

Synchrony in Human-Bonobo Dialog

Carl Vogel*, Maria Koutsombogera*, Justine Reverdy*, Anna Esposito†

*Computational Linguistics Group, Trinity College Dublin, The University of Dublin, Dublin 2, Ireland

Email: {vogel,koutsomm,reverdyj}@scss.tcd.ie

† Università degli Studi della Campania “Luigi Vanvitelli”, Caserta, Italy and

International Institute for Advanced Scientific Studies (IIASS), Vietri sul Mare, Italy

Email: iiass.annaesp@tin.it

Abstract—Past research has suggested that bonobos have demonstrated capabilities of successful linguistic interaction with humans. A transcript of interactions linked to such claims is analyzed using methods that have been deployed to assess mutual understanding in human communication. Evidence of dialog plan maintenance on the part of the bonobo is visible, but not of understanding of the human dialog partners (nor vice versa, mostly).

I. INTRODUCTION

Successful interspecies communication between humans and other creatures appears to be possible, to the extent that researchers in cognitive infocommunications¹ have seen ways to improve technology interfaces for human-human communication on the basis of human-dog communication [4]. Others seek to improve communication between humans and non-humans [5]–[7]. Language may be understood as a technology that is adapted for communication [8], [9]. The extent to which interspecies communication involving human language has been successful is an open question.

A published analysis of human-bonobo communication [10] focuses on the role of repetition in establishing cohesion in communication, towards a general argument attributing linguistic prowess to the relevant bonobo, Panbanisha: “we wish to argue that the bonobo’s discursive contributions are non-random and that this interspecies communication takes the form of a cohesive conversation with active co-construction by both parties” (p. 23). We take this assessment to be based on qualitative interpretation. The present work re-analyzes the data reported there, quantifying the extent to which the linguistic repetition behavior significantly diverges from what might be expected in a random process. The method of analysis builds on past analyses of conversations focused on assessment of synchrony and engagement among participants [11]–[15]. We build on past analysis of the bonobo data [10], [11]. Application of the method of analysis advocated here leads to the conclusion that there is insufficient support for the claim quoted above to the effect that Panbanisha’s conversational engagement was non-random.

Repetition provides an fallible index of interlocutor engagement in conversation. Repetition has many functions in conversation beyond signalling synchrony [16]. Further, the illusion of interlocutor participation may be sustained by repeating elements of discourse without attending to the content, and indeed, one way to fake engagement in conversation is to repeat the odd word or phrase uttered by an interlocutor. For

this to be a successful approach, it is necessary that measures of lexical and sub-lexical repetition in un-natural conversations are significantly different than in natural conversational interactions. A conversation in which interlocutors talk to each other about unlinked topics would be un-natural, and any lexical overlap would be random. Nonetheless, in natural dialog, it may happen that repetitions are distributed randomly.

The point has been made that language is not a random process [17]. When one considers distributional properties of natural language, in particular the facts of the frequency relations between successive pairs of items in rank-frequency orderings of words used, one would probably imagine that repetition in natural language is inevitable: it is hardly a surprise for successive statements by distinct speakers to contain the word “the”. It becomes interesting to consider how much and what kind of repetition in conversation is interesting and revealing of engagement and mutual understanding. Recent work has attempted to demonstrate the extent to which repetition in dialog correlates with task success [18]–[20]. Other work is developing profiles of dialogs according to levels of repetition effects (repetition of self, and repetition of others) as measures of interlocutor synchronization and relative involvement in dialog, in support of quantifying evidence of mutual understanding [11]–[15].

We use a Monte Carlo approach in which natural dialog transcripts are compared with randomized versions of the turns. Using past categorization of such approaches [21], the method constructs “full-sample” permutations. Counts of linguistic event types in the natural dialog are compared with counts of the same types in the randomizations. To obtain the randomizations the contributions of each participant are processed into a sequential data-structure, with one cell for each turn. The data-structure is then re-indexed using a random process without replacement, and this re-indexed structure is used to write revised transcripts in which the same turns ultimately happen, but nearly certainly in a different order. Measurements made for the actual dialog are compared with average measures made on the re-ordered dialogs, running the re-ordering process ten times. Here we focus on lexical and lexical sequence repetition counts in successive turns – any turn has a speaker: we consider the count of items shared between that turn and that speaker’s last turn (self-repetition) and the counts shared between that turn and each other speaker’s last turn (other-repetition). In principle, temporal overlap may also be considered as a metric of engagement; however, this data is not available in the bonobo-human data addressed.

