is affected by both the diversity of the alphabet $|A|$ (distinct communicators) and the length L of the block sequences (length of motifs) through the following equation:

$$|T| = n * 2^{log_2(|A|)*L} \qquad (5)$$

where n is the estimated number of times each motif will appear in the time series (the sampling of words). Equation (5) specifies the worst-case scenario, as it assures that each motif should appear at least the indicated number of times n. It could be relaxed by replacing $log_2(|A|)$ with h_μ (Marton and Shields, 1994).

We tracked the largest open teamwork datasets we could publicly find online since the beginning of their existence until February 2016. We analyzed the 10 largest GitHub projects (a web-based repository hosting a version control system for software development); the 2 largest OpenStreetMap collaborations (a collaborative platform to create free and editable maps); and 26 Wikipedia pages (a collaborative online encyclopedia), including the10 largest open- and the 16 largest semi-protected pages (participation requires an (auto)confirmed account) (Table 1).

It is important to point out that the nature of these group dynamics is different from most face-to-face decision-making processes in teams. First, these groups are open to the general online public. Second, they are big and with growing group size, the likelihood of parallel communications increases. Third, they are asynchronous, in a sense that anybody can chime in on any topic at any time. Someone might hold a long monologue without bothering anyone, in stark contrast to offline teamwork. For example, for GitHub we track ‘commits’, which keeps record of what changes were made to the collective product, when and by who (usually containing a brief description) and for Wikipedia we track ‘revisions’. Of course, also face-to-face communication can consist of parallel conversations and untethered monologues, but the decentralized and asynchronous nature of these computer-supported collaboration platforms provides more opportunity for flexibility, and therefore, for more uncertainty and potentially less predictability.

The mean of the number of contributions is 41,966 and the median, 23,456. On average, 5,426 different contributors communicate per project, with a median of 5,971. Naturally, an alphabet size of over 5,000 different symbols would not provide representative statistics in agreement with equation (5). We analyze three different coarse-graining mechanisms that combine certain kinds of communicators into groups.

Table 1. Included databases.

	Platform	Project	contributions	contributors
1	GitHub	Linux	573,066	14,641
2		Homebrew	60,242	6,371
3		Rails	55,914	3,553
4		Rust	49,666	1,445
5		Swift	31,944	314
6		Atom	27,404	316
7		Go	26,032	716
8		gitlabhq	23,144	1,009
9		django	22,007	1,198
10		docker	21,583	1,411
11	OpenStreetMap	Russia	34,034	90
12		Germany	18,587	259
13	Wikipedia Semi-protected	George W. Bush	45,875	14,456
14		United States	35,800	9,789
15		Wikipedia	33,992	13,209
16		Michael Jackson	27,677	6,522
17		Jesus	28,112	6,788
18		List of programs broadcast by ABS-CBN	25,239	5,409
19		Barack Obama	24,740	6,648
20		Adolf_Hitler	24,631	8,272
21		World War II	23,767	7,534
22		India	22,318	6,402
23		The Undertaker	22,196	7,207
24		Wii	21,743	7,022
25		United Kingdom	21,635	6,621
26		Roger Federer	20,993	7,024
27		FC Barcelona	20,702	6,099
28		Global warming	20,544	4,693
29	Wikipedia Open	List of WWE personnel	42,610	4,574
30		Catholic Church	26,523	6,108
31		List of Ben 10 aliens	21,077	5,843
32		ATP World Tour records	20,733	2,679
33		2006 Lebanon War	20,532	4,144
34		Kane (wrestler)	20,401	8,045
35		Jehovah's Witnesses	20,369	5,388
36		List of programs broadcast by GMA Network	20,312	4,018
37		PlayStation 3	20,069	6,609
38		List of Total Nonstop Action Wrestling personnel	18,477	3,758

We always coarse-grain according to the same schema: we track each one of the $(|A| - 2)$ most prolific communicators. For example, we distinguish among each one of the top 5 communicators with the most contributions. We then group all those communicators with only 1 and 2 contributions as one single group of communicators (ad-hoc user), and the remaining users into another group. As a result, we distinguish among $|A| = 7$ different entities: the five most prolific users, and two aggregated groups. Applying the sizes of our datasets to equation (5), using 7 different communicators ($|A| = 7$), allows for representative motif lengths between $L = 4$ and $L = 6$. Distinguishing among the top 15 contributors ($|A| = 17$), we obtain representative motif lengths of 3 and 4 consecutive turns, and when distinguishing among the 35 most prolific contributors ($|A| = 37$), we may merely work with motif length between 2 and 3 turns.

Results

RQ1: Length of Communication Motifs

The importance of motifs of different length (RQ1) is best seen with the example our largest database, the "Linux" project on GitHub (573,066 consecutive turns). It counts an average of about 39 interventions per contributor, and the most prolific contributor, the Linux creator Linus Torvalds, taking the proverbial talking stick some 3.5% of the time.

