Chapter 1
Methods for Analyzing Cognitive Architecture

Carl Vogel and Anna Esposito

Abstract We articulate aspects of our methods for analyzing human cognitive architecture. We focus on two approaches with a linked example. One approach is to analyze human cognitive abilities to isolate functions that have particular use in some processing domains (like language), with potential re-use in other domains. Here we focus on symmetry perception and production as may be realized in the function, *expect a copy* or a specialization of that function, *expect a complemented copy*. We illustrate the benefits of assuming such an operator is available to language processing by demonstrating that it allows reconciling empirical facts of human language processing with the natural predictions of formal language theory. The other approach is in simulation systems, parameterized to afford exploration of the effects of assuming that posited operators behave in particular ways. We link this to the *expect a copy* function through this being fundamental to habituation, which is requisite in simulation systems that model evolving behaviours and social constructs, like natural language.

1.1 Introduction

Cognitive infocommunications has been put forward as a scientific discipline that has as a focus the advancement of human cognitive capabilities [6, 7, 9]. The notion that advances to human capabilities may be assimilated into the concept of humanity has been analyzed, resulting in a theory being offered which holds that necessary pre-requisites for technologies to be successful in that sense is that they provide solutions to problems that many humans face privately and publicly [72, 73]. Human

---

Carl Vogel
Trinity College Dublin, the University of Dublin, e-mail: vogel@tcd.ie.

Anna Esposito
Università degli Studi della Campania "Luigi Vanvitelli", Caserta, Italy and International Institute for Advanced Scientific Studies (IIASS) Vietri sul Mare, Italy, e-mail: iiass.annaesp@tin.it

thought is a representative component of the private sphere of human existence, and communication is representative of the public sphere. Frequently, technologies are brought about in order to address issues in one sphere and are readily adaptable to distinct problems in the other sphere. While serendipity plays a role in this, it is important to create conditions that enable serendipity. Key serendipity-enabling conditions are intellectual curiosity and active pursuit of answers to questions, in this case, answers that yield clear ideas about problems that people face and deep understanding of humans as individuals and as they interact in groups. Constituent in that is having viable theories of human cognitive architecture. A theory of cognitive architecture amounts to a theory of the elements in human cognition that have a causal role in thought and behaviour.

Well-principled inventories and desirable feature lists (desiderata) for cognitive architecture are available in the literature [45, 68]. One influential argument includes the ten desiderata for autonomously developing agents, listed below in (1.1) [68, pp. 120-123].

- value systems and motives
- physical embodiment
- sensorimotor contingencies
- perception
- attention
- prospective action (1.1)
- declarative and procedural memory
- multiple modes of learning
- internal simulation
- constitutive autonomy

Our purpose here is to suggest methods for validating theories of human cognitive architecture. The approach that we suggest has both analytical and empirical elements, with attention to creating means for collecting positive and negative evidence for hypotheses. Having such methods in place supports verification of theories of cognitive architecture, and this in turn enables the identification of locations in which performance may be enhanced. This may be through enhancing the capabilities of robot systems or through providing humans with extended capabilities. For an analogy of the difference, one may consider the alternatives of sending a robot drone to inspect the state of fruit forming out of reach on a tree, or one may use binoculars to view the fruit oneself. The argument here draws upon and extends work that we presented recently [71].

Much of the demand for sophisticated human-machine interfaces originates in social needs that are emerging in a pronounced fashion in contemporary society. For example, the world's population is ageing, and this creates extreme challenges for maintaining social welfare, not least because of the cost of human labour. Robots that are acceptable for people to use may benefit from having intelligence levels that are close approximations of human intelligence. Seeking to achieve human level automaton intelligence may have a greater chance of success if the search is conducted in cognisance of methods that have been used in developing models of human cognitive architecture and to validate those models. Within this, one must: firstly, adopt or develop appropriate algorithms operating over multimodal data containing social signals to achieve the necessary pre-processing which individuates values of pertinent features; and, secondly, derive computational models that satisfy aesthetic principles of speed and simplicity that are effective, with accuracy similar or superior to that of humans in the same tasks, in the classification of highly structured objects and relations among objects, sensitive to change over time. Emotional expression of feelings is an example of a dynamic and deeply structured relation among sets of humans, who themselves are highly structured objects.

