Eskişehir Osmangazi Üniversitesi
Sosyal Bilimler Dergisi
Kasım 2023(3), 785-800
DOI: 10.17494/ogusbd.1351753

Emergence of Communication Systems in Collaborative and Joint Action[1]

Murat ULUBAY[2]

Ortak ve İşbirliğine Dayalı Eylem Bağlamında İletişim Sistemlerinin Gelişimi

Öz

İletişim sistemleri, kolektif bilişsel faaliyetin gerektiği ortak-eylem alanının önemli bir unsuru olmuştur. (Sebanz and Knoblich, 2006). Çalışmada sanal bir deneysel ortamda katılımcı çiftlerin iletişim imkanları metin tabanlı bir arayüzde çok kısıtlı sayıda harf ve sembole sınırlanarak, sekiz deney oturumunun her birinde benzer ama farklı ortak eylem görevleri verilmiştir. Katılımcılar deneyler ilerledikçe ortak eylem kapasitelerini geliştirmiş ve her çift kendine özel bir iletişim sistemi geliştirmiştir. Yapılan analizlerde deneyler ilerledikçe, birçok niceliksel parametrelerle ilgili olan trendlerle beraber, Yönlendirici ve Saptayıcı-Betimleyici söz-eylemleri kullanım oranlarında sırasıyla azalan ve artan trendler saptanmıştır. Bu durumun olgunlaşmış işbirliği kapasitesini sağlayan bilişsel psikolojik mekanizmaların ve etkin iletişim sistemlerinin gelişiminin göstergesi olduğu sonucuna varılmıştır.

Anahtar Kelimeler: Ortak-eylem, Kolektif zeka, Söz-eylem, Biliş, Bağıntı kuramı, Bilişsel psikoloji

Emergence of Communication Systems in Collaborative and Joint Action

Abstract

Communication is a focal point in studies of joint action, where collective cognitive activity required (Sebanz and Knoblich, 2006). In this study, a virtual experimental environment was utilized where participants' communication was restricted to text messaging interface with a limited set of letters and symbols. In eight experimental sessions, similar but different joint action tasks were given; each couple improved their joint action skills and developed their own private communication systems. Analysis shows, in addition to several quantitative parameters, trends of diminishing use ratio of directive speech acts and increasing use ratio of assertive speech acts had been observed which are claimed to be characteristics of development of a mature and effective communication system and cognitive mechanisms of joint action.

Keywords: Joint-action, Cognitive psychology, Speech acts, Relevance theory, Cognition

Makale Türü: Araştırma makalesi

Paper Type: Research article

1. Introduction

Even considered as an individually experienced phenomenon in its prototypical sense, in most cases, cognition happens as a result and for interacting and communicating with other agents. Therefore it's not extreme to claim that language and communication are essential dimensions of social cognition, joint action and collective cognition (Hutchens and Johnson 2009). This centrality of language and communication for cognition in social context, required an expansion of prototypical

[1]This study is derived from the PhD thesis by Murat ULUBAY, under the supervision of Annette Hohenberger titled as "Resilient Features of Re-Emerging Dyadic Communication Systems in An Interactive Virtual Environment" in the Department of Cognitive Sciences, Informatics Institute, Middle East Technical University.

[2] Asst. Prof. Dr., Ankara Yıldırım Beyazıt University, Business School, mulubay@aybu.edu.tr ORCID: 0000-0002-9775-5754.

Makale Gönderim Tarihi: 29.08.2023
Makale Kabul Tarihi: 02.10.2023

individual cognitive system paradigm to a novel research paradigm of a group of cognitive agents and group processes (Galantucci 2005; Sebanz et al., 2006). The central area of research is specified to be the contexts of joint-action where agents are collaborating and coordinating for a common task. Therefore main object of measurement and observation, in these studies happens to be communicative activities and communication systems (Steels, 2006). Action-perception couplings on the other hand was another novel conception to study cognition in such kind of collaborative contexts (Wagner et al., 2003). This concentration on communication materializes in research that investigates the cognitive processes of joint action by observing (and facilitating experimentally) emergence of simpler than natural language communication systems specialized for the shared task and goals (Shintel & Keysar, 2009). The experimental research on the emergence of communication can be exemplified as Galantucci, 2005, Scott-Phillips, Kirby, and Ritchie, 2009, where participants interacted in a computer environment for joint-action tasks and engaged each other via computer based communication channels which severely restricts language use if not completely.

In line with this past and current research trends in cognitive psychology and cognitive sciences, this study will inquire how human agents succeed at collaboration, coordination and effective communication in terms of cognitive mechanisms and theory proposed in the literature, by means studying of emergence and utilization of dedicated communication systems during this collaborative actions.

1.1 Problem Statements and Hypotheses

The problem statements of the research is expressed in the original thesis study as follows:

> "Are individuals capable and if so, to what extent are they capable of creating a new/alternative communication system under specific joint-action related task constraints, communication constraints and environmental/spatial conditions?
>
> Are there arbitrary variations among possible communication systems that are expected to emerge or are there resilient features and trends that materialize reliably across these communication systems?
>
> During After the emergence of a novel but yet preliminary communication system, how do individuals reach a convention and align their communicative actions in terms of symbols or signs?
>
> How is the emerging communication system related to the task at hand, the environmental constraints or features, namely to reach a certain goal in the participants' joint action? In other words, to what degree can these communication systems be considered as an adaptive system that is inherently integrated with/determined by the requirement of a general set of adaptive skills that are constrained by the requirements and necessities of the tasks and environment?" (Ulubay, 2013)

In concordance with these problem statements, the general hypotheses of the study are indicated as follows:

H1- The couples will be able to device unique ways of developing a communication systems under the task and environmental constraints (increasing trends of turn success ratio and task completion rate).

H2- The trends of change of characterizing features (use ratios of speech act categories, new lexical item generation rates) of communication systems will be consistent and resilient across the couples and during progression of experimental sessions.

H3- The types of communicative actions and their contents, and the lexical inventory will be sufficient and limited to the task and environmental constraints and requirements.

