The Evolving Nature of Human-Device Communication: Lessons Learned from an Example Use-Case Scenario

Adam Csapo and Peter Baranyi

***Abstract*—Cognitive infocommunication channels are structured multi-sensory forms of communication capable of transmitting high-level concepts between systems with various levels of cognitive capabilities. Theoretical questions on how CogInfoCom channels are best designed to ensure effectiveness and comfort of use have been studied extensively in the past. However, in modern ICT environments, where humans are increasingly merging together and becoming entangled with the infocommunications infrastructure, the question of how CogInfoCom channels should evolve through time, depending on the context and on past interactions, is becoming equally important. In this paper, we discuss this question from various aspects based on a use-case scenario, and draw conclusions for future CogInfoCom design.**

I. Introduction

Cognitive infocommunications (CogInfoCom) is an emerging field at the meeting point of infocommunications and the cognitive sciences [1], [2]. The goal of CogInfoCom is to provide a systematic view of how cognitive processes can co-evolve with infocommunications devices, with special focus on the merging process which is occurring between humans and the ICT network surrounding them [3].

The phenomenon which primarily motivates research on CogInfoCom can be referred to as the merging or entanglement between humans and the infocommunications network. This merging can be observed at various levels, ranging from low-level connectivity at the cellular and electrotechnical level, all the way to the highest level of sensing collective behaviors such as mass movements, mass habits etc. As a result, humans (more generally, living beings) and infocommunications will soon coexist as an entangled web, resulting in an augmentation of natural cognitive capabilities.

As a matter of fact, it can be argued that a human capacity to become entangled with technologies has always existed, and that a new form of entanglement between humans and infocommunications can already be observed today. Neurophysiological evidence on the one hand that humans treat technologies as if they were natural extensions of themselves [4], [5], and psychological evidence on the other that e.g. humans are capable of developing new artificial sensory 'modalities' [6], [7] together point towards such an entanglement.

The subject of this paper is motivated by research on CogInfoCom channels – structured multi-sensory messages which carry information on high-level concepts [8], [9]. Many investigations in the past – even before the notion of CogInfoCom channels was formulated – have focused on the design of explicit messages to users with well-defined semantics at well-defined points of interaction (e.g., [10], [11], [12], [13]). However, relatively little attention has been focused on how communication can or should evolve through time – i.e. during interactions over extended periods of time – so that various communication channels can be comfortably incorporated by humans into a set of unnoticeable but readily accessible cognitive capabilities.

In this paper, we make the case for several arguments. First, we argue that the design of temporally evolvable communication can be described by key phases and transitions in biological communication processes. Second, based on empirical support from a use-case scenario, we argue that interactions during these various phases are characterized by different degrees of willingness to communicate consciously and purposefully, and by different degrees of immediacy. We conclude that these observations create constraints as to the kinds of communication that are effective in different cases, and that they may therefore inform the effective design of evolvable communication in future technologies.

The paper is structured as follows. In Section II, the definition of CogInfoCom channels is re-iterated, and a brief discussion is given on the parallels between CogInfoCom channels and biological communication, and the challenges in integrating CogInfoCom channels into natural communication processes. In Section III, an example use-case scenario between a human and an infocommunications device is outlined and analyzed from several aspects. A discussion on observations made follows in Section IV.

Adam Csapo and Peter Baranyi are with the 3D Internet based Control and Communication Laboratory, a joint laboratory between the Budapest University of Technology and Economics, and the Institute for Computer Science and Control of the Hungarian Academy of Sciences (e-mail: csapo.adam@tmit.bme.hu; baranyi.peter@sztaki.mta.hu).

II. CogInfoCom Channels: What are They and How Might They Emerge?

A. Definition of CogInfoCom channels

We begin our discussions by briefly re-iterating the definition of CogInfoCom channels and related terms (i.e., CogInfoCom icons, messages and message-generated concepts) based on [8], [9].