Robust artificial cognitive architectures, for robots, need not be isomorphic to human cognitive architectures in order to be effective. For example, robots may be designed to achieve success when measuring success using criteria that depend on a particular task, ignoring the means by which humans approach the same tasks, without replicating the cognitive processing humans require to achieve them. A different methodology is to model robots on the basis of what is known about human cognition and behaviour. An important strand of research in cognitive science is in modelling, studying the properties of simulation systems, such as simulated interacting agents, to identify how properties of the simulations map on to properties of the corresponding human interactions. Thus, simulations provide a means of testing theories of interactions. It makes every sense to pursue an overall programme of research into possible and plausible theories of human cognitive systems and their underlying architecture making use of the findings that emerge from simulation systems – one may gain knowledge about the corresponding qualities of humans and one may gain knowledge that has the potential to transfer productively to task-oriented robotic systems.

Two methods of research into cognitive architectures, each complementing the other, are our focus here. In one approach, behavioural mechanisms are hypothesized and studied in order to determine the extent to which those mechanisms aptly describe human cognition and behaviour, directly. In another approach, such mechanisms are studied indirectly, through their instantiation or emergence in simulations, and the simulation behaviours are evaluated in terms of the degree to which they correspond to aspects of human cognition and behaviour. Both of these directions inherit insights from cognate disciplines, for example, cognitive psychology and linguistic theory. We demonstrate here the efficacy of these methods by attending to symmetry identification as a single behavioural mechanism, showing the role of this mechanism in analysis of linguistic data and in simulation systems that model the evolutionary emergence of language.

When adopting the first method, specific mechanisms, such as "symmetry identification", are treated as cognitive operations hypothetically available to the human cognitive system, and human behaviours are analyzed across a range of behaviour types to determine where and how often the proposed mechanisms are reasonable to invoke as explanatory. If sufficiently attested, the mechanism may be hypothesized as a component within human cognitive architecture and as a candidate for replication in artificial cognitive systems. A role for symmetry identification may be noticed in visual processing and in language processing. Symmetry perception has a number of distinct forms: identification of copies, identification of copies subject to transformations like mirroring and so on [66]. Human capacities for symmetry perception are isolated as a specific mechanism that may have a role in human cognitive architecture, hypothesized as accounting for anomalies in the alignment of predictions that arise from formal language theory [39] with reality, in data associated with human linguistic behaviours [2, 76, 78]. Behaviours associated with the production of symmetry are integral to synchrony in dialogue [27]. Interlocutors' self-repetition and repetition of each other have been explored in quantifying interpersonal engagement in dialogue and have been assessed as proxy measures of mutual understanding achieved in dialogue [53, 54, 70].

With the second method, simulations are configured and conducted, and their emergent behaviours and properties of those behaviours are analyzed to assess their goodness of fit with human counterparts. A sizeable community of scholars explore hypotheses about language evolution using simulations, typically of agents interacting in stylized representations of the actual world [40, 60, 42, 44, 61, 62, 15, 16, 50, 69, 3, 4]. As the parameters of simulations lead to interaction patterns that conform with or diverge from patterns visible in human languages or human interactions, the results contribute to support or refutation of theories about the role of those parameters in the actual human language evolution. While some of these works address simulations that are embodied and situated (e.g. [62]), a lesson from cognitive neuroscience is that neither physical embodiment nor complete knowledge of the processes that ground symbols to physical objects are necessary preconditions to increasing what is understood about the physical neural layer upon which human cognitive architecture depends [47].

This article is not the place for a comprehensive review of relevant literature from linguistics, psychology, sociology and cognate disciplines nor for an inventory of all relevant facts about human cognition and interaction. They are all important to cognitive architectures, and where there is an intent to dwell on cognitive architectures in support of robotics research it is urgent to pay particular heed to findings regarding modalities of sensation, both in isolation and in their fusion [14, 30, 26, 11]. One reason to note these is that sensory modalities have a clear biological substrate. Sensory modalities depend on mechanisms implemented in biology and which contribute to the processes that fuse information arising from each modality; however, at the same time, there is ample evidence of aspects of processing and integration of visual and auditory signals that appear to be anchored in gender and culture, sociological constructs, more than biological [56, 31, 55]. Therefore, autonomous robots developed for the purpose verifying theories about human thought and interaction require mechanisms for learning, with attention to signal fusion [28, 29].