Figure 1c shows the resulting entropy curves for a coarse-graining to $|A| = 7$. The theoretically maximum uncertainty among seven uniformly distributed contributors is $\log_2(7) \approx 2.8$ bits per turn. We ask how much the consideration of temporal motifs can reduce this uncertainty. Following equation (5), the size of the dataset assures an adequate statistical representation of n occurrences per motif for motif lengths up to $L = 6$ (see the vertical dashed lines in Figure 1c). The biggest drop in uncertainty reduction happens when we expand our motif from length 1 to length 2 ($h_\mu(L = 1) = 0.68$, $h_\mu(L = 2) = 0.53$). This is in agreement with previous findings (Dabbs & Ruback, 1987; Ellis, 1979; B. A. Fisher & Drecksel, 1983; Keegan et al., 2016; Krain, 1973; Parker, 1988; Stasser & Taylor, 1991). After that, we see that the entropy rate converges to about $h_\mu = 0.49$ bits per turn. This is the intrinsic randomness of the process that cannot be explained with our data. The uncertainty reduction contributed by expanding the motif dictionary from lengths 4 to 5, and from 5 to 6 is merely some 0.006 bits per symbol. This implies that there is more structure in shorter subsequences, while longer motifs contribute with sharply decreasing returns. It might as well be that we would discover more structure in longer motifs (maybe some recurring cycle that repeats every few dozen contributions?). The size of our dataset does not allow us to explore this in a meaningful way (increasing the length of motifs will quickly lead to overfitting).

Figure 1: (a) Sliding window of block length-4 sequences: S^4. (b) Entropy curves with increasing motif length L, for period-4 process: [ABBBABBB…]. Entropy curve for block entropies $H(L)$ and entropy rate with convergence $h_\mu(L = 4) = 0$. (c) Application to the Linux project on GitHub until 01/22/2016, with $|A| = 7$. Non-deterministic process with remaining uncertainty.

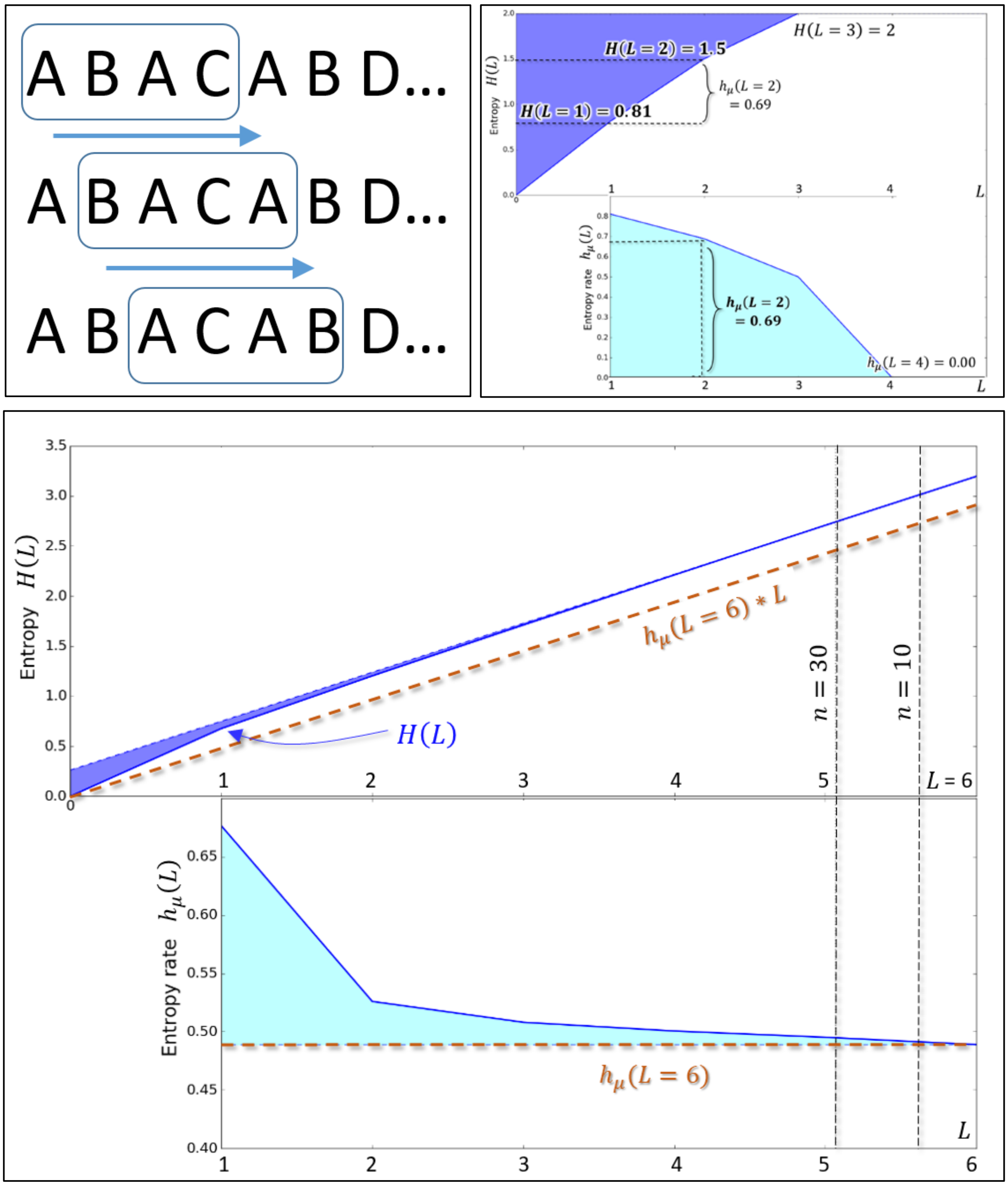
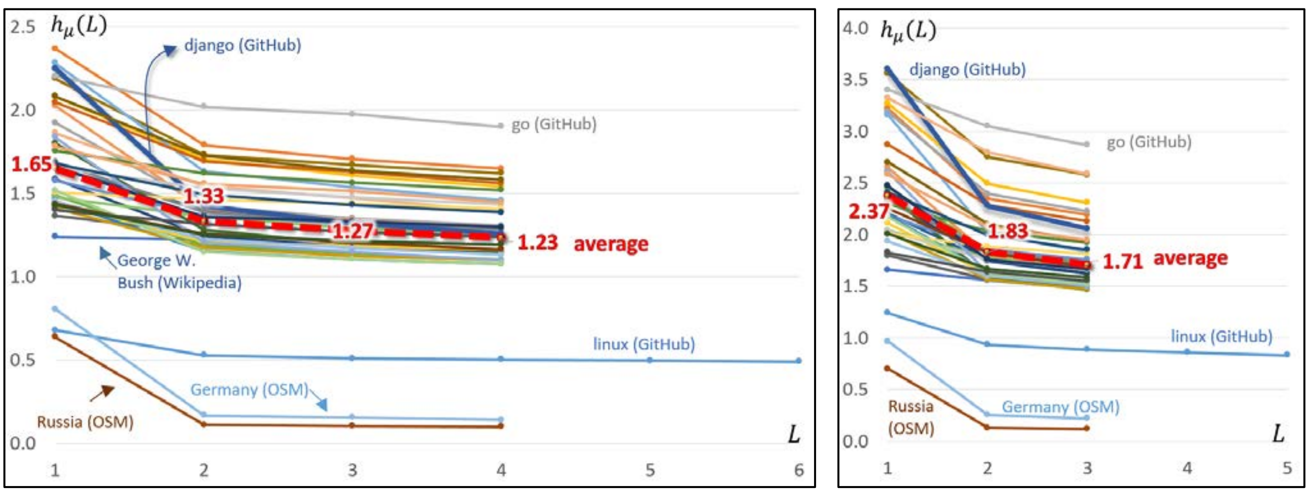