1.2. Significance of the Study

The past joint action and cognition of communication, emergence of communication studies used very strict constraints on existing communication skills and tools of the participants in their experimental scenarios. The experimentally available communicative channels disabled any natural language based communication but allowed only very rudimentary graphical (not even textual) signaling (Galantucci, 2005) or using the indexicality of the behavioral actions during joint action (Scott-Phillips et.al, 2009). This usually provide findings about whether the participants were able to communicate and the variations of rudimentary communication styles. In our study by adopting a mild restriction approach (which will be detailed in the method section), and several experimental sessions with the same dyad, we enabled more complex joint action contexts. This enabled tracking the gradual improvement and evolution from a rudimentary to a relatively developed, and then to a mature and adequate communication system accompanied by effective joint action, through the sessions.

Secondly the emerging communication systems were more real life ones and can be studies in comparison with a natural language based communication. For instance, we can inquire about the communicative function of a message by categorizing it in terms of speech acts.

Last but not the least, the insights provided by observations of variations of these communication systems and the strategies in creating them, by the participants, will shed light on the cognitive communicative capacities such as, the multi-modal communicative and other adaptive skills (textual, graphic, iconic, deictic, context awareness, embodied etc.) that may be at work in that particular experimental context which are otherwise invisible in previous research with similar experimental paradigm.

2. Literature Review: Cognition and Communication

The study of communication (language/language use) by cognitive science is generally in line with the mainstream conceptions of linguistics and psychology/psycholinguistics. On one side, in linguistics, language is studied as a formal external entity with its abstract features and on the other side is a cognitive capacity, an outcome of a cognitive module used to produce individual linguistic behaviour (Port, 1999). However, understanding and studying language as a component of multi agent cognitive activity became a modus operandi of cognitive science after criticism of computational representational model to be augmented by an interest on social and biological dimensions of language phenomena (Thagard, 2005). For example, the developments in the pragmatics field of linguistics, in its conception of communication from code model and conduit metaphor to the *inference model* (Grice, 1975), created a cognitive turn in the study of communication. In the code model, the message was coded into linguistic form by the speaker, carried via her utterance (*the conduit*) and the meaning was decoded by the hearer. In this study even the communication will occur in textual/computer environment, following this convention, the parties in conversation will be denoted as speaker (sender) and hearer (receiver).

The shortcomings of the code model is the observation of the phenomena of *under-specification of meaning by the message*, i.e. same message/content can carry multiple meanings, intentions, depending on the context, past history of the parties in communication, shared cognitions. To be able to understand mutual understanding in joint action, in addition to understand mutual intentions in terms of behavioral actions for coordination, this second task of expression and detection of intentions by the parties of the communication, transpires as a second dimension. This requires the study of mental capacity of context and situation awareness and ability to infer the intentions of the other party based on these information. This cognitive turn was ushered by Gricean notion of “implicature” and provided the inferential model of communication based of Grice’s four maxims of communication

(Levinson, 2000) including the maxim (principle) of relation which asserts that speaker and hearer assumes the *relevance* of the utterance to the context and inferential encoding and decoding relies on this assumption when processing the contextual information.

2.1 Cognitive Mechanisms and Theory of Relevance

Inferential model assumes the active, multiple step logical operations of speaker and hearer about the context and shared information of the both parties when coding the message and then decoding it, for the sentence meaning and then inferring the intention, i.e. Gricean *meant* meaning of the speaker. This assumption is considered to be theoretically sound but underdeveloped and inconsistent with how mind works in terms of economy of cognitive resources by cognitive studies of communication. Dan Sperber's *theory of relevance* builds upon the maxim of relation (relation) and is a good example of how cognitive science approaches communication in its own scientific terms:

> "In relevance-theoretic terms, any external stimulus or internal representation which provides an input to cognitive processes may be relevant to an individual at some time. According to relevance theory, utterances raise expectations of relevance not because speakers are expected to obey a Co-operative Principle and maxims or some other specifically communicative convention, but because the search for relevance is a basic feature of human cognition, which communicators may exploit." (Sperber and Wilson, 1995, p. 119)

So dubbed as *cognitive principle of relevance* by Sperber's theory of relevance, human cognitive capacity of searching for relevance, in the context of any interaction, including joint-action, and in communication, describes the tendency of selecting most relevant contextual information to infer meaning, and the most relevant is calculated by the least cognitive processing effort requiring to come up with an interpretation of the message within the context (Van der Henst & Sperber, 2006).

> "In relevance-theoretic terms, other things being equal, the greater the PROCESSING EFFORT required, the less relevant the input will be. Thus, RELEVANCE may be assessed in terms of cognitive effects and processing effort. Follow a path of least effort in constructing and testing interpretive hypotheses (regarding disambiguation, reference resolutions, implicatures, etc.). Stop when your expectations of relevance are satisfied (Wilson and Sperber, 2006, p 610)

In addition to the *cognitive principle of relevance,* theory of relevance introduces a second principle, *communicative principle of relevance,* claims that every communicative action, stimulus, including the gestures, behavioral actions to signal a message and of course linguistic ones, presumes its optimal relevance by the speaker and the hearer. This presumption of optimal relevance (Wilson and Sperber, 2006, p. 614) means that if a message is made visible apparent by the sender, it is assumed to be relevant to the context by both parties including the receivers, so that they can operate on the cognitive principle of relevance. The signal itself implies that I-) "this message is worth your (the receiver's) efforts for seeking a relevance", and ii-) the communicator used her best capacities to send the message in the most relevant form and meaning.

These two principles and theory of relevance is implemented in explanation of cognitive economy dimensions in comprehension in a communicative and joint-action context. Secondly it has suppositions about the mental architecture regarding it's relation to more general mental capacity namely theory of mind. Sperber and other practitioners of theory of relevance claim that conscious chains of logical operations on the information about relevance, i.e. using general mind reading abilities is not viable considering the real time, online processing of intentions and meanings during a communication and joint action. What they propose is there is a dedicated *module of intentionality*

detection which is a submodule of general reasoning system (among many like a continuously operating eye-direction detection module) performs this mind reading in the form of selecting the most relevant with the least effort. This is considered essential for detection of speaker's meanings from the content of their sent message, or utterance (Sperber, 1996; Origgi and Sperber, 2000).