Research using these complementary methods towards clarifying the essential and accidental qualities of human cognition and human communication is not yet complete; further work is necessary and must be conducted with attention to results that emerge from additional methods. We state this imperative not because we alone use these methods, but because we feel that these methods have provided much of what is considered knowledge about human cognitive architecture. Consider an example success of this approach. It is a hypothesis of Pinker [51] that humans have as part of their congenital endowment an innate knowledge of predicate-argument structure in language, something like the theoretical fundamentals of lexical functional grammar [43], and that this innate knowledge allows the bootstrapping of language acquisition during development. This innate knowledge may be thought of as a hypothetical mechanism that is part of the cognitive architecture. Support for the hypothesis that this mechanism may be a component within the human cognitive architecture is visible in the fact that many sorts of simulation system "hard code" structural preferences for predicate-argument structure in language, and which this, in turn, forms part of the basis for emergent properties that ensue, such as related syntactic constraints. In this example, both sorts of methods contribute. Examples like this demonstrate that the research method we support here is successful in contributing knowledge about cognitive architectures and in a manner that could move forward the state-of-art in robotic systems.

1.2 Adaptable functions

A robust and productive research tradition in psychology has followed from the individuation of an inventory of cognitive faculties. Within this tradition, Fodor's modularity of mind hypothesis [33] supposes that the human cognitive architecture includes a small number modules defined by highly specific functioning and called "input modules". Relevant properties include processing speed, whether voluntary or involuntary, whether information encapsulated, and so on. Of primary importance are the input and output specifications. Given such a specification of an inventory of modules functionally defined, an agenda for research follows: it becomes necessary to verify the input and output behaviours that are in turn used to define each faculty, to establish what processes occur within each faculty that impinge on the output behaviours given the input behaviours, to establish whether the functional specification is correct, and so on.

"Vertical" faculties [33, p. 21] are thought of as functionally specific, enabled by discernible substrate of neural mass, with designated computational resources ("computationally autonomous"), and with these physical dimensions determined by genetic encoding. "Computational autonomy" means that resources that support processing within the functional domain are not shared with other modules and this means that information required within a module is not by default shared with other modules, except through the input-output parameters. This is the essence of "information encapsulation". An example of a vertical faculty in this sense is vision. The "subsumption architecture" as used in the design and development of autonomous robots is well aligned with the modularity of mind hypothesis. This architecture avoids "shared memory between computational elements" [13, p. 3] – thus, that the subsumption architecture and vertical faculties both suggest information encapsulation. "Horizontal" faculties are defined as a complement to the vertical faculties, and thus lack specificity and share resources. In conceptualizing a cognitive architecture on the basis of horizontal faculties, one might think about memory, attention, judgement, etc. and explore, for example, the extent to which these faculties influence each other. From the definition, one is not led to the immediate conclusion that memory of visual scenes is fundamentally different to memory of auditory scenes, although individual differences may arise.

	Vertical	Horizontal
Resource autonomy	+	-
Designated neural substrate	+	-
Encapsulation of information	+	-
Candidate faculty	Vision	Memory

Table 1.1 Horizontal vs. Vertical: Differences in individuating cognitive faculties in accordance with Fodor's modularity of mind hypothesis