Figure 2 shows the entropy rate curves for our 38 datasets, for coarse-grained processes among 7 communication units (Figure 2A), and 17 communication units (Figure 2B). On average, without any temporal information ($h_\mu^H(L=1) = H(L=1)$), the distribution among 17 different communicators represents 2.37 bits of information (Figure 2B). That is, simply by knowing the skewness of the distribution, we already know that from the possible 17 communicators, any of about $2^{2.37} \approx 5.2$ do the talking. Static skewness reduced our uncertainty from 17 possibilities to about 5. Considering motifs of sequence length 2, we can reduce our uncertainty to any of $2^{1.83} \approx 3.6$ individuals, and with motifs of length 3, we can specify that at any given point it is the communicational turn of any one of $2^{1.71} \approx 3.3$ participants. As shown in Figure 2, our response to RQ1 is that largest process contribution to communicative structure consists in the registration of the dyadic exchanges (from motif length $L = 1$ to $L = 2$).

Figure 2: Motif length L against entropy rates for the 38 databases from Table 1, (a) with coarse-graining to $|A| = 7$, and (b) coarse-graining to to $|A| = 17$.



RQ2: Frequency-, Sequence-, and Total Predictability,

We now employ Fano's inequality from equation (3) to calculate the limits of predictability Π, which is what we need to meaningfully quantify the contributions of static frequencies and temporal dynamics (RQ2). Figure 3 shows the results for our different level of coarse-graining. Naturally, the next contributor is easier to predict if we distinguish among less contributors $|A|$.

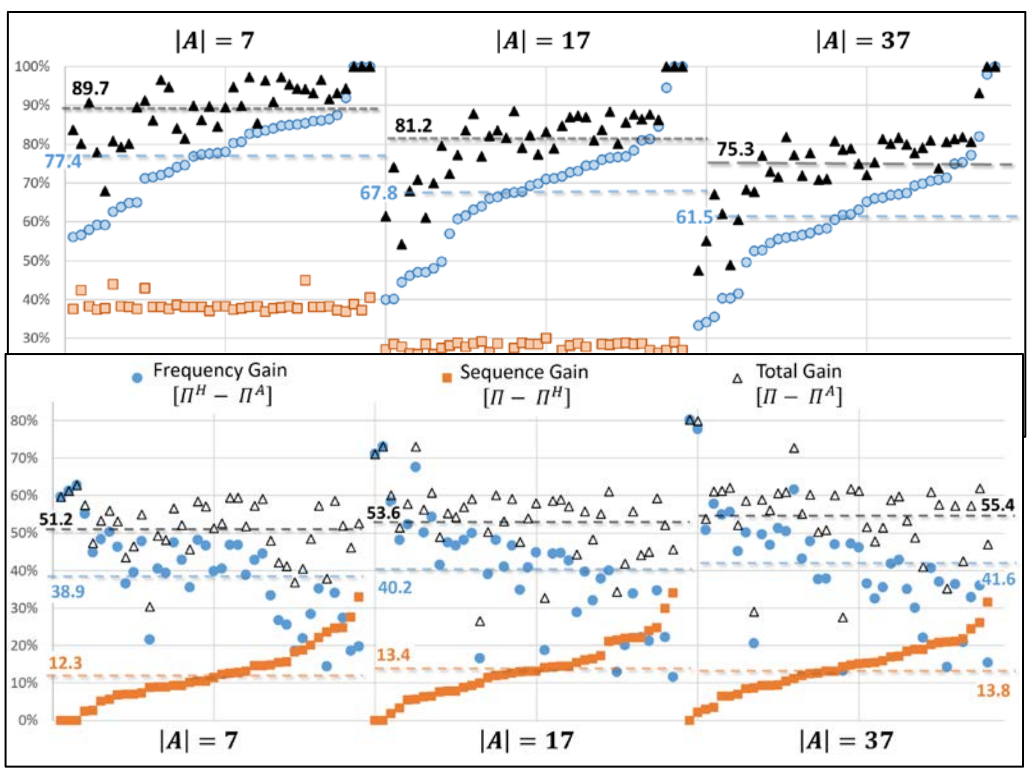
When distinguishing among 37 different groups, the best prediction algorithms cannot predict the next contribution with better than 75.3% accuracy (black triangles in Figure 3a). When coarse-graining among 17 different contributing units, the best predictive algorithms cannot exceed an accuracy of 81.2%. When coarse-graining to 7 different communicative units, we get a limit of maximal predictability of $\Pi = 89.7\%$. In other words, with almost 90% accuracy we can predict if the next contribution will come from one of the top-5 power-users, from a sporadic ad-hoc contributor, or from the group in-between. This shows that a considerable amount of communicative structure emerges in large-scale collaborative teamwork. Figure 3a also shows the constituents of this total predictability, the empirically derived baseline predictability Π^A and the static frequency contribution Π^H.