§II describes the data related to claims of interspecies

¹See [1]–[3].

1. PB: CARRY YES
2. SSR: you want Russ to carry you? ((quiet laughter)) instead of the dog. (1.0)
3. Panbanisha i'm going to tell you something (4.5)
4. Russ is going to CARRY the DOG because the DOG is SCARED of
5. PANBANISHA. the dog is scared of you.

Fig. 1. Initial transcript segment (5 of 95 lines) [10].

communication re-analyzed here. §III details the methods of analysis used in our assessment of the degree to which the dialog evidence diverges from expectations associated with randomness. §IV considers the human-bonobo interactions, where patterns characteristic of human-human dialogs are not necessarily expected. It is shown that the effects which are visible in the bonobo-human interactions cannot be confidently argued to be other than random.

II. BACKGROUND

Transcripts of bonobo-human interactions have been analyzed for repetition effects with the argument that these effects are not random [10]. The methods addressed involved scrutiny of context (for example, the role of a dog) and a particular focus on repetition in the transcripts (“carry”, “telephone”, “mad”). Instances of repetitions of forms across the entire transcript are reckoned as evidence of overall lexical cohesion. Linking the forms to context provides a foothold in arguing that the forms are contextually grounded, and to that extent non-random. The situation involves Sue Savage-Rumbaugh (SSR), an adult human male (Russ), a dog (Mocha) and Panbanisha, a bonobo (PB); the humans communicate vocally, and the bonobo uses a keyboard to convey lexigrams (along with uninterpreted vocalizations). The transcript and accompanying description reveal: PB has frightened the unfamiliar dog; PB wishes to be carried; the bonobo adheres to a plan of conveying this wish; evidence of this plan is in repetitions of the lexigram for “carry” in distinct dialog turns, after intervening contributions from the other interlocutors. The initial instance of the use of the lexigram by PB is shown in the transcript excerpt in Fig 1. Pedersen and Fields observe [10, p.32], “The laughter [of SSR] and the pauses in SSR’s turn might have indicated to PB that SSR is more prone to her request than the literal message reveals.” It is consistent with the holistic method of analysis advocated by the researchers [10] to speculate about the meaning for a bonobo of a human laugh in the context of the ten-word utterance in which it occurred and for which the propositional content speculated to be derived by the bonobo depends on a preposition that expresses a negated relation (“instead”). We doubt the robustness their suggestion.

The set of arguments presented by [10] are legitimate; however, it is appropriate to scrutinize the data further, if only to bolster the case that its conclusion does not incur a type I error, an inappropriate rejection of the null hypothesis. The null hypothesis here is that the linguistic alignment behavior exhibited in the dialog and its context of interactions is non-purposive, not significantly different from random alignment. Note that purely mechanical systems, like a functioning system of dynamos operating in context, will fall into alignment without purposive intervention [22]. Further, not all instances of human behavioral alignment are about interaction: spectators at a tennis match will tend to align head movements with the

motion of the ball in play rather than with each other, yet the outward spectacle is consistent with purposive alignment.

The intent of the present work is to re-examine the same transcript studied by [10] in order to assess the support that the transcript alone provides for the argument that the bonobo repetition behaviors are other than random. The transcript is taken from the appendix of that work, and treated solely in the elimination of parenthetical descriptions of scenes and line numbers associated with a particular typesetting of turns, using the decisions of the transcriptionist to determine what amounts to a discrete turn for a speaker (thus, the excerpt from Fig 1 becomes that in Fig. 2; notice that SSR has four successive lines of contributions). An indication of which of the three individuals participating in the interaction is responsible for any linguistic contribution (“speaker”) is used to index turns. The facts of synchronization in the behavior of mechanical systems and humans just mentioned are sufficient to support a trivial argument that synchronization in the bonobo-human interaction is non-random. However, it is safe to understand that the intent of that prior work is to show that the interaction exhibits more than trivial non-randomness. The scrutiny provided here is to assess whether non-trivial non-randomness may be confidently ascribed. The conclusion is that the transcripts do not support that ascription.