2.2. Cognitive Mechanisms of Joint Action and Communication

Taking a social perspective on cognition is necessary in the case of joint action where mechanisms of task sharing, joint and/or shared attention, action observation, coordination and agency (Sebanz et al., 2006) are essential. This perspective can take the form of paradigms of distributed cognition (Hutchins, 1995) and extended cognition (Clark and Chalmers, 1998) where the unit of analysis is the group, i.e. the collectivity of individuals is considered to be the cognitive system to be analyzed. However in joint-action research the focus is on "the cognitive and neural processes within the boundaries of individual minds, acknowledging the major force of interaction to shape processes of joint action" (Galantucci and Sebanz, 2009, p. 256). Such joint-action theories studies the cognitive underpinnings of action alignment, perception-action couplings during interaction, when at least two agents trying to achieve coordination. These mechanisms are not tailored specifically for communication but includes it as an indispensable dimension of joint action, and comprises the general capacity of inferring the intentions of the partner, guessing mental states (who knows what, what is visible to the other party, what does partner wish to do as a coordination action) i.e. mind reading. In contrast with the relevance theoretic assumption of a dedicated relevance tracking, intention detection cognitive module as a part of mental architecture, joint-action researchers assert that there may be several low level processes that enables the detection of the intention in the communication, hence the comprehension of the intended meaning along with other requirements of intention detection like behavioral actions (in addition to communicative actions). One of the low-level mechanisms/processes is interactive alignment where agents synchronize cognitive processes related to joint motor/behavioral actions, their conceptions (synced mental models), and finally linguistic alignment of lexical items and semantic representation etc.:

> "Interaction provides interlocutors with many cues that can support coordination of meaning, even when they are neither produced intentionally for that purpose nor interpreted as signaling speakers' intention. In many circumstances, interlocutors can take advantage of these cues to adapt their behavior in ways that promote coordination, bypassing the need to resort to deliberative inferential processes" (*Shintel* and Keysar, 2009, p. 260).

Joint action theory considers the shared context of coordination and collaboration as a productive source for harvesting cues for the intentions of the speakers and predicting the comprehension performance of the hearers when producing an utterance. The shared implicit information provides the means to diminish the load of being as much as communicative possible for the speaker and to provide detailed linguistic expressions of the intent and offloads the cost of deliberative inferential processes.

> "Another advanced cognitive requirement is being able to use several forms or protocols of communication, as well as communicative content. These protocols are then negotiated and selected during the online interaction ("on the fly"). This is the ability to modify communication behaviors on the basis of the potential targets of communication acts (Nolfi, 2005, p. 242). For example, a human agent can limit communication acts, filter useful ones, can regulate communication flow (turn taking), and the agent has the ritualized forms of communicative interaction, like read-back rituals to guarantee or to receive confirmation from the speaker about the correctness of the reception of the signal." (Ulubay, 2013)

3. Method

In this study a computer environment was utilized (Active Worlds, see Figure 1). ActiveWorlds is an online, interactive, multi agent and customizable virtual world software. In this environment participants have their avatars, can wander in the universe from a first or second person view, can establish visual contact and chat with other participants. A 3D maze like environment was designed for this experiment. Each participant "plays" the experimental session via an ActiveWorlds installed PC, in a controlled environment, in separate rooms monitored by the experimenter.

Participants were teamed as couples and their natural language (NL) use was hindered by restricting them to a limited set of letters and symbols (namely, q,w,e,r,t,i,y,o,p, →, ?) for the text based messaging interface. In eight experimental sessions, similar but different collaborative tasks were given, such as finding each other or an object. Across these sessions, each couple increased their skills in coordination and using the given constrained communication channel to help complete these collaborative tasks.

The main goal and general research questions were i) Will participants be capable of creating a simple communication system to help coordinate and share necessary information? ii) How will the participants align their communicative and behavioral actions to accomplish the joint action tasks? iii) Variations of emerging communication systems, and what are the effect of environmental and task constraints in the structure of these systems as a common denominator?