It is not an entailment of Fodor's theory that the architecture of human cognition comprises exclusively horizontal or solely vertical faculties. Cognitive faculties understood as syntheses of the horizontals and verticals also make sense to consider. "Diagonal faculties" are those which have properties shared with both horizontal and vertical faculties. Some of these may depend on a designated neural substrate but lack autonomous control over that substrate, thereby either being subject to resource competition or information sharing, for example. Diagonals might lack content domain specificity, but with varying processing priority in each domain of use, for example – sound perception may be a diagonal that has differing processing urgency depending on whether the aural source derives from music or from language, or from the natural environment. Some evidence suggests that concepts tied to affordances of objects (for example, "decorate self" is an affordance of necklaces, combs, etc.) are in some cases linked to regions of neural substrate *distinct* from the neural regions linked to recognition of the artifacts themselves that may be manipulated to achieve their afforded functions [46]. In those cases, the neural substrate integral to recognizing object function is distinct from the substrate that manages the manipulable objects' recognition. However, that distinct neural substrate is involved in each case is not support for the conclusion that the best-fitting cognitive architecture is one in which there are two separate vertical faculties at stake, one that enables inference regarding object function and another that supports inference regarding how to actually achieve that function through manipulation of the object. A cognitive architecture that includes an inventory of diagonal faculties might be more easily motivated. This could include some such faculties that require information encapsulation that depends on content domain in the way just described but which is also in competition for resources with other faculties to achieve required computation and processing. The desirable features of cognitive architectures (such as in (1.1), as indicated above) may be achieved through modules that are vertical, horizontal or diagonal. The space of viable models of of cognitive architecture is is further clarified by analyzing the space of defensible diagonal faculties, since these ultimately may be necessary components of the human cognitive architecture. For this reason, we next consider one such candidate diagonal faculty in the role of symmetry recognition to visual perception, a vertical faculty, and language, a horizontal faculty.

1.3 Symmetry

Recognizing that two objects are instances of the same type is a form of copy identification, and copy identification is an example of being aware of symmetry [66]. The ability to identify copies cuts across desiderata (see (1.1) above) for human cognitive architectures [68]. That is, this kind of symmetry awareness is relevant to perception, to both procedural memory and declarative memory, and it is also integral to cognitive simulation. Copy identification has been posited as a cognitive function that can then reconcile empirical facts about human language processing with the complexity predictions suggested by formal language theory [76]. Using the worst-case asymptotic time-complexity analysis of grammaticality judgements for fully abstract families of languages corresponding to the Chomsky hierarchy,[1] one would expect that languages that avail of constructions requiring more than strong context-free expressivity should require more processing time of their native speakers or greater error rates for native speakers of those languages, in comparison with native speakers of languages requiring no more than context-free expressivity, using balanced materials for both groups.[2] The prediction anchored in formal language theory is as follows: human languages for which context-free expressivity is insufficient [1, 52, 58, 22] will impose greater computational resource demands during language processing than those languages for which context-free expressivity suffices. These predictions, however, were not met in empirical testing [2] of speakers of Dutch and German. Other work [76] attempted to reconcile the divergence of conclusions reached through asymptotic worst-case complexity analysis and empirically determined facts by making reference to the existence of operators that provide computational modules which operate outside the constructs that are adopted as theoretical primitives in classical formal language theory [39].

A quintessential quality of the expressivity enabled by context-free languages is the capacity to constrain arbitrarily deep nested dependencies – center-embedded recursion. Nested matching of brackets in programming languages and palindromes (strings concatenated with their reversals) are examples of recursive center embedding that may be regarded as canonical: see (1.1).[3] One sentence in this language

---

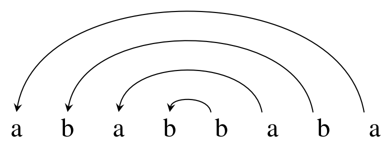
[1] For context-free languages, this is $O(n^3)$, and for the mildly context-sensitive languages, this is $O(n^6)$, where n denotes the terminal symbol token count within the sentence being processed.

[2] *Strong* containment in a fully abstract language family amounts to parse-trees for grammatical sentences to be appropriate to that language family; weak membership corresponds to grammatical sentences (strings) being appropriate. A language like Dutch may be weakly context-free at the same time as strongly context-sensitive, which is exactly the argument made by [12] about Dutch.

[3] Let the operator r denote the application of reversal to the string it superscripts. The symbol w stands for any arbitrarily ordered finite sequence of terminal symbols; here, this means sequences containing a and b finitely many times and in any order.