Definition II-A.1 (Icon layer)**.** *The set of sensory percepts that give rise to immediate and unique semantic interpretations. The* ***compound icon layer*** *contains sensory data combined from several modalities which give rise to immediate and*

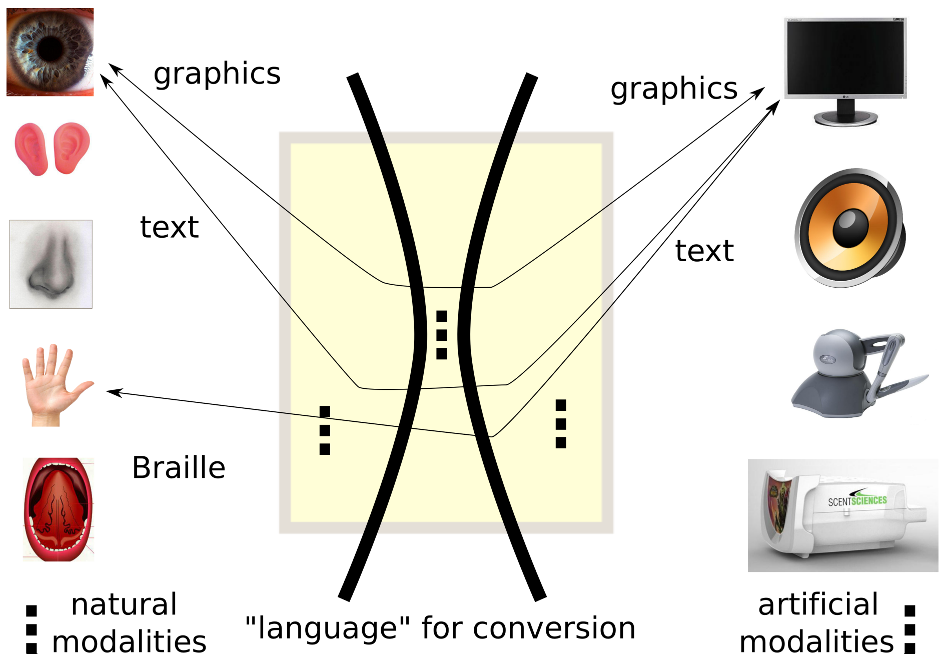

Fig. 1: The definition of CogInfoCom channels incorporates the human-centered, system-centered and representation-centered points of view.

unique semantic interpretations. Percepts contained in the (compound) icon layer are referred to as ***CogInfoCom icons****.*

Definition II-A.2 (Message layer)**.** *The set of abstract, sequential messages which are built up of elementary icons from the (compound) icon layer. Messages contained in the message layer are referred to as* ***CogInfoCom messages****.*

Definition II-A.3 (CogInfoCom concept layer)**.** *The set of all concepts which can be generated by the message layer in a given sensory modality. Concepts within the layer are referred to as* ***CogInfoCom message generated concepts****.*

Definition II-A.4 (CogInfoCom channel)**.** *A group of CogInfoCom messages used to carry information on the state of a CogInfoCom message generated concept.*

CogInfoCom channels offer a modality-independent view of structured and semantically meaningful sensory signals. In fact, CogInfoCom channels de-emphasize the interpretation of modalities: irrespective of whether the concept of 'modality' is interpreted as a characteristic of the human sensory modality which is used to perceive signals ("*human-centered view*" [14]), as a characteristic of the device used to generate the signals ("*machine-centered view* [14]), or as a characteristic of the way in which semantic meaning is mapped onto the perceiving human sensory modality (which we may refer to as the "*representation-centered view*", cf. [15]). In this way, there is no confusion in characterizing modalities in cases where the user reads text displayed on a screen, or reads text through tactile perception and Braille encoding, or views images on a screen (cf. Figure 1).

B. Phases of evolution in CogInfoCom channels

CogInfoCom channels – according to their original definition [8] – represent only a very specific kind of communication which occurs at an explicit and conscious level. However, based the pervasiveness and success of more implicit forms of communication in biology, it seems to be a viable approach to broaden the scope of artificial behaviors which could be interpreted as lower-level communication capabilities in infocommunication technologies[1].

Two concepts which emerge in biology and which have no well-defined parallel in engineering design are *cues* and *signals* [19]. Cues are behaviors which do not in themselves qualify as a form of communication *per se*, but which can evolve – through a process referred to as *ritualization* – into purposefully generated *signals* if they are perceived as useful in eliciting predictable and useful responses from other individuals in the population.