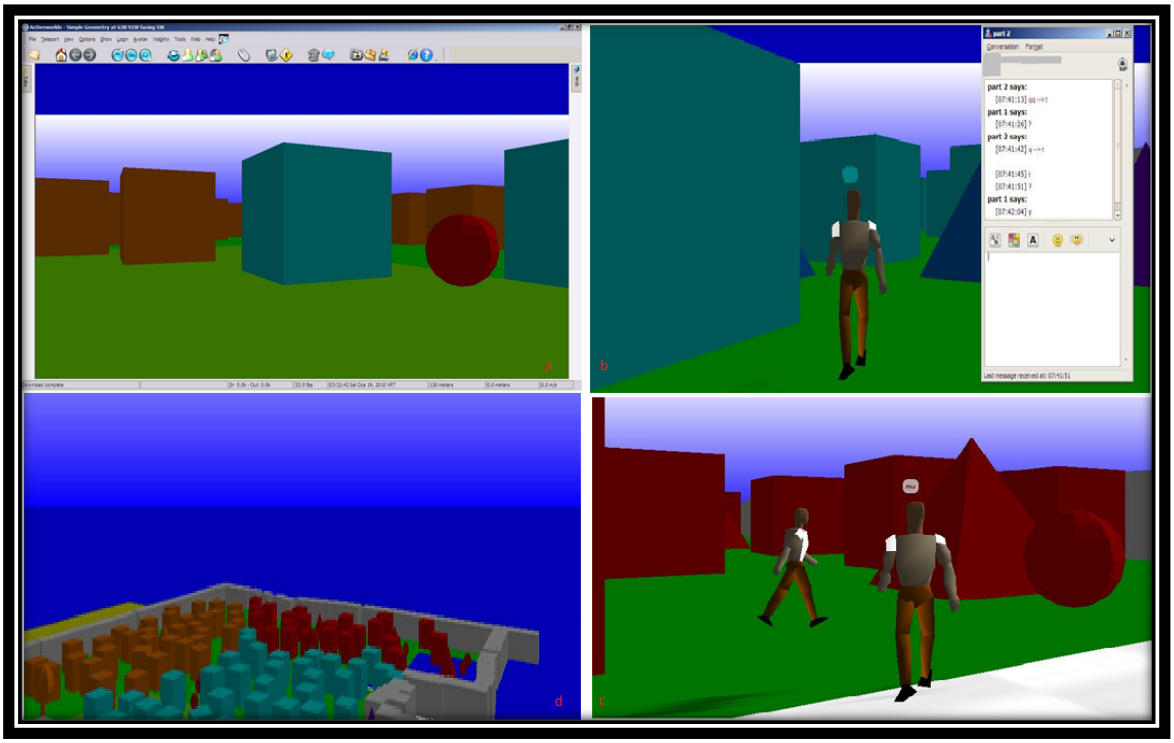

Figure 1. Images in clockwise direction (a) First person, (b) second person view of the ActiveWorlds environment (c) visual interaction of participants (d) panoramic view of the environment

The assigned tasks were finding a unique object in the maze. To facilitate and require communication; in the first three single task sessions, the target was given to one of the participants (interchangeably across the sessions) and the other participant was informed that the partner has the task goal. The maze plan kept same across the sessions but target objects are inserted to tricky locations as novel objects before each session. Starting from session 4, until the last 8th one, both

participants are given target objects, so they need to communicate their target object to the partner and collectively find both of them. This created an increase in the amount of communication content independent from the maturation of communication system or couples increase in efficiency of joint-action. This effect of the manipulation is accounted for by analyzing the trends after 4th experiment separately (for 4-5-6-7-8th sessions) in the analysis section. In addition to increasing task object number from one to two, another intervention was giving a 2 months break, after conducting the first 6 sessions within a period of around 8 days. This was to see whether there are any temporal resilience in the communicative and coordinative skills they have developed. See Appendix A for instructions for the participants and tasks of each experiments.

A total of 22 participants volunteered for the experiments. Of eleven couples, data from eight of them were used due to lack of completion of the sessions, or other logistical problems made their data available after the analysis were completed for the rest of the participants. All participants (11 female, 5 male, mean age 27.5, range 21-34, std. 4.21) were undergraduate or graduate students, and were Turkish and English (second language) speaking.

3.1 Data Collection

After each experimental session, session information is exported (success, failure, time stamps) separately and communication logs are recorded. Each session lasted 10 minutes but for the data collection purposes, after each session a post experimental interview is conducted with both of the participants separately. In this interview, for each turn taking of the speaker, speaker reported the intention of the utterance and in the other interview hearer reported the comprehension for the same utterance. A turn comprehension success (1 or 0) is registered if the comprehension matches the intention (see Table 1). The researchers monitored the session and took critical contextual notes for each utterance instance like whether there was a visual contact between the participants or about target is found or not, some of the contextual notes are acquired during the interviews.

Table 1. Sample from post experiment interview records

1	2	3	4	5	6	7	8
LOG	*partcp1 Interpretation*	*partcp2 Interpretation*	*partcp1 Intention*	*partcp2 Intention*	*context*	*S/F*	*Speech Acts*
(4:53:46 PM) partcp1: yepıy prwt		yeşil piramit	yeşil piramit (bulalım)		no visual contact	Success	Commissive
(4:53:55 PM) partcp1: put		bul	bul			Success	Commissive
(4:54:00 PM) partcp2: qeywe req qure put	kahve renk küre bul			kahve renk küre bul		Success	Directive
(4:54:16 PM) partcp1: oq		ok	ok (for turn1 of p2)			Success	Commissive
(4:54:16 PM) partcp2: irq yepıy prwt oq	ilk yeşil piramiti (bulalım) ok			ilk yeşil piramiti (bulalım) ok	no visual contact	Success	Commissive

4. Analysis and Results

The quantitative data gathered from communication logs annotated, categorized and analyzed. List of all 24 parameters measured and analyzed in the original study is presented in Table 2. The common properties, characteristics of the communication systems developed by the participants is described. The dissimilarities and idiosyncratic issues are explained. Speech act categories are used for describing the types of communicative actions that took place during the experiments. Only the results and relevant discussions about turn success, new lexical items, and speech act categories parameters of the original study will be presented in this study:

Table 2. List of Quantitative Parameters for a Couple's session data

1. Turn Number	13. Total time (sec.)
2. Time	14. Task Result: Success/Fail
3. Participant	15. Number of P1 Turn takings
4. Message	16. Number of P2 Turn takings
5. Gap [interval] btw. turns (sec.)	17. Total Number of Turns
6. Turn Success / Failure	18. Number of Successful Turns
7. Character Length (for each turn)	19. Number of Failed Turns
8. Number of lexical items	20. P1 ∑ lex. İtems
9. Complexity (of sentence)	21. P2 ∑ lex. İtems
10. P1 ∑ character length	22. ∑ new lexical items in each exp.
11. P2 ∑ character length	23. ∑ lexical items (type)
12. ∑ character length P1P2	24. ∑ lexical items (token)

The selected variables, use ratio of speech act categories across sessions, number new lexical items in each session, token number of lexical items in each session and turn success ratio, are subjected to statistical tests of variance. Repeated Measures Anova (RMA) is used to detect whether the change on the dependent variable across different time intervals, due to the effect of some independent variable, like a medicine, treatment, education or learning effect is significant and to compare the variation over time with variation within the treatment group (Field, 2009). In our research design, trough successive sessions, participating couples' exposition to same virtual environment, with slightly different but same type of tasks is claimed to create a change on the communication systems characteristics (a maturation effect). The dependent variable here is the selected parameters of the communication systems; such as use ratio of certain speech act categories or number of new lexical items invented in each session. If the turn comprehension success ratio significantly increases as the sessions proceed, we can conclude that the emerging communication system is maturing and becoming adequate to support the needed coordination in joint action.