Figure 3b takes the differences between these constituents of total predictability. The combined total predictability gain is the difference between our baseline and the total limit of predictability $[\Pi - \Pi^A]$. It is between 51.2% and 55.4%. This total consists of the sum of the contribution of static frequencies, $[\Pi^H - \Pi^A]$, plus the predictive gain from temporal sequence: $[\Pi - \Pi^H]$. Figure 3b shows that static frequencies contribute with a predictability gain of some 40% (or some ¾ of the total), while the consideration of temporal sequence contributes an additional 13% (the remaining ¼ of the total predictability).

Figure 3: Comparing for $|A| = 7$; $|A| = 17$; $|A| = 37$. (a) Limits of predictability Π, baseline Π^A (knowing the alphabet), and Π^H (knowing the static distribution). (b) Predictive gains of static frequencies: $[\Pi^H - \Pi^A]$; temporal sequences: $[\Pi - \Pi^H]$; and combined total: $[\Pi - \Pi^A]$.

(a)

(b)



A quite surprising finding is that these predictability gains are very similar among the different level of coarse-graining: ¾ and ¼ of the total predictability. While the averages of the obtained limits of predictability among the 38 projects from Figure 3a are all significantly different (all paired sample t-tests $p \leq 0.000$), the averages in Figure 3b are not (with the sole exception of the difference of the total gain between $|A| = 17$ and $|A| = 37$).

A closer inspection of Figure 3 shows that the predictability gain based on frequencies, $[\Pi^H - \Pi^A]$, correlates positively with total predictability gain, $[\Pi - \Pi^A]$ ($r(112) = 0.865$, $p \leq 0.000$), while the predictability gain based on the sequence, $[\Pi - \Pi^A]$, correlates negatively with the total predictability gain ($r(112) = -0.339$, $p \leq 0.000$). The higher the gain in predictability due to a skewed frequency distribution, the lower the gain in predictability due to temporal sequence. There is a certain trade-off in predictability gains.

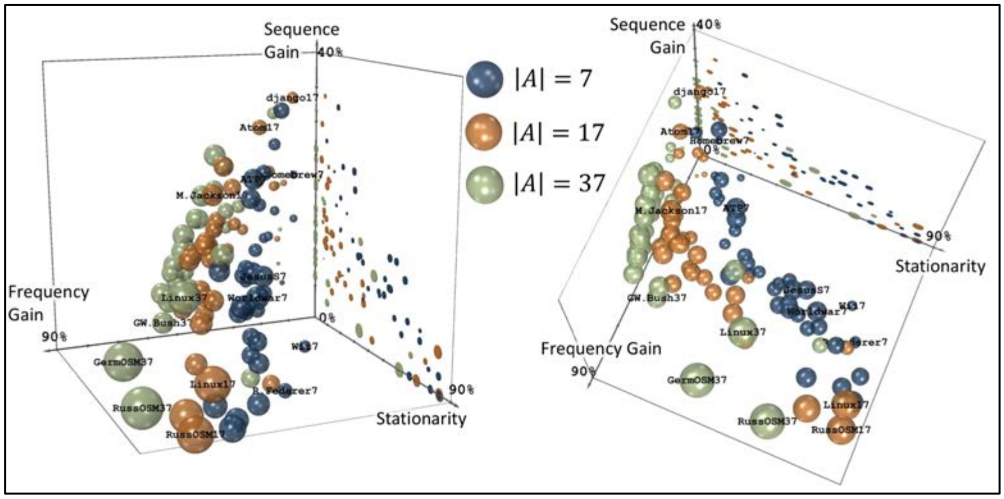
Therefore, we can answer and even fine-tune RQ2. On average, static structure contributes about ¾ to the predictability of teamwork turns and temporal patterns ¼. This is the average of a trade-off between predictability from static structure and temporal dynamics.

RQ3: Frequencies, Sequences, and Stationarity

Figure (4) combines the frequency and sequence gains with the results for stationarity tests. We obtain the percentage of stationary sub-periods by counting how many of those pairs are not significantly different at $p \geq 0.01$. For example, Linux can be meaningfully tested for stationarity when divided into 2,339 equal length sub-periods for $|A| = 7$ (resulting in 2,734,291 χ^2 tests among possible combinations), and in 84 groups for $|A| = 37$ (3,486 paired combinations).