is depicted in (1.3): the string is $ababbaba$, where $w = abab$ and $w^r = baba$; the arcs demonstrate the nested dependencies. The string is shown again using indexed bracketing in (1.4). A set of production rules that satisfy the constraints of context-free grammars and which generate this language is illustrated in (1.5) – uppercase letters are non-terminal symbols, lowercase letters are terminal symbols and ϵ denotes the empty symbol. Greater than context-free expressive capacity is required to capture cross-serial dependencies. The string copy language (1.6) is an iconic representative of cross-serial dependencies. A sentence from this language, $babbbabb$ (where $w = babb$) is shown in (1.7) with an indexed bracketed equivalent provided in (1.8).[4]

$$\{ww^r | w \in \{a, b\}^*\} \tag{1.2}$$

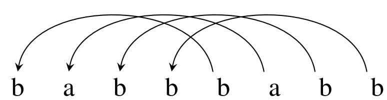


(1.3)

$$({}_l a({}_k b({}_j a({}_i bb)_i a)_j b)_k a)_l \tag{1.4}$$

$$\{S \rightarrow bSb, S \rightarrow aSa, S \rightarrow \epsilon\} \tag{1.5}$$

$$\{ww | w \in \{a, b\}^*\} \tag{1.6}$$

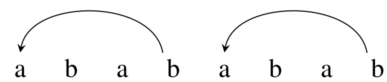


(1.7)

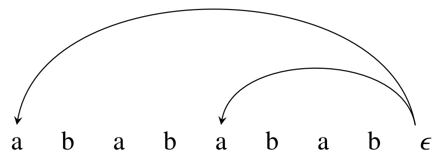
$$({}_l b({}_k a({}_j b({}_i bb)_l a)_k b)_j b)_i \tag{1.8}$$

$$\{w^i | w = (abab)^*\} \tag{1.9}$$

[4] The appendix provides a mildly context free grammar that generates the copy language and demonstrates a corresponding tree for the sentence in (1.7).



(1.9a)



(1.9b)

$$({}_i abab)_i ({}_j abab)_j \tag{1.10}$$

$$\{S \rightarrow ababS, S \rightarrow \epsilon\} \tag{1.11}$$

It is helpful to contrast the cross-serial dependencies and nested dependencies with a superficially similar but less complex construction that arises from tail recursion. Consider the example in (1.9): the ability to recognize strings consisting of i copies of w appears more complicated than a single copy, but in this example, w is fixed (to $abab$) while in (1.6) w consists of an arbitrary sequence of instances of a and b. Thus, the dependencies are extremely local, as depicted in (1.9, 1.9a) and (1.10),[5] or at best, bound to one edge as depicted in (1.9, 1.9b). A set of productions satisfying the constraints of regular grammars is provided in (1.11).

Cross-serial semantic dependencies can be demonstrated in English with "respectively" constructions, as in (1.12), which has the meaning given in (1.13),

Sue, Kim and Bobby went to medical school, law school and dentistry school, respectively. (1.12)

Sue went to medical school, Kim went to law school, and Bobby went to dentistry school. (1.13)

However, the asymptotic worst-case complexity results for the language recognition problem do not address semantic dependencies; those arise for the interpretation problem. Rather, the recognition problem is that of establishing that an input sentence satisfies syntactic and morphosyntactic constraints articulated by the grammar – it is the determination of grammaticality.[6] While reversal languages ww^r and

---

[5] Notice in (1.10) that the brackets do not nest or cross.

[6] An example of genuine cross-serial dependencies in natural language syntax is available in data from Swiss-German noted by Shieber [58]. A dialect spoken in Zurich allows arbitrarily many levels of embedding in subordinate clauses which can be idealized by the language $\{w|w = a^i b^j c d e^i f^j g, i, j \geq 0\}$. In the idealized string set specification, the instances of a, iterated i times, represent nominals with accusative case marking, and are in correspondence with instances of e, also iterated i times, and represent verbs that require accusative-marked nominal arguments.

copy languages *ww* are canonical examples of context-free and indexed languages, respectively, natural languages at either level of expressivity also have other constructions which do not require the full expressive power available. Nonetheless, in order to provide a model of grammaticality for those languages, it is necessary to accommodate their canonical constructions.