RMA is regarded to be robust in the case of moderate violations of the normality requirement (Blanca et al. 2017). Therefore, the results of analysis are reported even if the Saphiro-Wilks test of normality was not satisfied. However, to support the significance of RMA analysis results, non-parametric statistics i.e. Friedman's Test was used (normal distribution is not a requirement in Friedman's Test). This test was used in both of the cases of non-extreme violation and extreme violation of normality. In each cases, the significant results of Repeated Measures ANOVA is corroborated by Friedman's Test. This finding can be interpreted as additional support for the overall robustness of Repeated Measures ANOVA in cases where normality is not satisfied. Nevertheless, it's worth considering that this lack of normal distribution, albeit improbable, as a limitation in specific scenarios when generalizing the findings"

4.1. Number of token and new lexical items,

In in the initial phases of the experiments, since the participants has no common ground for communicating the target given, or to make requests and to ask questions ("follow me", "where are you?"), there is a period of creation, negotiation and alignment (lexical alignment) of new lexical items to express intentions, meanings. The main strategy was creating shortcuts with the available letters and symbols, for most desired and required words, similar to chat abbreviations used for practicality. Some couples used single letters to denote required words and expected to other party to guess the meaning from the context and task requirement and then a mutual alignment on the use of the created lexical item achieved. Some other couples used iconicity of the available letters to denote objects (introducing letter "o" to denote "sphere" as a lexical item). The expected trend was an increase in the number of new lexical items created in the earlier phases, and after a while, since the task and

environmental structure is similar, the created lexical inventory renders adequate and the trend of creation of new lexical items flattens or decreases in the later phases. This increase in the size of lexical inventory is also parallel with the trend of token number lexical items (words) used in each experimental session. But after efficient coordination is learned the amount of communicative action tends to decrease in the later phases of 4-5-6-7-8th sessions.

According to the RMA, the change of *token number of lexical items* across the eight experimental sessions was significant (F (7, 49)=5.40, p<.0001). The decreasing trend of token number of lexical items in 4-5-6-7-8th (after the second target mode is introduced) sessions was also, for 4-5-6-7-8 is a steady decreasing one, which is verified by the Repeated Measures ANOVA with marginal significance (F (4, 28)=2.60, p=.058). The change of *number of new lexical items* across the eight experimental sessions was significant p<.0001).

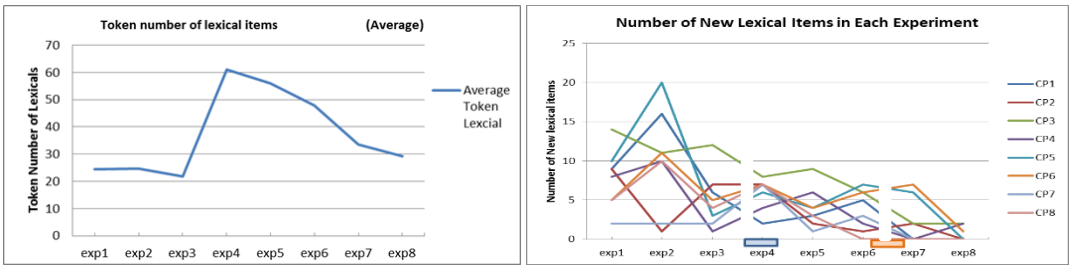

Figure 2. Average Token number of lexical items and new lexical items for all 8 couples (CP1-CP8).

4.2. Number of Turns and Turn Success ratio

Total number of turns in each session is a similar indicator of amount of communication content or communicative action, as token/type number of lexical items used (in 4.1.1). As expected, total number of turns in each session initially increases as the communicative skills improves (hence the communication system develops), and is effected by transitioning to 2-target task mode, but then similar to the case of token number of lexical items; the efficient coordination decreases the need to communicate so we observe a flattened or decreasing trend. The expected increase in the successful comprehension in each session was also observed and confirmed. The change in the Turn success ratios and the numbers of turns were significant, as shown by RMA (F (7, 49)=16.12, p<.0001) for turn success and (F (2.9, 20.9)=3.29, p=.043) for average number of turns.

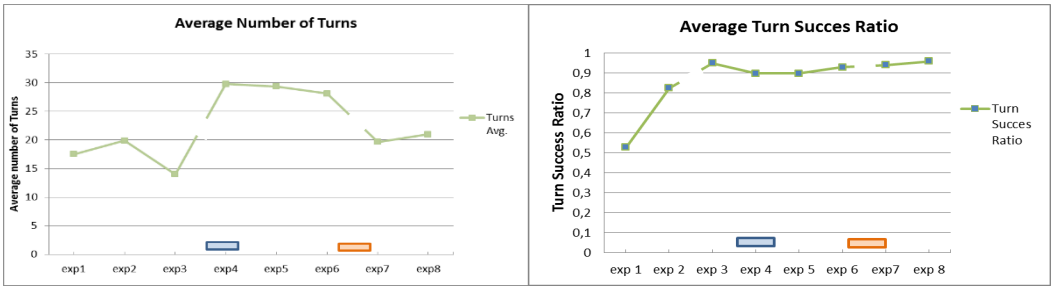

Figure 3. Average Number of Turns and Av. Turn Success Ratio of the 8 couples across the 8 experiments

4.3. Speech act categories

Speech acts are used to categorize the communicative effect or function of utterances in pragmatics. Searle's 1969 classification of speech acts is widely accepted and used in natural language processing studies as well as in pragmatic analysis (Riloff and Qadir, 2011; Appling et al. 2013). Bach's simplified four category version in his 1979 study, contains, "assertive" speech acts which are

declarations about facts, which can be true or false, i.e. has a propositional content. Directive speech acts are communicative actions that aim influence the hearer in terms of causing a behaviour or providing a communicative action in response. Expressive speech acts are the utterances expressing the mental-emotional state of speaker to the hearer or acknowledging hearer's mental-psychological state. Commissives are communicative actions with the function of making commitments about upcoming joint actions like assurances, agreements (i.e. agreeing to do something) or soliciting a collaborative action. Table 3 summarizes how possible communicative actions are categorized in to speech acts.