CogInfoCom channels can be added as a third level to this framework of cues and signals, and conceived of as an explicit form of communication which evolves based on variations in signals: each minor variation would then convey a different state of the same high-level concept (e.g., in much the same way as different pitches and speeds of human speech can convey different levels of calmness, anxiety, etc.).

The need for such a hierarchy in the communication between humans and artificial systems, both in terms of qualitative and temporal aspects, can be supported based on several considerations:

- It is unlikely that CoginfoCom channels can emerge spontaneously without any form of prior interaction (unless

[1] It should be noted that several authors have advocated in the past to incorporate biologically inspired communication mechanisms in artifical systems [16], [17], [18].

they are designed explicitly and taught to the user before his or her interaction with the system); cognitive systems need to communicate for some time before they are capable of picking up on subtleties which will lead to the emergence of CogInfoCom channels.
- While it is plausible that signals can exist without prior use of cues (e.g., if the designer a system implements a set of signals which the user will then become accustomed to), a system without a minimal aptitude for adaptivity (i.e., unable to adapt to the user's habits and requirements) will generally be perceived as rigid, inflexible and even annoying.
- The different levels of communication defined by cues, signals and channels require different levels of attention. A system that does not demand all of the user's focused attention immediately, but rather adapts to elicit increasing amounts of attention as the user becomes accustomed to using it will in general be regarded as more pleasant to work with [2].

C. Challenges behind implementing evolution in CogInfoCom

The evolution from cues to signals to channels can be regarded as a natural tendency in CogInfoCom. However, the question still remains: how can this tendency be implemented in a way that could unfold in the communication between humans and artificially cognitive systems, irrespective of the specific details of the interaction and of the application domain?

The challenge which lies behind this question is significant. If the starting point of the designer is to consider solely the definition or meaning of cues, signals and channels, then enabling an artificially cognitive system to develop its own signals and channels would require the designer to also implement some way for the system to recognize what external events (e.g., human behaviors) are "useful" to it and which of its cues were successful in eliciting those events. Clearly, just the definitions of these stages of communication will not help engineers in making good design choices. Instead, finding a set of characteristics which transcend the notions of cues, signals and channels, but which can nevertheless be brought into connection with these categories may be a viable solution. More plainly: if there exist a set of characteristics which are applicable to cues, signals and channels, but which are also different in each of these communication forms, then they can be taken as a basis for CogInfoCom design. Such characteristics will be described based on an example use-case scenario in the following section.

III. Example Scenario

The example scenario focuses on the functionality of an intelligent alarm clock that keeps track of how much the user has slept in the previous few days and also monitors how refreshed or tired the user is several times during the day. Although much of this information can be manually kept track of by the user, or can otherwise be queried explicitly by the device, the purpose of the scenario is to demonstrate that communication can occur at various cognitive levels, and that different levels of communication are better suited to different contexts and situations.

A. When the device queries the user

The first important observation is that information can be queried by the device using different levels of communication. Let us consider the following ways in which the device may learn the time at which the user decides to go to sleep:

1) The device may ask the user explicitly about bedtime based on the current time.
2) The device may ask the user explicitly about bedtime based on the user's actions, such as dimming the light, spending more time in the bedroom, etc.
3) The device may ask the user explicitly about bedtime while the user is updating the settings of the alarm clock.
4) The device may make implicit efforts to ascertain the user's bedtime by making statements about its assumption that the user is about to go to sleep.
5) The device may ask the user implicitly about bedtime by making statements about how tired it is, asking the user if its services are still needed for the day.

It is clear that several aspects influence the perception of each of these different kinds of queries, e.g.:

- Whether the user has to respond to the query immediately. In cases 1, 2, 3 and 5, the way in which questions are posed demands that the user respond to the query. In case 4, the user will not necessarily feel that a response is needed, and can more easily decide not to take notice of what the device is saying.
- Whether the user's conscious actions elicited the query. In cases 2 and 3, the user's conscious actions serve to elicit the query. Further, it may potentially be that the user's conscious actions serve to elicit the query in case 1 as well, if the user instructs the device beforehand to make the query at the given time. In the remaining cases an event that is outside of the user's conscious influence is the culprit in eliciting the query.
- Whether the user's purposeful actions elicited the query. In case 3, the user's purposeful actions serve to elicit the query (in the sense that the purpose of the actions is to work with the alarm clock). Further, it may potentially be that the user's purposeful actions serve to elicit the query in case 1 as well, if the user instructs the device beforehand to make the query at the given time.
- Whether the user expects to receive a query at the time when it is received, or if the query causes surprise. This aspect is somewhat related to the previous questions on consciousness and purposefulness, however, the time at which the conscious and/or purposeful action takes place may or may not directly precede the time of the actual query, which in turn can influence the degree of surprise caused by the query. For example, if the user sets a time