III. METHODS

The dialog analyzed here has already been independently transcribed [10]. Therefore, one issue of data treatment that would otherwise be contentious is resolved through independent prior transcription of the dialogs: individuation of tokens within the dialogs that are understood as distinct turns (1). This is made difficult by overlap. If a contribution of Speaker \mathcal{A} is accompanied by Speaker \mathcal{B} interrupting Speaker \mathcal{A} in the middle, it is necessary to decide whether this amounts to one turn from each of \mathcal{A} and \mathcal{B} , or whether this is better understood as two turns from \mathcal{A} and one from \mathcal{B} . In availing of extant transcripts, we rely on the transcriptionist’s decisions about turn individuation and ownership.

We emphasize that we use a method that we have described and employed before [11]–[15]; for clarity, we restate it. We use a ‘register’ (3) for each speaker containing the tokens that comprise their most recent utterance (2).

- (1) Ξ is the set of interlocutors (α).
- (2) $u_j = \langle \tau_b, \tau_e, \alpha, \sigma \rangle$ is the j -th utterance of the transcript: τ_s , start-time; τ_e , end -ime; α , interlocutor; σ , utterance. At u_j , $\alpha^{u_j} = actor(u_j)$; $\sigma^{u_j} = statement(u_j)$; etc.
- (3) \mathcal{R}_α , for each $\alpha \in \Xi$, is a register that records the start-time, stop-time and content of the last contribution of α .

\mathcal{R}_Ξ is the set of registers;

$\mathcal{R}_{\Xi/\alpha}$ is the set of registers for all interlocutors but α ;

PB: CARRY YES

SSR: you want Russ to carry you? instead of the dog.

SSR: Panbanisha i'm going to tell you something

SSR: Russ is going to CARRY the DOG because the DOG is SCARED of

SSR: PANBANISHA the dog is scared of you.

Fig. 2. Initial transcript segment, treated [11].

$(g(n, \sigma)[i])$ is the i -th element of $g(n, \sigma)$.

For each utterance, a count is made of tokens shared with immediately preceding turns for each speaker – the speaker’s own last turn (6) and each interlocutor’s last turn (5), in each case determined by (4) for the each n of the n -grams. These are repetition counts. Repetition counts in actual dialogs are then compared with the average counts obtained from ten randomized turn re-orderings. Even in randomizations, the sequence of tokens that comprise original turns remain in the same order in the re-ordered turns – the randomization is of the turns orderings, not the internal contents of turns.

$$(4) \quad \kappa(n, \sigma^1, \sigma^2) = \sum_{i=1}^{g(n, \sigma^1)[max]} (g(n, \sigma^1)[i] \in g(n, \sigma^2))$$

$$(5) \quad other-shared(u_j, \mathcal{R}_{\Xi}, n) = \sum_{\alpha}^{\Xi/\alpha^{u_j}} \kappa(n, \sigma^{u_j}, \sigma^{\mathcal{R}_{\alpha}})$$

(6) a. first count with respect to former value for self:

$$self-shared(u_j, \mathcal{R}_{\Xi}, n) = \kappa(n, \sigma^{u_j}, \sigma^{\mathcal{R}_{\alpha^{u_j}}})$$

b. update register for the agent’s own current utterance:

$$\mathcal{R}_{\alpha^{u_j}} := \langle \tau_s^{u_j}, \tau_e^{u_j}, \sigma^{u_j} \rangle$$

The random-reordering of the dialogs is effected by re-indexing the utterances (8).

- (7) given $u_1 \dots u_{max}$
- $starttime = \tau_s^{u_1}$
 - $stoptime = \tau_e^{u_{max}}$
 - $maxoverlap =$ maximum temporal overlap of u_i and \mathcal{R}_{Ξ} noted at time of shared n -gram computation.
- (8) for each u_i ,
- $\tau_s = rand(0, stoptime)$
 - $\tau_e = \tau_s + rand(0, maxoverlap)$
 - $u_i = \langle \tau_s, \tau_e, \alpha^{u_i}, \sigma^{u_i} \rangle$

In the re-ordered dialog, as in the actual dialog, we count other-shared tokens and self-shared n -gram tokens. Here we rely on n -gram sequences from $n = 1$ to $n = 5$; we contrast “reality” with averages over ten randomizations.