Table 3. Speech act categories adopted from (Bach, 1979) as cited in Ulubay, 2013.

	Definition of Speech Acts
Assertive	affirming, alleging, announcing, answering, attributing, claiming, classifying, concurring, confirming, conjecturing, denying, **disagreeing**, disclosing, disputing, identifying, informing, insisting, predicting, ranking, reporting, **stating**, stipulating **(also categorized as *Constative*)**
Directive	advising, admonishing, **asking**, begging, dismissing, excusing, **forbidding**, instructing, **ordering**, permitting, **requesting**, requiring, suggesting, **urging**, warning.
Expressive	apologizing, condoling, congratulating, greeting, thanking, **accepting** (acknowledging an acknowledgment) **(also categorized as *Acknowledgment*)**
Commissive	**agreeing**, guaranteeing, inviting, **offering**, **promising**, swearing, volunteering

Each turn-taking of the participants during the chat conversations was annotated including the speech act categorization. A guideline for speech act category judgments was provided to two independent raters. An inter-rater agreement test was conducted to check the used categorization guideline's fitness to elicit speech acts categories reliably (Average Kappa = 0.83). For each experimental session (all 8 experimental sessions are abbreviated as Exp1,…,Exp8), category totals of each speech act category calculated and normalized by dividing category totals by the total number of turns in that session (Figure 4).

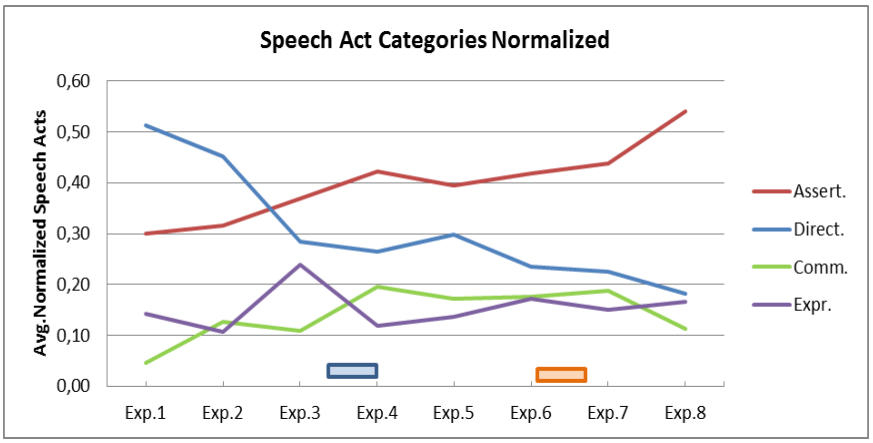

Figure 4. Speech Act Categories in average (use ratio) of 8 couples

The experimental environment and tasks, required users to learn to navigate in the environment and communicate with the team member. In the early experimental sessions, participants needed to familiarize themselves with the environment and develop necessary conventions which means common lexical inventory. At this early stage, the main challenge was to resolve ambiguity: "What may my teammate be meaning with her utterance?" After some trial and errors the lexical inventory began to flourish and became saturated eventually. Creation of new lexical items was less necessary, the task completion time and the amount of communication decreased. This means that the teams were adopting to the environment and to the tasks and used the existing communication channel more efficiently by means of the developed communication system.

Directive speech acts are usually orders, requests and questions. These are required when there is ambiguity in the received message or the visual field and/or about navigation, or a problem with getting the necessary information from the teammate. Directive speech acts include requests of following each other, asking for the location, confirming the intended meaning. Assertive speech acts are defined as statements, propositions with truth value. The increasing trend of assertives and decreasing trend of directives (Figure 4) as observed at later stages of the experiments, describes a phase of increased efficiency of the communication system, whose lexicon and pragmatic power are sufficient for meeting the communicative challenges and succeeding on the collaborative tasks, due to efficient resolution of ambiguity in communication and coordination. The change in the use ratios and the use numbers were significant, as revealed by a Repeated Measures ANOVA ($F(2.4, 16.7)=3.51$, $p=.047$ for assertives and ($F(2.3, 16.2)=4.95$, $p=.018$ for directives). At this phase, there is less ambiguity in communication and coordination. The less the challenge the less the teammates needed directives and could use assertives instead. Having become skillful teammates, they then just needed to exchange or share information; the required coordination was virtually spontaneous and effortless and was the automatic outcome of the uncovered new information.

5. Discussion and Conclusion

In this joint action study, emergence and development of dedicated and private communication systems for each couple was observed. This development started from an elementary phase to a, nearly, established phase serving for the needs of the agents needs for accomplishing task requirements. This development has been observed and characterized by measuring and analyzing linguistic and pragmatic/communicative features of the emerged communication systems for each couple. In case of each couple a lexical inventory was created very quickly in the first half of the eight experimental sessions and needed few supplements during the later phases as a result of meeting the requirements of the task and environmental constraints. The participants, though each couple using varying strategies, succeeded in the interactive alignment on the common strategy of lexical item creation. This alignment is not only about the mode or the specific strategy used for developing the communication system but also includes mutual lexical alignment, which means, if one member invented a comprehensible, novel and useful new lexical item, the other member showed a higher tendency of using it in next relevant context. Actually it's extremely rare if not absent, where a member offered a synonymous lexical item after a word is already convened upon

Similar to the situation of less new lexical items required in the later phases of the series of experimental sessions; the trend about the amount of communication required has been a flat or decreasing trend in comparison to intermediary phases. While the first trend of less new lexical item requirement characterizes the saturation and maturity of the communication system, being concise and parsimonious in using the system is the indicator of being more and more efficient collaborators for each couple.

Starting from the earliest experimental sessions, couples managed to complete tasks and turn success ratio increased to above 90% for most of the couples. As a qualitative observation, as mentioned in section 4.1, different couples used different strategies ("tricks") to invent ways of communicating using the restricted communication channel. These findings indicate that "H1- The couples will be able to device unique ways of developing a communication systems under the task and environmental constraints" was confirmed.