[2] The emphasis here is on focused attention relevant to a functionality. In other words, while no system should require the user's devoted attention all the time, the attention necessary to achieve a task can and should vary: later on in the user's interaction history, the achievement of tasks should be perceived as purposeful and smooth.

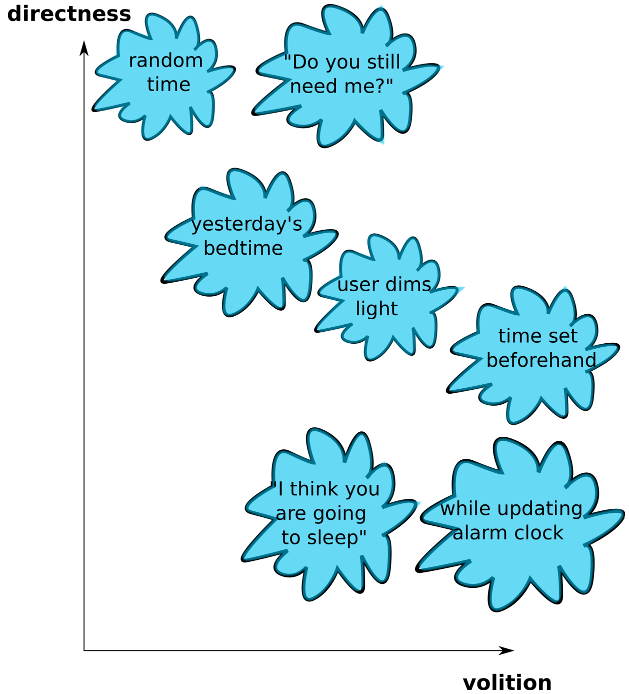

Fig. 2: There are several ways in which a device might try to obtain information from the user as to when he or she plans to go to sleep. The different approaches are characterized by unique combinations of degree of volition and degree of directness.

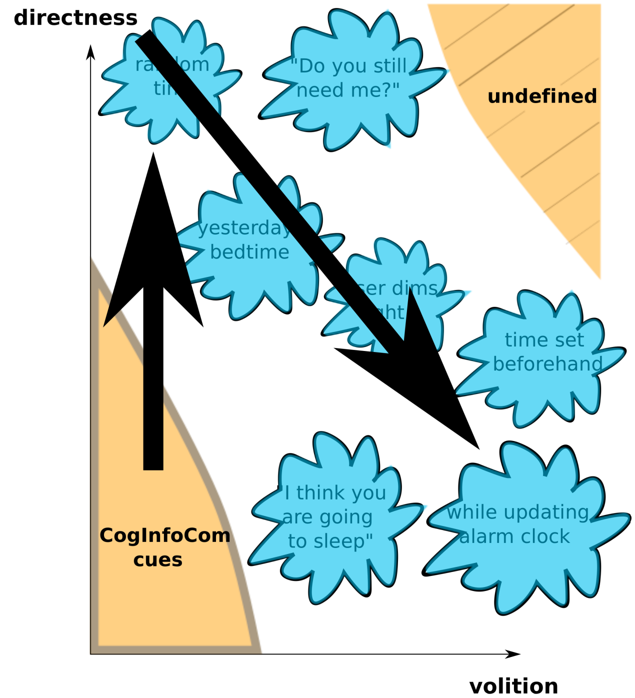

Fig. 3: Behaviors which qualify as involuntary and indirect (from the user's perspective) can be regarded as non-communicative behaviors – e.g., noises, vibrations, etc. – that are unique to the device (i.e., this region contains CogInfoCom cues). The natural flow of communication is depicted by the arrows from involuntary and indirect, through involuntary and direct, to semi-voluntary and semi-direct.

at which the query should be generated, and does so hours beforehand (e.g., case 1), the query may cause more surprise than cases 2 and 3. Case 2 may in turn cause more surprise than case 3. Similarly, if the user does not set a time at which the query should be generated, but the timing of the query nevertheless correlates with an event in the past – for instance, the time at which the user went to sleep the day before – then the query will still cause less of a surprise than in the case when it occurs at a time that seems completely random to the user.