The process outputs a file containing a sequence of turns, each owned by exactly one speaker and with measures for each line taken as described above, one line for each level of n -grams analyzed. Repetition values of n -grams are noted as counts with respect to the values in the registers for both SELFShared and OTHERShared tokens. Thus, in addition to those counts, we consider the corresponding proportions SELFSharedPROP and OTHERSharedPROP.² In addition to the count values we also consider ratios that relativize counts

²We abbreviate these as SSProp and OSProp.

to the total number of n -grams that could have been shared between a dialog turn and the immediately preceding turns stored in registers SSREL and OSREL. See Table I for a consolidated presentation of the quantities monitored.

TABLE I. QUANTITIES ANALYZED

Variable	Interpretation
DIALOGTYPE	actual vs. randomized
OTHERSPEAKERS	total number of participants, minus one
N	length of N -grams (1, 2, 3, 4 or 5)
N'	N -gram length as a two-level factor (“1” or “2+”)
NBAR	N'
NGRAMS	total number of N -grams in a turn
SELFShared	count of tokens from turn shared with own prior contribution
OTHERShared	count of tokens from turn shared with prior contribution of other
SSREL	SELFShared \div NGRAMS
OSREL	OTHERShared \div (NGRAMS \times OTHERSPEAKERS)
SSPROP	SELFShared/(NGRAMS-SELFShared)
OSPROP	OTHERShared/((NGRAMS \times OTHERSPEAKERS)-OTHERShared)

IV. THE PANBANISHA DATA

The 97-contribution transcript of a recording from the appendix of work by [10] is analyzed. The human participants are adults, Sue Savage-Rumbaugh (SSR) and a male named Russ. The bonobo Panbanisha (PB) is the agent whose contributions to the interaction are under scrutiny. Of the contributions, 74 are “spoken”, the remainder of the turns are descriptions of actions that took place. Russ had 5 speaking turns; PB, 13; SSR, 56. Ten turn-randomizations of this transcript are considered.

First, we profile the repetition with respect to the actual and randomized data. Overall repetition effects may be viewed in relation to phrase length. Focusing upon OSREL, the mean of this value by dialog type and speaker is shown in Table II. The number of other-shared n -grams relative to the number that could have been shared, for other-repetition is depicted in Figure 3. There is less repetition of unigrams in the actual dialog than the randomized counterparts, but slightly more repetition of longer sequences in actual dialog than random. Attending to SSREL, the number of self-shared n -grams relative to the number that could have been shared, for self-repetition (Figure 4) and for both sequence length levels, greater levels of repetition are noted in the actual dialog than are evident in the randomizations. Focusing upon OSREL, Figure 5 shows that only Russ had greater levels in actual than randomized dialog. With respect to SSREL, the mean of this value by dialog type and speaker is shown in Table III. Figure 6 shows that only Russ had lesser levels in actual dialog than randomized counterparts.

TABLE II. MEAN OF OSREL BY SPEAKER AND DIALOGTYPE

Dialog Type	Speaker		
	PB:	Russ:	SSR:
Randomized	0.0007692308	0.006685714	0.005940709
Actual	0.0000000000	0.022857143	0.004666583

TABLE III. MEAN OF SSREL BY SPEAKER AND DIALOGTYPE

Dialog Type	Speaker		
	PB:	Russ:	SSR:
Randomized	0.02076923	0.017066667	0.02775327
Actual	0.03846154	0.009333333	0.07046652

Our basic questions here relate to whether there is an interaction between dialog type (actual vs. random), speaker (Panbanisha, Russ and Sue) and the split between n-grams that were actually shared and which could have been shared. These possible interactions, we analyze next.

A. Results

For each speaker, we test the difference between actual and randomized dialogs for the values of OSREL and SSREL, using a Wilcox test. The results are shown in Table IV: the relative other-sharing for Russ approaches significance; relative self-sharing for SSR is significant; no other values are.