The total predictability gain is positively correlated with the stationarity of the time series, although surprisingly weak ($r(112) = 0.309$, $p \leq 0.000$). Stationarity is positively correlated with the predictability gain from static frequencies ($r(112) = 0.530$, $p \leq 0.000$) and negatively correlated the gain obtainable from temporal sequences ($r(112) = -0.598$, $p \leq 0.000$). Figure 4 also visualizes the statistically significant trade-off between predictive gains from static frequency and temporal sequence ($r(112) = -0.765$, $p \leq 0.000$).

Figure 4: Predictive gains vs. stationarity. Bubble size shows total predictability gain: $(\Pi - \Pi^{A})^2$



Discussion

We found that static first-order variance contributes three times as much predictability in open computer-supported collaborations on crowdsourcing platforms than higher order temporal processes (¾ - ¼); that there is trade-off between the predictability gains from static structure and temporal dynamics; and that static structure predicts more strongly in a stationary process and dynamics in a nonstationary process. The following discusses each finding.

The ¾ - ¼ Rule of Predictability in Collaborative Turn Taking

The fact that we found a surprisingly stable share of predictability gains from static and temporal signatures independent from the level of course-graining has to be qualified by three issues: the chosen indicators, the chosen coarse-graining, and the chosen datasets.

Our turn-taking indicator could be complemented by additional externally related variables (e.g. socio-economic attributes of the communicator or exogenous events) (Butts, 2008; Pilny et al., 2016), or by specifying content type (e.g. length or meaning of contribution). This could explain more or different kinds of structures and therefore change the finding.

Our chosen coarse-graining might also confound this result. Having a closer look at the data reveals that static skewedness originates for different reasons: in OpenStreetMap projects from the dominance of one power-user, in the Linux project, from the dominance of the diverse group of regular contributors, which are neither power-users, nor sporadic ad-hoc contributors. This latter case is an artifact of joining different users into a common group (dealing with equation (5)). Figure 4 reveals that more fine-grained approaches increase the contribution of temporal dynamics within the Linux project (compare Linux17 and Linux37 in Figure 4).

As for the chosen datasets, we worked with the contributions or commits of open computer-supported collaborations on crowdsourcing platforms. Of these, we chose the cases with the most contributions (again, dealing with equation (5)). It is important to notice that those two consecutive Wikipedia or GitHub contributions do not necessarily have to be connected. This leads to the longstanding question of the importance of coherence in group communication (Poole, 1985). “The claim that a group's communication plays a significant role in the outcomes of its discussion presupposes that the members of the group were in actuality communicating with one another” (Pavitt & Johnson, 1999, p. 303). We did find that the vast amount of structure is contained in the dyadic exchange between two consecutive conversational turns (RQ1), which is in agreement with previous findings (Dabbs & Ruback, 1987; Ellis, 1979; B. A. Fisher & Drecksel, 1983; Keegan et al., 2016; Krain, 1973; Parker, 1988; Stasser & Taylor, 1991). Given that these open and large-scale platforms allow for asynchronous parallel communication, it is not always guaranteed that two consecutive turns constitute a reply and response dynamic among different participants. In some cases, we find that consecutive dyads more often originate by two consecutive contributions of the same user. For example, in the case of the Russian map, the most prolific user makes on average 127 consecutive contributions. At one point, the user holds on to the proverbial ‘talking stick’ for 3,062 consecutive turns. Such concatenated contributions could be collapsed, and future studies should look into the arising differences. Our presented methods lends itself for the detection of differences. As for our study, we set out not to collapse dynamics, and asked about predictability in temporal processes as observed. If the user feels that these are different contributions, we accept this choice as a unit.

The Structure-Dynamic Trade-Off

As for the detected trade-off, we offer an empirical and two mathematical explanations. Taking a closer look at the data reveals that in all 38 empirically analyzed cases the amount of sequence gains is quite stable. The magnitude of their contribution is more similar than frequency gains (over all 38 cases: frequency gain (M = 0.40, SD = 0.14); sequence gain (M = 0.13, SD = 0.07), see also axis in Figure 4). This means that a similar amount of sequential structure is found in communication processes with much and little static skewedness.

The first mathematical reason that converts this empirical particularity into the detected trade-off is that the limit of predictability is capped at 100%. In those cases where almost all predictability originates from the static skewedness, such as in the Russian and German OpenStreetMaps and Linux17 (see Figures 2 and 4), the most prolific group makes some 90% of the contributions. If first-order static structure already provides much predictability, second-order dynamic structure cannot provide much more. In general, if static frequency already explains more than half of the total predictability, the math assures that temporal dynamics cannot add more than the other half of the total 100%.

On the contrary, if the frequency of contributions is more uniform, the stable amount of temporal structure accounts for a lot, in relative terms. In django_|A|=17, the most prolific contributor makes 13% of the contributions, the second and third both 9%, the fourth 8%, etc. This static frequency distribution is quite uninformative, and only contributes 13% to the predictability gain. The tracking of temporal motifs, however, contributes 34%, mainly stemming from consecutive dyads ($h_{\mu}(L = 1) = 3.6\ bits$, and $h_{\mu}(L = 2) = 2.3\ bits$) (Figure 2b). A stable amount of dynamic structure is relatively larger in a smaller total predictability.