A conclusion drawn from the analysis of the empirical data described at the start of this section, which compared the error rates between native German and Dutch speakers in making judgements of sentences with increasing levels of subordinate clause embedding, is that the computation-theoretic equivalent of context-free grammars, push-down automata, do not supply the computational model that underlies human processing of natural language [2]. In response to these observations, it has been hypothesized that the human cognitive architecture has access to an operator, *expect a copy*, which is availed of in human processing of natural language [76]. In the context of formal language theory, this involves a *meta-grammatical* analysis. Recognition of repeating sequences is easily captured if the *expect a copy* operator is available. Recognizing that a series is repeated by a counterpart series is also an instance of the *expect a copy* process. Local complementation operators may be involved in order to match items with counterparts. A "local" complement is one that applies at the level of terminal symbols or pre-terminal symbols in a grammar, but not for arbitrary non-terminal symbols. Suppose that local complement of a is z and that the local complement of b is y, it follows that given the sequence $aabbb$, an expected complemented copy would be the sequence $zzyyy$. Addressing the natural language data associated with Swiss-German, the nominals with accusative case marking and verbs that require nominals marked with accusative case are local complements of each other, as are nominals with dative case marking and berbs that require nominals marked with dative case. Making local complementation operators available to *expect a copy* processes and invoking these as meta-grammatical processes gives both regular grammars and context-free grammars the capacity to generate and sentences that are beyond context-free syntactic expressivity.

The language processing facts that have been noted [2] are not easily explained with reference to fully abstract language families as defined by classical formal language theory [39]. Formal language theory suggests that processing constructions involving greater than context-free expressivity should be more difficult than processing constructions that require only context-free expressivity. The copy languages (1.6) require more than context-freeness while the reversal languages (1.1) require only context-freeness, so copies should be more difficult than reversals. Inspecting those canonical languages, though, suggests that copies (*ww*) *look* easier than rever-

---

The instances of b represent nominals with dative case marking, and the instances of f represent verbs that require dative-marked nominal arguments; the correspondence between b and f is established, again, through equal iteration, j times for each. If the number of nominals with dative case marking fails to correspond with the verbs that require dative marked nominals or if the number of nominals marked with accusative case does not correspond to the number of verbs that require nominals marked with accusative case, then the sentence fails to be grammatical [58]. Syntactic expressivity greater than supplied by context-free grammars is required in order to capture the data described here.

sals (ww^r).[7] The psycholinguistic data reported [2] also appears to suggest that the copies are easier than the reversals. To see this, contrast (1.12) above with (1.14) – many find it easier to derive the meaning (1.13) from (1.12) than to compute (1.15) from (1.14). However, (1.14) embodies nested-semantic dependencies and should, reasoning from formal language theory, be easier than (1.12).

Sue, Kim and Bobby went to medical school, law school and dentistry school, anti-respectively. (1.14)

Sue went to dentistry school, Kim went to law school, and Boby went to medical school. (1.15)

According to the meta-grammatical approach, a language like ww^r needs an operator additional to apply prior to the copy or complemented-copy operator – one that performs a reversal first – This additional process is then put forward as the explanation of the greater error rates during equal processing time for the greater error rate for the speakers of German (whose relevant constructions are homomorphic to ww^r) than of Dutch (whose relevant constructions are homomorphic to ww) [2, 76].

Within linguistic theory, the *expect a copy* function would be understood as meta-grammatical – something that operates orthogonally to the distinctions made available by the Chomsky hierarchy of expressivity. It effectively gerrymanders the Chomsky hierarchy to yield regions that may provide more appropriate characterization of the expressive requirements of natural language syntax.[8] Meta-grammatical processes other than *expect a copy* have been proposed in relation to other syntactic phenomena, such as extraposition and coordination resolution, and as such, indirectly support this analysis [23, 49]. Evidence used to support invoking a meta-grammatical analysis might come from prosody, the way humans modulate the sound of what they say in order to achieve grouping, for example. In general, the same sequence of words pronounced with two different intonation patterns may achieve two distinct syntactic analyses.[9] Inasmuch as intonation patterns are not typically understood as terminal symbols in a language, they may be characterized as meta-linguistic.[10]

It has been argued that symmetry identification is a general purpose process available throughout the human cognitive architecture, and that the use of the *ex-*

[7] It must be noted that this relationship is not evident in all empirical tests, however [32].

[8] The fully abstract families of languages that define discrete levels within the Chomsky hierarchy are characterized by closure under union, homomorphism, intersection with regular sets, and so on [48]. Given that the meta-grammatical operator when added to some abstract family of languages yields an irregular region, it may be expected that the resulting language family no longer respects the usual closure properties.

[9] The easiest way to see this is to recall that otherwise seemingly declarative sentences may be understood as interrogative if intoned thusly.