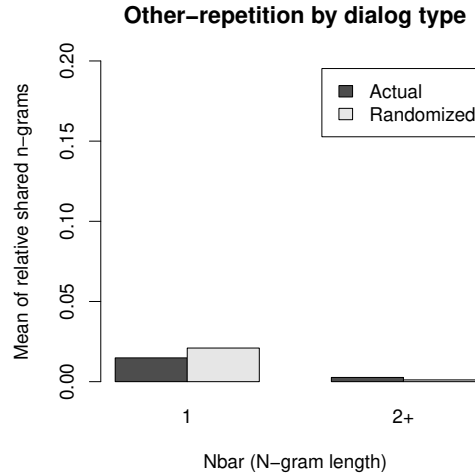
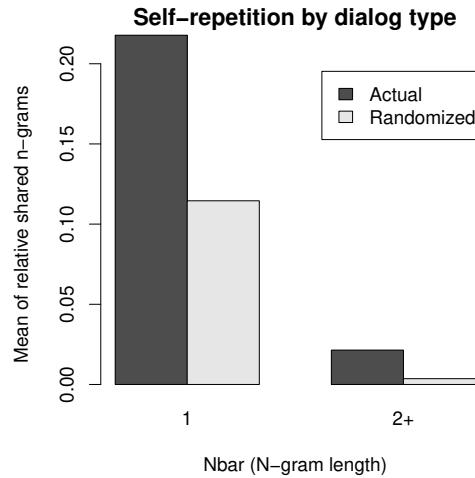
TABLE IV. WILCOX'S W AND p VALUES CORRESPONDING TO THE TEST OF THE DIFFERENCE BETWEEN ACTUAL AND RANDOM DIALOG IN MEAN OSREL AND SSREL FOR EACH SPEAKER

Speaker	OSREL		SSREL	
	W	p	W	p
PB:	21190	0.657	20700.5	0.3498
Russ:	2826	0.08729	3070.0	0.7306
SSR:	401257	0.1354	334427.5	2.695e-11

Figure 7 and Figure 8 provide the association plots showing Pearson residuals related to the contingency table of counts of n-grams that were repeated or could have been repeated but were not. Positive values correspond to quantities that exceed the values one would expect in the absence of interactions among cross-tabulated counts, and negative values result from quantities that are less than would be expected if there were no interactions. Pearson residuals with magnitude between 2 and 4 are significant ($p < 0.05$) and residuals with magnitude greater than 4 are highly significant ($p < 0.001$). In the plots, darkness of shading corresponds to increasing significance. They show the interaction between the dialog type (actual or random), speaker and repetition (other-repetition is shown in Figure 7 and self-repetition is depicted in Figure 8). In particular, they highlight the degree to which cells in the interaction diverge or not from expectations. None of the three speakers show the same pattern for other-repetition or self-repetition. Only SSR provided highly-significant actual self-repetition (in contrast to random).

B. Discussion

The data provides evidence that a dialog plan may be ascribed to SSR and that synchrony with the remarks of others is visible for Russ only. Inspecting the dialog qualitatively gives us the sense that this is appropriate. If PB wanted to express anything it was the desire to be carried, and this persisted. SSR led the conversation, with comment on the evident desire of PB to be carried and about PB's relationship with a dog. Qualitative view of the transcript does not give us the impression that significant mutual understanding was exhibited, except, perhaps by Russ.

Fig. 3. Actual dialog vs Randomized turns: sharing others' N -grams, by N Fig. 4. Actual dialog vs Randomized turns: sharing own N -grams, by N

V. GENERAL DISCUSSION & CONCLUSION

A reasonable objection to our method is that it is not natural to consider transcripts only instead of making use of the underlying recordings. Indeed, we note that others have productively analyzed interaction in human-human encounters without recourse to linguistic content of recordings at all [23]. On the other hand, the transcripts themselves are the basis of analysis in the foundational work that we reconsider here [10] and constitute the data available to us. While our methods embody the sort of non-holistic method of analysis argued against in previous examination of the data [10], comparably abstract analysis is also conducted within that research team [24], considering the temporal patterns of sound and silence during interactions. It is a sound scientific method to analyze factors in isolation, as has been done here.

The bonobo data as profiled above demonstrates a vast imbalance in the contributions of the participants, but effects of repetition that provably transcend randomness only in ways that do not correspond to engaged dialog among humans.