The second mathematical reason is more subtle. A few dominant power-users result not only in much static predictability, but also in less temporal patterns, because the same power-users are involved in almost all the patterns, thus limiting the number of possible patterns. A more uniform distribution among individual contributions will result in richer set of occurring temporal combinations. This can increase the role of dynamic signatures. Both of these reasons show that the general trade-off between static and dynamic structures is actually to be expected.

Stationarity

Stationarity is the elephant in the room of dynamical analysis of categorical variables. Our large datasets and today's computational power allowed us to confront it head-on and we found that stationarity plays an important role in our trade-off: the more stationary, the less important temporal dynamics, and the more important static frequencies (Figure 4).

This is to be expected, as stationary processes enable skewed participation structures to carve out its peaked distribution over time. Two periods are likely to be similar if both are dominated by the same power-user, who is the same power-use that also carves out the skewedness of the static frequency distribution. On the contrary, a changing time series dynamics result in a more uniform distribution through mixing. For example, consider the pseudo time series AAABAAAB BBBABBBA. If the dynamic of the first half would continue, power-user A would dominate and create static predictability. However, stationarity breaks in the middle of the series, which flattens the distribution of contributions. Over the entire period, the static structure is most uninformative: 50% chance for A and 50% for B. This kind of mixing leads to the proportionally smaller contributions of static structure in less stationary processes (Figure 4). Additionally, if the first half would continue, we would never see the temporal patterns BB, BBB, BBA, BAB, and ABB. Less stationarity produces more diversity in temporal sequences. Combined, these two effects result in temporal dynamics playing a more important role in less stationary processes, and static statistics in more stationary processes.

While this was again a mathematical reason, the empirically detected universal persistence of temporal motifs through dyads can also play a role. Continuing with our non-stationary pseudo time-series, we clearly have much dyadic structure throughout, especially AA during the first half and BB during the second. We see the same signature in our data. For example, the GitHub django project was mainly driven by a small and cohesive group during its first half (2005-2012), and by one/two-time contributors during the second (2012-2016). This change reduces stationarity, but still provides dyadic temporal structure of two consecutive contributions: power-users during the first half, and sporadic users during the second. Some kind of dyad still dominates the sequence vocabulary, even with low stationarity. In cases where total predictability is small, this represents a larger share of total predictability in relative terms (see the small bubble size of django in Figure 4).

Outlook

Going forward, the detected role of stationarity shows that it is essential to further deepen our understanding of the issue of stationarity for the categorical variables of communicative exchanges (as for example done in DeDeo, 2016). This should consider the length and type of the involved time series. Naturally, more fine-grained time series are also less stationary (see the three almost parallel diagonal alignments in three dimensions in Figure 3).[1]

Second, the presented approach can be used to study more practical problems of dynamics in group communication. Quantifying the structural and temporal patterns can help developing procedures that aim at increasing teamwork productivity, and in the design of related information systems. For example, our method can contribute to the search for the optimal sweet spot between hierarchical structure and decentralized randomness in digitally enabled peer production (Kreiss, Finn, & Turner, 2011; Shaw & Hill, 2014). A better understanding of stationarity in social communication dynamics also allows to tackling potential lock-in effects from the application of machine learning to time series data. If we assume that the past is equal to the future, machine learning is useful. If not, filter bubbles and personalized price discrimination can lock humans into their own past (Hilbert, 2014; Madsen et al., 2016).

A technical point is that our method is confined to consecutive act sequences and so will miss sequences in which irrelevant acts intervene between meaningful ones. Future studies could extend the presented methodology to test for the role of intermittent or other irregular structures.

Last but not least, another outstanding question is where these different communication patterns come from. They are surely the result of a mix of mutually entangled social-, communicative-, and psychological characteristics, such as local and global coherence (Pavitt & Johnson, 1999), habits (Fisher, 2004), reciprocity in trust building (Ostrom and Walker, 2003), path-dependencies and reciprocity (Butts, 2008), and the role of exogenous incidents (Poole et al., 2000). Our approach to quantify static and temporal patterns allows to explore if different generative mechanisms can be linked to the emergence of different patterns of communicative structure. As such, our approach contributes to deepening our understanding of the complementary importance of static structure and temporal dynamics in communication patterns.

[1] Correlations are $r(36)$ at $p \leq 0.000$, ordered according to $|A| = [7; 17; 37]$: stationarity and frequency gain: [0.881; 0.740; 0.567]; stationarity and sequence gain: [-0.811; -0.695; 0.567].

References

Anderson, T. W., & Goodman, L. A. (1957). Statistical inference about Markov chains. *The Annals of Mathematical Statistics*, *28*(1), 89–110.

Attneave, F. (1959). *Applications of information theory to psychology: A summary of basic concepts, methods, and results* (Vol. vii). Oxford, England: Henry Holt.

Barnett, G. A., Chang, H., Fink, E. L., & Richards, W. D. (1991). Seasonality in television viewing: A mathematical model of cultural processes. *Communication Research*, *18*(6), 755–772. https://doi.org/10.1177/009365091018006003

Begole, J. "Bo," Tang, J. C., Smith, R. B., & Yankelovich, N. (2002). Work Rhythms: Analyzing visualizations of awareness histories of distributed groups. In *Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work* (pp. 334–343). New York, NY, USA: ACM. https://doi.org/10.1145/587078.587125

Berlo, D. (1960). *Process of communication: An introduction to theory and practice*. Harcourt School.