The example demonstrates the fact that various forms of communication can be categorized as different depending on whether the user is surprised by a query and/or is required to respond immediately, and whether the user's conscious and/or purposeful actions are needed to elicit the query. We propose to use the terms *directness* and *volition* to describe these aspects of interaction.

The different cases outlined above are shown in terms of these two concepts in Figure 2. Also conceptually depicted on the figure are two general regions in the volition-directness plane: first, the voluntary and direct region has no meaning (at least not in the system-queries-user scenario, cf. Section III-C), and second, the involuntary and indirect region comprises "non-communicative behaviors" – i.e., behaviors which are naturally exhibited by the device in various contexts but which as yet have no meaning attributed to them. It is interesting to note the correspondence between this region and the notion of *cues* in biological communication.

B. The flow of communication

Based on the example, an important question arises: if there are so many ways in which the device can elicit the user's attention, is there a preferred method which should always be used, or should the way in which the device queries the user evolve through time? Further, if communication should evolve, then is there a preferred "trajectory" which could inform the designer of the device as to which forms of communication are preferred when? Potential answers to these questions are discussed in Section IV.

Based on extensive research on the way humans like to communicate with each other, and also with machines (cf. e.g. [20]), it can be established that humans prefer communication to evolve through time. Having a sense that the communicating partner understands us better than when we first met is always a positive experience. Having a sense that a communicating partner is willing to change his or her personality traits (e.g., verbosity, mood, etc.) to match ours – even if our personality traits do not match to begin with – is also a positive experience. Conversely, if the flow of communication goes against these tendencies, we are left with a feeling of discomfort and even frustration.

Returning to the above example, it can be conjectured that users will be left with the best experience if communication evolves from the top-left corner (i.e., low volition, high directness) towards the lower-right half of the diagram (i.e.,

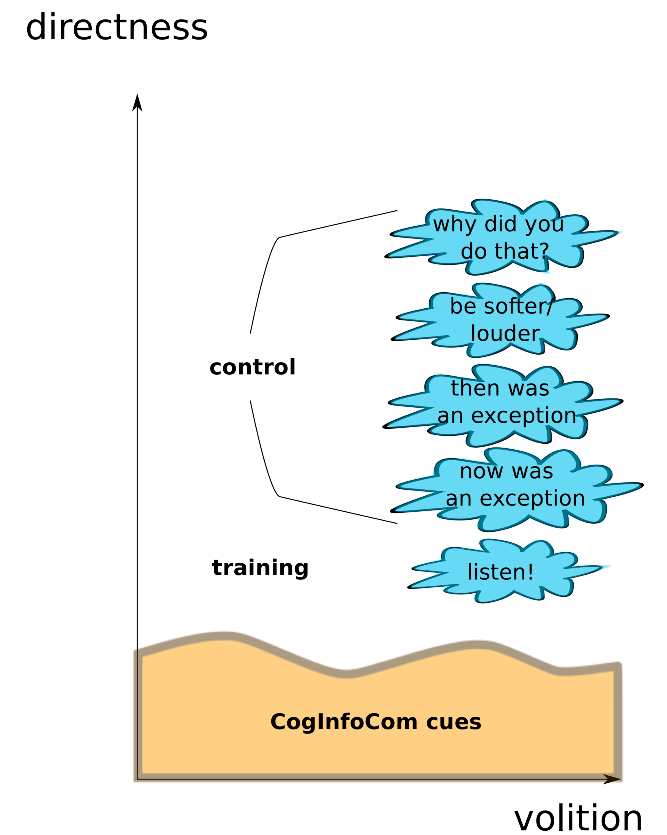

Fig. 4: When the user aims to understand the functionality of the device and to influence its future behavior, all control and training messages are voluntary, and range from semi-indirect to semi-direct.