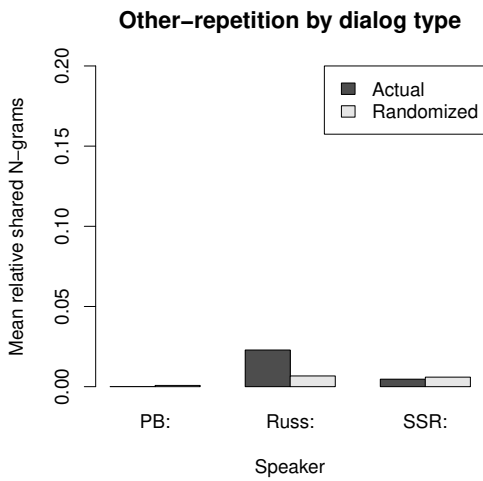


Fig. 5. Actual ordering vs random and sharing others' N -grams, by speaker

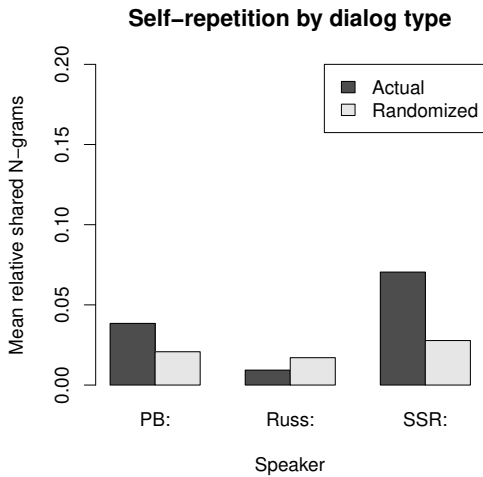


Fig. 6. Actual ordering vs random and sharing own N -grams, by speaker

We have assumed that sharing one's own phrases between successive dialog turns is indicative of maintenance of a dialog plan, and that sharing other's phrases is indicative of engagement. On the basis of these observations, the repetition data reported by [10] and reanalyzed here do not support the widest ramifications of non-random behavior.

In fact, a simpler analysis can be conducted to show non-random behavior by the bonobo, and has been noted [10, p.36]:

An objection to these conclusions could hold that Panbanisha's repetitions are rather due to statistical chance rather than to communicative skills. This is very improbable both given the number of lexigrams available on the keyboard and given her knowledge of the referential content of the lexigrams as demonstrated in previous research. The keyboard has about 390 lexigrams available and that a lexigram, or a small number of lexigrams, should come up repeatedly and systematically as in this conversation due to chance is unlikely. It is even more im-

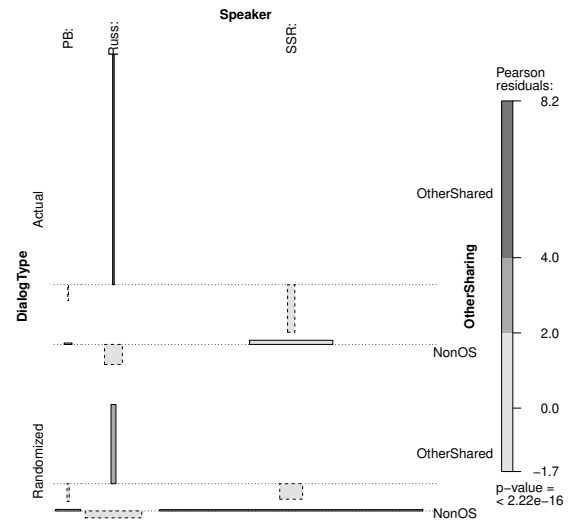


Fig. 7. Proportions of Other Repetition by DialogType and Speaker

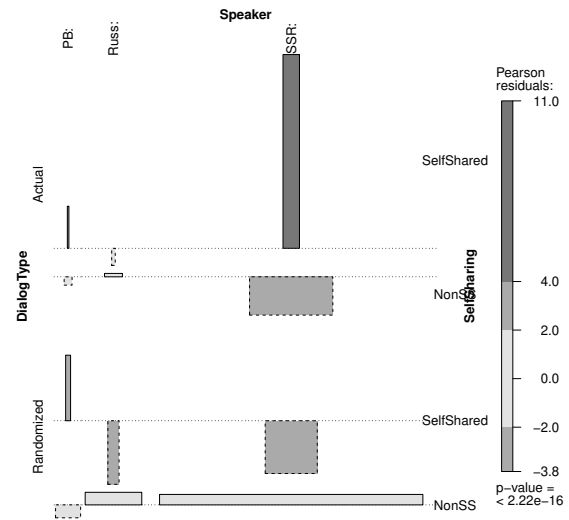


Fig. 8. Proportions of Self Repetition by DialogType and Speaker

probable that appropriate compounds such as "YES TELEPHONE" would be chosen as a matter of chance, and be followed incidentally by a repetition of "CARRY". The most likely interpretation is that Panbanisha uses the lexigrams in active conversational participation to reach her goal of being carried.