Bialek, W., Nemenman, I., & Tishby, N. (2001). Predictability, complexity and learning. *Neural Computation*, *13*, 2001.

Butts, C. T. (2008). A relational event framework for social action. *Sociological Methodology*, *38*, 155–200.

Cappella, J. N. (1979). Talk-silence sequences in informal conversations I. *Human Communication Research*, *6*(1), 3–17. https://doi.org/10.1111/j.1468-2958.1979.tb00287.x

Cappella, J. N. (1980). Talk and silence sequences in informal conversations II. *Human Communication Research*, *6*(2), 130–145. https://doi.org/10.1111/j.1468-2958.1980.tb00133.x

Cappella, J. N., & Planalp, S. (1981). Talk and silence sequences in informal conversations III: interspeaker influence. *Human Communication Research*, *7*(2), 117–132. https://doi.org/10.1111/j.1468-2958.1981.tb00564.x

Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory* (2nd Edition). Hoboken, NJ: Wiley-Interscience.

Crutchfield, J., & Feldman, D. (2003). Regularities unseen, randomness observed: Levels of entropy convergence. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *13*(1), 25–54.

Crutchfield, J. P. (1994). The calculi of emergence: computation, dynamics and induction. *Physica D: Nonlinear Phenomena*, *75*(1–3), 11–54. https://doi.org/10.1016/0167-2789(94)90273-9

Crutchfield, J. P., & Packard, N. H. (1983). Symbolic dynamics of noisy chaos. *Physica D: Nonlinear Phenomena*, *7*(1–3), 201–223.

Dabbs, J. M., & Ruback, R. B. (1987). Dimensions of group process: amount and structure of vocal interaction. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 20, pp. 123–169). Academic Press. https://doi.org/10.1016/S0065-2601(08)60413-X

DeDeo, S. (2016). Conflict and computation on Wikipedia: a finite-state machine analysis of editor interactions. *Future Internet*, *8*(3), 31. https://doi.org/10.3390/fi8030031

Ellis, D. G. (1979). Relational control in two group systems. *Communication Monographs*, *46*(3), 153–166. https://doi.org/10.1080/03637757909376003

Ellis, D. G., & Fisher, B. A. (1975). Phases of conflict in small group development: a Markov analysis. *Human Communication Research*, *1*(3), 195–212. https://doi.org/10.1111/j.1468-2958.1975.tb00268.x

Fano, R. M. (1961). *Transmission of information: a statistical theory of communications*. M.I.T. Press.

Fisher, B. A. (1970). Decision emergence: phases in group decision-making. *Speech Monographs*, *37*(1), 53–66. https://doi.org/10.1080/03637757009375649

Fisher, B. A., & Drecksel, G. L. (1983). A cyclical model of developing relationships: A study of relational control interaction. *Communication Monographs*, *50*(1), 66–78. https://doi.org/10.1080/03637758309390154

Fisher, D. (2004). *Social and temporal structures in everyday collaboration (dissertation)* (PhD Thesis). University of California, Irvine, Irvine. Retrieved from https://www.microsoft.com/en-us/research/publication/social-and-temporal-structures-in-everyday-collaboration-dissertation/

Gibson, D. R. (2003). Participation shifts: order and differentiation in group conversation. *Social Forces*, *81*(4), 1335–1380. https://doi.org/10.1353/sof.2003.0055

Gibson, D. R. (2005). Taking turns and talking ties: networks and conversational interaction. *American Journal of Sociology*, *110*(6), 1561–1597. https://doi.org/10.1086/428689

Gleick, J. (2011). *The information: a history, a theory, a flood*. Random House Digital, Inc.

Goffman, E. (1981). *Forms of Talk*. University of Pennsylvania Press.

Grassberger, P. (1986). Toward a quantitative theory of self-generated complexity. *International Journal of Theoretical Physics*, *25*(9), 907–938. https://doi.org/10.1007/BF00668821

Hawes, L. C., & Foley, J. M. (1973). A Markov analysis of interview communication. *Speech Monographs*, *40*(3), 208–219. https://doi.org/10.1080/03637757309375798

Hilbert, M. (2014). *Big data requires big visions for big change*. London: TEDxUCL, x=independently organized TED talks. Retrieved from www.youtube.com/watch?v=UXef6yfJZAI

Jurgens, D., & Lu, T.-C. (2012). Temporal motifs reveal the dynamics of editor interactions in Wikipedia. In *Sixth International AAAI Conference on Weblogs and Social Media*. Retrieved from http://www.aaai.org/ocs/index.php/ICWSM/ICWSM12/paper/view/4573

Keegan, B. C., Lev, S., & Arazy, O. (2016). Analyzing organizational routines in online knowledge collaborations: a case for sequence analysis in CSCW. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (pp. 1065–1079). New York, NY, USA: ACM. https://doi.org/10.1145/2818048.2819962

Kolmogorov, A. N. (1958). A new metric invariant of transient dynamical systems and automorphisms in Lebesgue spaces. *Dokl. Akad. Nauk SSSR.*, *119*(861–864), 2.

Kolmogorov, A. N. (1959). Entropy per unit time as a metric invariant of automorphisms. *Dokl. Akad. Nauk SSSR.*, *124*(4), 754–755.