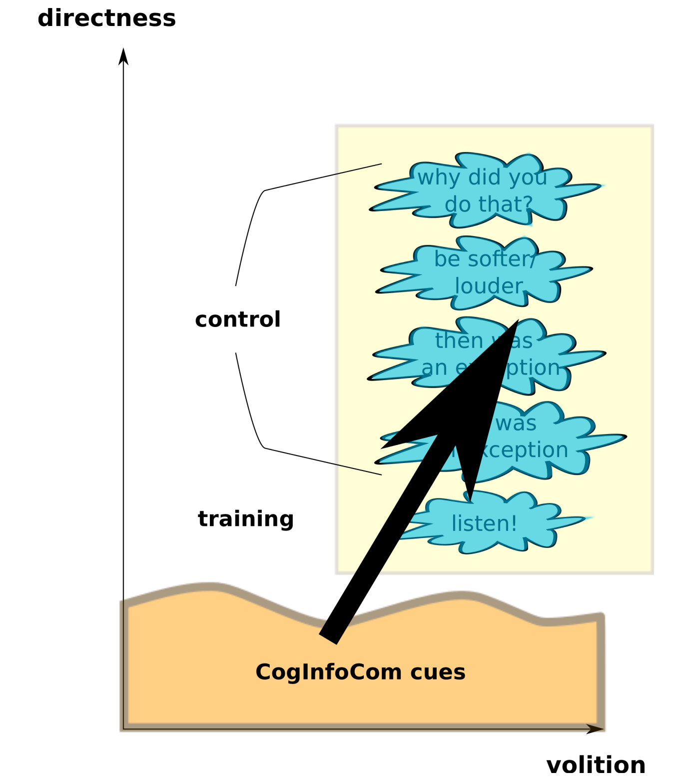

Fig. 5: The natural flow of communication from the user to the device is depicted by an arrow from indirect to semi-direct and voluntary communication.

relatively high volition and relative indirectness). The explanation for this observation is that early on in the interaction history between the user and the device, having the device try to ascertain what the user is doing based on environmental cues can be prone to errors, and will lead to frustration if attempted too early. In the early phases of interaction (and also during error recovery, should something go wrong), it is inevitable that the device query the user irrespective of whether the user wants it or not, and expects it or not; and as long as this phase does not draw out for too long, users will appreciate the progress made in the communication process rather than be frustrated with early transients. The role of CogInfoCom cues is to precede (and overlap somewhat) with the early phase of involuntary and direct communication, so as to allow the device to "tease out" from its environment those behaviors which are effective in gaining the user's attention (Figure 3).

C. When the user queries the device

Communication is a bi-directional process, hence, the example outlined in section III-A describes the entire communication process only partially.

If we consider the other direction of interaction, namely the case where the user tries to obtain information from – or otherwise tries to influence the behavior of – the device, analogous observations can be made. First, let us consider as a starting point those events – in other words, "stations" during interaction – which may prompt the user to intervene in various ways:

1) In general, the user will wish to know why, or based on what sensors, the device generated the query.
2) *When the device emits CogInfoCom cues:* the user may wish to know what those cues represent, or ignore some of them altogether instead. The device will adapt to apply those cues as CogInfoCom signals which successfully elicit the user's attention.
3) *When the device asks the user explicitly about bedtime:*
    a) *based on bedtime the day before:* the user may wish to communicate that the day before was exceptional, or that the current day is exceptional.
    b) *based on "random" time – i.e., for no reason that is related to the user:* the user may ignore the query, and later wish to tell the device to make similar queries softer or louder in the future.
4) *When the device asks the user explicitly about bedtime based on the user's actions such as dimming the light, spending more time in the bedroom, etc.:* the user may wish to calibrate the device's sensitivity if its event recognition was a false positive. The user may also wish to "train" the device through the use of *corroborative stimulation* [8], i.e. by letting the device know that he or she is about to sleep and performing various actions (e.g., dimming the light, snapping a finger, etc.) at the same time.

The example interactions can also be divided into different categories: generic ***control messages*** (e.g.,"why did you do that?", "yesterday was an exception", "now is an exception", "be louder next time", "be softer next time") and specific ***training messages*** (e.g., "listen, I am going to sleep now", "I am paying attention to you now because that cue was subtle and I liked it").