Indeed, if one thinks of the keyboard as a 390-sided die, then for a side marked "CARRY" to emerge on top in more than three tosses out of 13 leads to the conclusion that the die is not fair with 99.5% confidence (Bernoulli schema, $p < 0.005$).

The deeper analysis above undermines the claim that the bonobo's use of lexigrams, particularly repetitions, in the interaction with humans is non-random. The bonobo's repetition behaviour in the actual dialog is not significantly different than in randomized counterparts. On the other hand, it is clear

that the researchers involved in the transcript analyzed here perceived that mutual understanding was achieved, and it is clear as well that independent researchers reach the same conclusion [10]. Arguably, such perceptions are among the goals of those who seek to build artificial dialog agents. We do not doubt that Panbanisha has more general intelligence than every artificial dialog agent. We also suspect that most commercial dialog agents, using the analysis we conduct here, would be deemed to be engaged, such that, if they were not artificial, we would be willing to ascribe mutual understanding. Among the challenges of artificial agents is making interactions with them feel natural rather than artificial [25], such as may happen when emotions seem genuine in all parties, as suggested in the transcript analyzed here. We conclude that much can be learned towards improving the viability of artificial agents from studying more deeply human-human interactions and human interactions with intelligent natural agents.

One may see the approach of Sue Savage-Rumbagh as akin to the program of cognitive infocommunications [1]–[3] in attempting to increase the cognitive capabilities of another species by providing it with a capability available to humans, human language. Our analysis suggests that the attempt to provide Banbanisha with human language is not yet measurably successful, even though co-located humans and independent researchers perceive non-random behaviours.

ACKNOWLEDGMENT

This research has been supported through funding from the ADAPT Centre (Grant 13/RC/2106), the European Regional Development Fund, the EU H2020 programme under the Marie Skłodowska-Curie grant No 701621.