The findings about the trends of parameters of speech act categories (section 4.3) and new lexical item generation (section 4.1) points out that they were nearly universal to all couples during the experiments, which indicates that the hypotheses that "the trends of change of characterizing features of communication systems will be consistent and resilient across the couples and during progression of experimental sessions" (H2) was confirmed.

How can we relate this maturation of communication systems and the acquisition of efficient coordination skills tailored for the task environment, to the cognitive mechanisms introduced by psychology of communication and joint action literature? The theory says that there are continuously operational cognitive mechanisms assigning the potentially ambiguous messages an *optimal relevance* and selecting the most effortless interpretation to maintain cognitive economy during the real time joint action (the cognitive principle of relevance). The performance of the couples as the data suggests, the observations and the post experimental interviews show that agents operated under the contextual constraints which were the source of implicit information about which was the optimally relevant interpretation of any message. These constraints, were actually very productive since they were narrowing down the expectation space about what might be the intention of the speaker in that particular context. As a shared experience from the previous sessions "the task must be some object in the maze", "this novel lexical item", or unfamiliar message body must be about that, or "this request must be about coordinating our movement in the maze", or "location of the partner", or "the location of the object" etc. the general qualitative interpretation of the responses of the participants in the interviews indicate that these were not done after explicit chain of reasoning but more or less checking the fit of the possible interpretation to the activated mental model of the situation. This effect of constraints also indicate that the hypotheses that "types of communicative actions and their contents, and the lexical inventory will be sufficient and limited to the task and environmental constraints and requirements" (H3) was confirmed.

The second dimension of the optimal relevance, the participants acted and communicated under the accurate assumption that the partner is operating under the same narrowed down expectation space. This increased the confidence of communication under uncertainty and in cases of failure of correct comprehension, we observed a persistence of using same proposed but not settled down novel lexical items and they were comprehended in the second attempt of use in the same session or in the next experimental sessions. This second dimension of the observations can be interpreted as validation of *communicative principle of relevance* which was also referred to the presumption (both by the speaker and the hearer) of optimal relevance of the message, so that this confidence mitigated the confusion as a result of miscommunication and motivated trial and errors in communication.

The *interactive alignment* process as suggested by joint action theory is also universally observed during the experiments. The process itself being not limited to communicative actions, is materialized in spontaneous actions of following each other in cases of visual contact, and in terms of behavioral actions. But also, as for the communicative behaviors, the participants immediately adopted the lexical items proposed by one of the members of the dyad (i.e. couple), or communication strategies once convened upon mutually or just only after being used by one side and comprehended by the other party. This swift adoption of communicative conventions can be considered that the interactive

alignment is a universal cognitive mechanism for coordination of both behavioral and communicative actions.

5.1 The significance of the Results.

We are born into, and we utilize readymade communication systems and a set of conventions to collaborate in the context of usually well predictable real life task scenarios. However in the experimental scenario of the present study, we had the chance to compare the style and level of success of coordination at the earliest stages of the emerging communication system, when the couples lacked conventions and a shared lexicon. By the help of the findings on the trends of use ratio of speech acts, this study also enabled us to compare how a mature communication system functions for sharing information and coordination. When a communication system is mature, the parties do not need to order or request; the coordination is more like dancing, parties take necessary actions by only receiving the information contained in the asserted message. This is a good example of the "under-determination of meaning problem" described and explained by the "optimal relevance" concept of Dan Sperber (Wilson and Sperber 2006); meaning is recovered by and from the repository of shared world knowledge, experiences of coordination and collaboration.

The findings of the present study can also be utilized to detect and understand situations of collective decision making and problem solving. The analysis of communicative content from a speech act classification perspective can reveal a situation of crisis where the collective does not have conceptual or linguistic tools to understand and define a given problem. In such cases, we can expect to see directive speech acts to negotiate the meanings of the lexical items to deal with the new situation; requests and orders (depending on the power relation within the collective) serve to establish the patterns of coordinated action. There are several NLP studies trying to manage automated speech act classification using the syntactic and semantic features of surface lexical content (Ezen-Can and Boyer 2015, Appling et al. 2013). These studies generally just aim to achieve automated classification or to infer individual personality traits. The findings of this study however, can provide two kinds of inferences when the communicative content is analyzed in terms of speech act categories. First, abundance of directive or assertive speech acts indicates the level of efficiency of communication taking place in the collective (a team, or organization). Secondly, the speech act analysis may help detect a crisis phase (more directive speech acts) or solution phase (more assertive speech acts) when dealing with a novel problem.

5.2. Conclusion

The experimental constraints on the communication, task environment and task structure played a crucial role on the development and final shape of the communicative systems. Participants did not invent general skills of communication, having these skills already, couples utilized the shared constrains as cues to create and interpret the communicative behaviour of each other accordingly. As a result, shared task structure and shared requirements functioned as not only a starting point but always a common ground for joint action. The emerged communication systems were not a separate tool or formal system, they emerged as an integrated part of the problem/solution complex, i.e. the collaborative effort for the solution of the problem. The communication systems' scope, function and complexity is determined by context and shared constraints, which is in line with the conception of language as a complex adaptive system (Thagard, 2005).

> "This study exposes that current and past studies on the emergence of communication sharing the extended cognition, distributed cognition and joint action paradigms seem to be in dialogue and converge in terms of terminology, discussion of common theoretical problems and research agenda. However there is a lack of communication between joint action literature and the pragmatic/linguistic accounts of communication and emergence of

communication, which attempt to offer a cognitive architecture account of the cognition and communication problem, by presenting a dedicated sub-module of comprehension which includes intention detection and working principles of relevance theory. This study attempted to integrate the *cognitive/communicative principles of relevance* to the explanatory scheme of the joint action studies on the emergence of communication" (Ulubay, 2013).

Research and publication ethics statement

The research had been conducted in accordance with the ethical policies and regulations of Middle East Technical University and Informatics Institute in effect at the time.