It is important to note that the above message types are also amenable to categorization in terms of *directness*. For example, the control message which queries the reason why the device emitted a behavior (*"Why did you do that?"*) is somewhat direct in the sense that it shows the user's surprise and perhaps frustration as well (at the same time it is not entirely direct as it is not tied to real-time constraints). Conversely, training messages in general do not reflect surprise, but are instead "future-oriented" in the sense that through them the user can expect better modified behavior from the device in the future.

In terms of *volition*, it can be ascertained that all of the control and training messages given as examples are voluntary, however, involuntary cues are still relevant in user-to-device communication (these are the cues, for example, which are taken as a basis for various behaviors in case 4). These aspects of the example messages are shown in Figure 5.

IV. Discussion

Based on the example, we can make the strong hypothesis that a natural flow of communication exists within the directness-volition plane when users interact with artificially cognitive devices. Through the example, we have demonstrated that the direction of this flow is different when the user aims to influence the behavior of the device, from when the device aims to obtain feedback from the user. In simple terms, these differences reflect the fact that when communication is targeted at humans, direct queries can be perceived as more intimate – and, hence, intrusive if done without improper preparation, whereas in the opposite direction (from humans to artificial systems), direct interaction can be more efficient even in earlier stages of communication.

The dimensions of volition, directness and the interaction types outlined by the terms *control messages* and *training messages* are the kinds of characteristics mentioned earlier in Section II-C; namely, characteristics which are interpretable in terms of cues, signals and channels, and which at the same time are suitable to distinguish between the three levels of communication. By relying on these distinguishing characteristics, engineers may in the future szbe able to design CogInfoCom systems that are adaptive in their communication capabilities.

The discussions in the paper also support the idea that channels can be amenable to a process of gradual emergence, whereby communication forms that were signals at an earlier time develop into more graded, subtle forms of communication. Apart from explicit, goal-oriented training, this is achievable if a minimal level of understanding exists between the user and the system following the use of cues and signals.

V. Conclusion

In past research on cognitive infocommunication channels and human-device communication in general, relatively little attention has been focused on how communication can or should evolve through time, so that various communication forms can be comfortably incorporated by humans into a set of unnoticeable but readily accessible cognitive capabilities.

In this paper, we argued that the design of temporally evolvable communication can be described by the key phases and transitions in biological communication processes, defined by cues, signals and channels. Based on empirical support from a use-case scenario, we argued that interactions during these various phases are characterized by different degrees of willingness to communicate consciously and purposefully, and by different degrees of immediacy and surprise. We referred to these aspects as volition and directness. We concluded that these observations create constraints as to the kinds of communication that are effective in different cases, and that they may therefore inform the effective design of evolvable communication in future technologies.

Acknowledgment

This research was realized as part of the Ányos Jedlik PhD candidate scholarship, in the frames of TÁMOP 4.2.4. A/1-11-1-2012-0001 "National Excellence Program – Elaborating and operating an inland student and researcher personal support system", which was subsidized by the European Union and co-financed by the European Social Fund.

References

[1] P. Baranyi and A. Csapo, "Cognitive infocommunications: CogInfoCom," in *Computational Intelligence and Informatics (CINTI), 2010 11th International Symposium on*, Budapest, Hungary, Nov. 2010, pp. 141–146.

[2] ——, "Definition and synergies of cognitive infocommunications," *Acta Polytechnica Hungarica*, vol. 9, pp. 67–83, 2012.

[3] G. Sallai, "The cradle of cognitive infocommunications," *Acta Polytechnica Hungarica*, vol. 9, no. 1, pp. 171–181, 2012.

[4] J. M. Carmena, M. A. Lebedev, R. E. Crist, J. E. O'Doherty, D. M. Santucci, D. F. Dimitrov, P. G. Patil, C. S. Henriquez, and M. A. L. Nicolelis, "Learning to control a BrainMachine interface for reaching and grasping by primates," *PLoS Biol*, vol. 1, no. 2, p. e42, Oct. 2003.

[5] A. Maravita and A. Iriki, "Tools for the body (schema)," *Trends in cognitive sciences*, vol. 8, no. 2, pp. 79–86, 2004.