REFERENCES

- [1] P. Baranyi and A. Csapo, "Cognitive infocommunications: Coginfo-com," in *2010 11th International Symposium on Computational Intelligence and Informatics (CINTI)*, 2010, pp. 141–146.
- [2] —, "Definition and synergies of cognitive infocommunications," *Acta Polytechnica Hungarica*, vol. 9, no. 1, pp. 67–83, 2012.
- [3] P. Baranyi, A. Csapo, and P. Varlaki, "An overview of research trends in CogInfoCom," in *IEEE 18th International Conference on Intelligent Engineering Systems*, ser. INES, 2014, pp. 181–186.
- [4] T. Kifor, T. Gottdank, Á. Hajnal, P. Baranyi, B. Korondi, and P. Korondi, "Smartphone emotions based on human-dog interaction," in *2011 2nd International Conference on Cognitive Infocommunications (CogInfoCom)*, 2011, pp. 1–6.
- [5] A. D. Cheok, R. T. K. C. Tan, R. L. Peiris, O. N. N. Fernando, J. T. K. Soon, I. J. P. Wijesena, and J. Y. P. Sen, "Metazoa ludens: Mixed-reality interaction and play for small pets and humans," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 41, no. 5, pp. 876–891, 2011.
- [6] G. Lemasson, S. Pesty, and D. Duhaut, "Increasing communication between a man and a dog," in *2013 IEEE 4th International Conference on Cognitive Infocommunications (CogInfoCom)*, 2013, pp. 145–148.
- [7] S. Ondáš, J. Juhár, M. Pleva, P. Ferčák, and R. Husovský, "Multimodal Dialogue System with NAO and VoiceXML Dialogue Manager," in *Proceedings of CogInfoCom 2017: the 8th IEEE International Conference on Cognitive Infocommunications*. Debrecen, Hungary: IEEE, 2017, pp. 439–444.
- [8] C. Vogel and A. Esposito, "Linguistic and behavior interaction analysis within cognitive infocommunications," in *10th IEEE Conference on Cognitive Infocommunications*, 2019, pp. 47–52.
- [9] —, "Interaction analysis and cognitive infocommunications," *Infocommunications Journal*, vol. 12, no. 1, pp. 2–9, 2020.
- [10] J. Pedersen and W. M. Fields, "Aspects of repetition in bonobo-human conversation: Creating cohesion in a conversation between species," *Integrative Psychological and Behavioral Science*, vol. 43, no. 1, pp. 22–41, 2009.
- [11] C. Vogel and L. Behan, "Measuring synchrony in dialog transcripts," in *Cognitive Behavioural Systems*, A. Esposito, A. M. Esposito, A. Vinciarelli, R. Hoffmann, and V. C. Müller, Eds. Springer, LNCS 7403, 2012, pp. 73–88.
- [12] C. Vogel, "Attribution of Mutual Understanding," *Journal of Law and Policy*, vol. 21, no. 2, pp. 377–420, 2013.
- [13] J. Reverdy and C. Vogel, "Measuring synchrony in task-based dialogues," in *Proceedings of the 18th Annual Conference of the International Speech Communication Association (INTERSPEECH2017)*, 2017, pp. 1701–1705.
- [14] J. Reverdy, A. Hayakawa, and C. Vogel, "Alignment in a multimodal interlingual computer-mediated map task corpus," in *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, H. Koiso and P. Paggio, Eds. Paris, France: European Language Resources Association (ELRA), 2018, pp. 55–59. [Online]. Available: http://lrec-conf.org/workshops/lrec2018/W20/pdf/14_W20.pdf
- [15] J. Reverdy, M. Koutsombogera, and C. Vogel, "Linguistic repetition in three-party conversations," in *Neural Approaches to Dynamics of Signal Exchanges. Smart Innovation, Systems and Technologies*, A. Esposito, M. Faundez-Zanuy, F. Morabito, and E. Pasero, Eds. Singapore: Springer, 2019, pp. 359–370.
- [16] D. Tannen, *Talking voices: repetition, dialogue, and imagery in conversational discourse*. Cambridge: Cambridge University Press, 2007.
- [17] A. Kilgarriff, "Language is never, ever, ever random," *Corpus Linguistics and Linguistic Theory*, vol. 1-2, pp. 263–275, 2005.
- [18] D. Reitter, F. Keller, and J. Moore, "Computational modeling of structural priming in dialogue," in *Proceedings of the Human Language Technology Conference of the North American Chapter of the ACL Association for Computational Linguistics*, 2006, pp. 121–124.
- [19] D. Reitter and J. Moore, "Predicting success in dialogue," in *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*. Association for Computational Linguistics, 2007, pp. 808–815.
- [20] J. Reverdy and C. Vogel, "Linguistic repetitions, task-based experience and a proxy measure of mutual understanding," in *8th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, P. Baranyi, A. Esposito, P. Földesi, and T. Mihálydeák, Eds., 2017, pp. 395–400.
- [21] F. Ramseyer and W. Tschacher, "Nonverbal synchrony or random coincidence? How to tell the difference," in *Development of Multimodal Interfaces: Active Listening and Synchrony*, A. Esposito, N. Campbell, C. Vogel, A. Hussain, and A. Nijholt, Eds. Springer, LNCS 5967, 2010, pp. 182–196.
- [22] V. Y. T. Yolong and P. Wofo, "Synchronization in a ring of mutually coupled electromechanical devices," *Physica Scripta*, vol. 74, pp. 591–598, 2006.
- [23] M.-M. Bouamrane and S. Luz, "An analytical evaluation of search by content and interaction patterns on multimodal meeting records," *Multimedia Systems*, vol. 13, pp. 89–102, 2007, DOI10.1007/s00530-007-0087-8.
- [24] S. Savage-Rumbagh, W. M. Fields, and T. Spircu, "The emergence of knapping and vocal expression embedded in a pan/homo culture," *Biology and Philosophy*, vol. 19, pp. 541–575, 2004.
- [25] A. Esposito, A. M. Esposito, and C. Vogel, "Needs and challenges in human computer interaction for processing social emotional information," *Pattern Recognition Letters*, vol. 66, pp. 41–51, 2015.