Krain, M. (1973). Communication as a process of dyadic organization and development. *Journal of Communication*, *23*(4), 392–408. https://doi.org/10.1111/j.1460-2466.1973.tb00957.x

Kreiss, D., Finn, M., & Turner, F. (2011). The limits of peer production: Some reminders from Max Weber for the network society. *New Media & Society*, *13*(2), 243–259. https://doi.org/10.1177/1461444810370951

Leenders, R. T. A. J., Contractor, N. S., & DeChurch, L. A. (2016). Once upon a time: Understanding team processes as relational event networks. *Organizational Psychology Review*, *6*(1), 92–115. https://doi.org/10.1177/2041386615578312

Lu, X., Bengtsson, L., & Holme, P. (2012). Predictability of population displacement after the 2010 Haiti earthquake. *Proceedings of the National Academy of Sciences*, *109*(29), 11576–11581. https://doi.org/10.1073/pnas.1203882109

MacKay, D. J. C. (2003). *Information theory, inference and learning algorithms* (1 edition). Cambridge, UK ; New York: Cambridge University Press.

Madsen, A. K., Flyverbom, M., Hilbert, M., & Ruppert, E. (2016). Big data: issues for an international political sociology of data practices. *International Political Sociology*, olw010. https://doi.org/10.1093/ips/olw010

Marton, K., & Shields, P. C. (1994). Entropy and the consistent estimation of joint distributions. *The Annals of Probability*, *22*(2), 960–977. https://doi.org/10.1214/aop/1176988736

Miller, G. A., & Frick, F. C. (1949). Statistical behavioristics and sequences of responses. *Psychological Review*, *56*(6), 311. https://doi.org/10.1037/h0060413

Miller, J. H., & Page, S. E. (2007). *Complex adaptive systems*. Princeton, New Jersey: Princeton University Press.

Monge, P. R., Farace, R. V., Eisenberg, E. M., Miller, K. I., & White, L. L. (1984). The process of studying process in organizational communication. *Journal of Communication*, *34*(1), 22–43. https://doi.org/10.1111/j.1460-2466.1984.tb02983.x

Ortega, F., Gonzalez-Barahona, J. M., & Robles, G. (2008). On the inequality of contributions to Wikipedia. In *Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 2008)* (pp. 304–304). https://doi.org/10.1109/HICSS.2008.333

Ostrom, E., & Walker, J. (2003). *Trust and reciprocity: interdisciplinary lessons for experimental research*. Russell Sage Foundation.

Parker, K. C. (1988). Speaking turns in small group interaction: A context-sensitive event sequence model. *Journal of Personality and Social Psychology*, *54*(6), 965.

Pavitt, C., & Johnson, K. K. (1999). An examination of the coherence of group discussions. *Communication Research*, *26*(3), 303–321. https://doi.org/10.1177/009365099026003002

Pierce, J. R. (1980). *An introduction to information theory: symbols, signals and noise* (2nd Revised ed. (1st 1961)). New York, NY: Dover Publications.

Pilny, A., Schecter, A., Poole, M. S., & Contractor, N. (2016). An illustration of the relational event model to analyze group interaction processes. *Group Dynamics: Theory, Research, and Practice*, *20*(3), 181.

Poole, M. S. (1985). Tasks and interaction sequences: A theory of coherence in group decision-making interaction. In R. L. Street & J. N. Cappella (Eds.), *Sequence and Pattern in Communicative Behaviour* (pp. 206–224). Edward Arnold.

Poole, M. S. (2007). Generalization in process theories of communication. *Communication Methods and Measures*, *1*(3), 181–190. https://doi.org/10.1080/19312450701434979

Poole, M. S., & Roth, J. (1989). Decision development in small groups IV. *Human Communication Research*, *15*(3), 323–356. https://doi.org/10.1111/j.1468-2958.1989.tb00188.x

Poole, M. S., Ven, A. H. V. de, Dooley, K., & Holmes, M. E. (2000). *Organizational change and innovation processes: theory and methods for research* (1 edition). Oxford: Oxford University Press.

Schramm, W. L. (1955). *The process and effects of mass communication*. Urbana; Chicago; London: University of Illinois press.

Shannon, C. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*, 379–423, 623–656. https://doi.org/10.1145/584091.584093

Shannon, C. (1951). Prediction and entropy of printed English. *Bell System Technical Journal*, *30*, 50–64.

Shaw, A., & Hill, B. M. (2014). Laboratories of oligarchy? How the iron law extends to peer production. *Journal of Communication*, *64*(2), 215–238. https://doi.org/10.1111/jcom.12082

Smith, G., Wieser, R., Goulding, J., & Barrack, D. (2014). A refined limit on the predictability of human mobility. In *2014 IEEE International Conference on Pervasive Computing and Communications (PerCom)* (pp. 88–94). https://doi.org/10.1109/PerCom.2014.6813948

Song, C., Qu, Z., Blumm, N., & Barabási, A.-L. (2010). Limits of predictability in human mobility. *Science*, *327*(5968), 1018–1021. https://doi.org/10.1126/science.1177170

Stasser, G., & Taylor, L. A. (1991). Speaking turns in face-to-face discussions. *Journal of Personality and Social Psychology*, *60*(5), 675.

Watt, J. H., & VanLear, C. A. (1996). *Dynamic patterns in communication processes*. Thousand Oaks, Calif. : Sage Publications. Retrieved from http://trove.nla.gov.au/version/26467801

Wooldridge, J. M. (2008). *Introductory econometrics: A modern approach, 4th Edition* (4th edition). Mason, OH: South-Western.

Yates, D., Moore, D., & McCabe, G. P. (1998). *The practice of statistics AP*. Macmillan.