[6] E. Platzer and O. Petrovic, "An experimental deprivation study of mobile phones, internet and TV," *Computer Technology and Application*, vol. 2, no. 8, pp. 600–606, 2011.

[7] M. Drouin, D. H. Kaiser, and D. A. Miller, "Phantom vibrations among undergraduates: Prevalence and associated psychological characteristics," *Computers in Human Behavior*, vol. 28, no. 4, pp. 1490–1496, Jul. 2012.

[8] A. Csapo and P. Baranyi, "A unified terminology for the structure and semantics of CogInfoCom channels," *Acta Polytechnica Hungarica*, vol. 9, no. 1, pp. 85–105, 2012.

[9] ——, "CogInfoCom channels and related definitions revisited," in *Intelligent Systems and Informatics (SISY), 2012 IEEE 10th Jubilee International Symposium on*, Subotica, Serbia, Sep. 2012, pp. 73–78.

[10] M. Blattner, D. Sumikawa, and R. Greenberg, "Earcons and icons: Their structure and common design principles," *Human Computer Interaction*, vol. 4, no. 1, pp. 11–44, 1989.

[11] J. J. Kaye, "Making scents: aromatic output for HCI," *interactions*, vol. 11, no. 1, pp. 48–61, 2004.

[12] N. B. Sarter, "Multimodal information presentation: Design guidance and research challenges," *International journal of industrial ergonomics*, vol. 36, no. 5, pp. 439–445, 2006.

[13] D. L. Riddle and R. J. Chapman, "Tactile language design," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 56, 2012, pp. 478–482.
[14] K. Jokinen, "User interaction in mobile navigation applications," in *Map-based Mobile Services*, ser. Lecture Notes in Geoinformation and Cartography, L. Meng, A. Zipf, and S. Winter, Eds. Springer Berlin Heidelberg, Jan. 2008, pp. 168–197.
[15] F. Vernier and L. Nigay, "A framework for the combination and characterization of output modalities," in *Interactive Systems Design, Specification, and Verification*, ser. Lecture Notes in Computer Science, P. Palanque and F. Paterno, Eds. Springer Berlin Heidelberg, Jan. 2001, no. 1946, pp. 35–50.
[16] C. Kline and B. Blumberg, "The art and science of synthetic character design," in *Proceedings of the AISB 1999 Symposium on AI and Creativity in Entertainment and Visual Art*, 1999.
[17] T. Fong, I. Nourbakhsh, and K. Dautenhahn, "A survey of socially interactive robots," *Robotics and autonomous systems*, vol. 42, no. 3, pp. 143–166, 2003.
[18] G. Lakatos and A. Miklosi, "How can the ethological study of dog-human companionship inform social robotics?" *Crossing Boundaries: Investigating Human-Animal Relationships*, vol. 14, p. 187, 2012.
[19] T. C. Scott-Phillips, R. A. Blythe, A. Gardner, and S. A. West, "How do communication systems emerge?" *Proceedings of the Royal Society B: Biological Sciences*, vol. 279, no. 1735, pp. 1943–1949, 2012.
[20] C. I. Nass and C. Yen, *The man who lied to his laptop: what machines teach us about human relationships*. Current Trade, 2010.

Adam Csapo is a PhD candidate at the Dept. of Telecommunications and Media Informatics of the Budapest University of Technology and Economics. He is affiliated with the 3D Internet based Control and Communication Laboratory, a joint laboratory between the Budapest University of Technology and Economics, and the Institute for Computer Science and Control of the Hungarian Academy of Sciences. His research has focused primarily on the design and analysis of cognitive infocommunication channels (http://www.coginfocom.hu).

Peter Baranyi is professor at the Budapest University of Technology and Economics, and head of the 3D Internet based Control and Communication Laboratory, a joint laboratory between the Budapest University of Technology and Economics, and the Institute for Computer Science and Control of the Hungarian Academy of Sciences. He holds a PhD degree in computer engineering and a DSc degree at the Hungarian Academy of Sciences. His work focuses primarily on nonlinear system control design (related to qLPV and LMI based theories), cognitive infocommunications (http://www.coginfocom.hu) and 3D Internet based control and communications (http://www.3dicc.sztaki.hu).