[10] Of course, it is possible to build models that lexicalize intonation patterns, or to associate finite sets of features with non-terminal symbols (functionally equivalent to having a finitely expanded inventory of non-terminal symbols) in order to encode prosodic constraints.

pect a copy function as a meta-grammatical process is an example of language processing availing of an operator that is available more widely in cognition [76]. We hypothesize that these specializations of symmetry identification, *expect a copy* and *expect a complemented copy* constitute diagonal faculties of the sort described above. We expect them to be available across domains functionally characterized – language, vision, emotion, haptic sensation, etc. We hypothesize that they involve shared resources (hence, competition for resources and little information encapsulation). These are empirical hypotheses with consequent empirical predictions. Apart from being able to explain the empirical facts that have been described, additional empirical predictions follow the availability of the operator. For instance, as noted [76], one would expect to find differential rates of "string-copy disfluencies" when comparing speakers whose native language can be meaningfully seen as homomorphic to ww with speakers whose native language is similarly homomorphic to ww^r. That is, where an operator is available, and one linguistic community makes more use of it than another, then one also expects to see differences between those communities in the degree to which that function goes awry.

We imagine that the function in its most general form interacts with behaviours outside language. For example, consider the *expect a copy* function as fundamental to habituation: humans sometimes repeat activities expecting repeated outcomes. Further, we imagine that the *expect a complemented copy* specialization of the function is fundamental to certain kinds of persistence: humans sometimes repeat activities hoping for a desired outcome distinct from what the activity last produced. Operators that promote habituation are fundamental to simulations that are used to learn about the possible past trajectory of human language evolution, and discovery of habituation behaviours is one way of conceptualizing the goal of machine learning applied to data originating from human cognition and interaction. We turn next to simulation as a method of gaining insight into the human cognitive architecture.

1.4 Simulation systems

Analysis of the evolutionary trajectories of species, including their cognitive architecture, is a useful complement to attending to the development of individuals, in theories of cognitive architectures appropriate to autonomously developing agents [68, 67], even if ontogeny does not replicate phylogeny. Within this, it is important to understand the evolution of language as a representational system useful for thought and communication. The human cognitive architecture could have much of the structure that it does but without it having a role in processing language. Humans would be humans even if they lacked language as a medium for thought and communication. Therefore, it seems incumbent to explain how language could have evolved, what functions were required of the human cognitive architecture in order to be able to process language, and how the distinctive properties of natural languages emerged. Mental states could be communicated more efficiently and effectively through smell or telepathy than natural language allows. Histories and desired futures would be easier and more complete if narrated via telepathy than can be imparted using natural language. Critical to the relative inadequacy of natural language for purposes of communication is that ambiguity in natural language is rife. Speaker meaning and hearer interpretation for the same sentence may differ. If natural language were optimized for communication in the way the eye is for vision, then one would imagine it would be as effective as telepathy. However, it is not. We feel that natural language is optimized for thought, and has a useful adaptation for communication (see [72, 73]). As a medium for conducting thought, natural language is very effective. It supports representation of *irrealis* states, negations, conditionals, and so on, that are more difficult to capture in a system that may record only what is positively present or positively remembered. Natural language is rather less good at supporting communication. It is very good for obfuscated communication – the capacity of natural languages for creating vagueness, ambiguity and deception are impressive, and allow speakers to hide true meanings from hearers. There is value in natural selection terms in being adept at obfuscated communication.

An account of language evolution must be able to provide a theory of the features of the first expressions of natural language and the contexts in which they emerged to account for both how those utterances were precipitated and what conditions were necessary for the effort of natural language use to be habituated. With respect to the effort of use and habituation, it is also necessary to be able to account for the speed of habituation, since something that requires significant effort without quickly obtaining effectiveness is more likely to be abandoned than something that reaches an effective state quickly. Considerations like these have shaped the language evolution workbench (LEW) for building simulations of agents and emergent languages and research availing of that system. The LEW is a platform that supports simulation of individual agents interacting as well as groups of agents interacting. Through the interactions, symbolic systems that approximate natural languages emerge, becoming available for communication [74, 75, 69, 3, 4, 5]. The LEW platform supports experimentation about early communication, using assumptions that appear pessimistic about the likelihood of interlocutors understanding utterances in the same way. Given that humans interacting with each other using language as it has evolved have little security that they have understood each other [64, 37, 70], it is reasonable to imagine that this condition has been endemic since the very first uses of language. The platform does not require the assumption that first communications succeeded in achieving positive transfer of information nor feedback about the level of success in their attempts at information sharing, even while providing an effectively infinite meaning space (cf. [41, 51, 44]).[11] The opposite assumption, that agents are capable