Authors' Contributions to the Article

Murat Ulubay (100%). Since the co-authorship of supervisor for an article from doctoral theses with the doctorant is customary in Turkiye, the following explanation is deemed necessary: The article originated from the author's (Dr. Ulubay) doctoral thesis and written by Ulubay. The thesis advisor stipulated that for co-authorship, the advisor would need to make additional contributions beyond supervising the thesis. Therefore, the advisor recommended that the article be published solely by the author of the thesis study.

Statement of Funding

No funding support received. Not applicable.

Statement of Conflict of Interest

There is not any conflict of interest between the author or other stakeholders.

References

Appling, D.S., Briscoe, E.J., Hayes, H & Mappus, R.L. (2013). Towards Automated Personality Identification Using Speech Acts. Association for the Advancement of Artificial Intelligence.

Axelrod, R. (1981). The evolution of cooperation. Science, 211, 1390-1396.

Bach, K. (2005). Speech Acts. Routledge Encyclopedia of Philosophy. Retrieved from http://online.sfsu.edu/kbach/spchacts.html

Blanca, M. J., Alarcón, R., & Arnau, J. (2017). Non-normal data: Is ANOVA still a valid option? Psicothema, 29.4, 552–557. https://doi.org/10.7334/psicothema2016.383

Clark, A., & Chalmers, D. (1998). The Extended Mind. In Analysis (Vol. 58, Issue 1, pp. 7–19). Oxford University Press (OUP). https://doi.org/10.1093/analys/58.1.7

Clarks, H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. Cognition, 22, 1-39.

Field, A. (2009). Discovering statistics using SPSS (3rd ed.). SAGE Publications.

Galantucci, B. (2005). An experimental study of the emergence of human communication systems. Cognitive Science, 29, 737-767.

Galantucci, B. (2005). An Experimental Study of the Emergence of Human Communication Systems. In Cognitive Science (Vol. 29, Issue 5, pp. 737–767). Wiley. https://doi.org/10.1207/s1551670

Galantucci, B., & Sebanz, N. (2009). Joint Action: Current Perspectives. In Topics in Cognitive Science (Vol. 1, Issue 2, pp. 255–259). Wiley. https://doi.org/10.1111/j.1756 8765.2009.01017.x

Gerard, S., & Pickering, M. (2004). Why is Conversation so easy? Trends in Cognitive Science, 8(1), 8

Grice, P. (1975). Logic and conversation. (P. Cole, & J. Morgan, Eds.) Syntax and semantics: Speech Acts, 3, 41-58.

Grice, P. (1989). Studies in the Way of Words. Harvard University Press.

Hutchins, E. (1995). Cognition in the wild. The MIT Press.

Hutchins, E., & Johnson, J. (2009). Modelling the Emergence of Language as an Embodied Collective Cognitive Activity. Topics inn Cognitive Science, 523-544.

Levinson, S. C. (2000). Presumptive Meanings. The MIT Press.

https://doi.org/10.7551/mitpress/5526.001.0001

Nolfi, S. (2005). Emergence of communication in embodied agents: co-adapting communicative behaviours. Connection Science, 231 - 248.

Origgi, G., & Sperberger, D. (2000). Evolution, communication and the proper function of language. In P. Carruthers, & A. Chamberlain (Eds.), Evolution and the Human Mind: Language, Modularity and Social Cognition (pp. 140-169). Cambridge: Cambridge University Press.

Qadir, A., & Riloff, E. (2011). Classifying Sentences as Speech Acts in Message Board Posts. Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, 748-758. https://aclanthology.org/D11-1069

Port, R. (1999). Language as a Formal System.

https://legacy.cs.indiana.edu/~port/pap/daviton/formal.lg.html

Scott-Phillips, T. C., Kirby, S., & Ritchie, G. R. S. (2009). Signalling signalhood and the emergence of communication. In Cognition (Vol. 113, Issue 2, pp. 226–233). Elsevier BV. https://doi.org/10.1016/j.cognition.2009.08.009

Searle, J. (1969). Speech acts: an essay in the philosophy of language. Cambridge: Cambridge University Press.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. Trends in Cognitive Sciences, 10(2), 70-76.

Shintel, H., & Keysar, B. (2009). Less Is More: A Minimalist Account of Joint Action in Communication. In Topics in Cognitive Science (Vol. 1, Issue 2, pp. 260–273). Wiley. https://doi.org/10.1111/j.1756-8765.2009.01018.x

Shockley, K., Richardson, D., & Dale, R. (2009). Conversation and coordinative structures. Topics in Cognitive Science, 1(2), 305–319.

Sperber, D., & Wilson, D. (1995). Relevance: Communication and cognition. Oxford: Blackwell.

Sperber, D., & Wilson, D. (2006). Pragmatics. (F. Jackson, & M. Smith, Eds.) Oxford Handbook of Philosophy of Language, 468-501.

Steels, Luc (2003). Language re-entrance and the 'inner voice'. Journal of Consciousness Studies (4-5):174-185.

Steels, L. (2006). Experiments on the emergence of human communication. In Trends in Cognitive Sciences (Vol. 10, Issue 8, pp. 347–349). Elsevier BV. https://doi.org/10.1016/j.tics.2006.06.002

Stenning, K., Lascarides, A., & Calder, J. (2006). Introduction to Cognition and Communication. The MIT Press.

Thagard, P. (2005). Introduction to Cognitive Science. The MIT Press.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions The origins of cultural cognition. Behavioral and Brain Sciences, 28(5), 675-691.

Ulubay, M. (2013). Resilient Features of Re-Emerging Dyadic Communication Systems in an Interactive Virtual Environment. Ankara: METU. Retrieved from http://etd.lib.metu.edu.tr/upload/12615502/index.pdf

Vogt, W. (1999). Dictionary of Statistics and Methodology: A Non-Technical Guide for the Social Sciences (2nd ed.). Sage Publications.

Wagner, K., Reggia, J., Uriagereka, J., & Wilkinson, J. (2003). Progress in the Simulation of Emergent Communication and Language. Int*ernational Society for Adaptive Behavior, 11*(1), 37-69.