[11] In contrast, a landmark work in this area [41], assumes that the initial space of possible meanings is finite and small. Additionally, agents are able to observe both what other agents express and what they meant. This assumption means that the initial agents had direct independent access to the meanings denoted by expressions produced by other agents, and this is the essence of telepathy. Thus, their first uses of language require telepathy as an initial condition in the evolution of language. In discussing human language development, the process by which children acquire natural language in its current evolutionary state, similar assumptions are recorded [51, p. 29, emphasis in the source retained]: "A third possible source of input that I *will* exploit ... depends on the assumption that the child can infer the meaning of adults' utterances from their physical and discourse contexts

of determining the extent to which their attempts at communicating via language have been successful, presupposes that the language used to communicate was not necessary to the evaluation, and that effectively, they started off having access to telepathy (or something functionally equivalent). If perfect communication were available via other channels to begin with, then it is difficult to see why natural language would be adopted as a primary vehicle for clear communication.[12]

The LEW platform provides a range of parameters relevant to natural languages as used in thought and communication, and supports experimentation with alternative values for those parameters: the total agent population size, whether agents participate in groups and if so, alternatives for group dynamics based on individual dominance, the number of phonemes agents are capable of discriminating, the space of distinct types of events that may occur (and precipitate linguistic description), the extent of memory imperfection suffered by agents, the potential for feedback, *et cetera*. It is an assumption built into the LEW system that events happen with selections of agents witnessing those events, each according to their own perspective. That agents participate in shared events though potentially view those events very differently makes it even harder for an initial language to get off the ground than if language emerged bootstrapped by fully shared experience. The basic interaction involves an event happening, a speaker commenting on that event, and a hearer decoding that utterance.[13] When an event happens, an agent perceives that event as a set of components. It is an assumption of the LEW model that when a speaker produces an utterance, that expression contains a label for each event component that exists within that agent's perspective on the event that happened. The utterance is produced as a sequence of phonemes, segmented according to the speaker's individuation of labels. Another pessimistic assumption is that agents available to "hear" a speaker's utterance take that utterance in without the interference of noise – they observe the exact sequence of phonemes produced by the speaker[14], but that hearers segment the sequences of phonemes that speakers produce themselves, possibly in a manner that is distinct to that of the speaker. Thus, speakers and hearers may not agree on what the meaning-bearing units (groupings of phonemes) are in any utterance. Symmetric with the constraint that speakers assign to each component in their perspective on the witnessed event some utterance segment, for each meaning-bearing unit in the hearer's segmentation of the same phoneme stream, the hearer finds an interpretation for that utterance segment. The primary source of potential meanings is in the hearer's perspective on the shared event, but also the hearer may avail of memory and the past mappings that the agent has formed between sequences of phonemes and possible meanings. If the feedback parameter is enabled, then the interlocutors have the possibility of detecting whether they have been successful

---

and from the meanings of individual words in the sentence.... Thus the input to the acquisition mechanism might be a pair consisting of a sentence and its meaning."

[12] However, as noted above, there would remain value in being able to avail of a medium in which communication could be obfuscated.

[13] Modelling with only one agent acting as both speaker and hearer allows study of the emergence of language for thought, in isolation from communication.

[14] This assumption merits relaxation and re-configuring as an experimental parameter.

in communicating or not, but if the feedback is not successful, there is no further feedback regarding the aspects of the communication that were unsuccessful. This provides a model of aspects of task-based human dialogue in which participants may notice that their shared task is not succeeding, but not have direct evidence of which linguistic expressions are at fault. From outside the system, the researcher may monitor the simulation behaviour to externally assess the degree of communicative success that is managed. The simulation system provides a number of ways in which interlocutors may come to understand the meaning of utterances differently. Therefore, it is an interesting outcome for stable communication systems to emerge from these pessimistic assumptions.

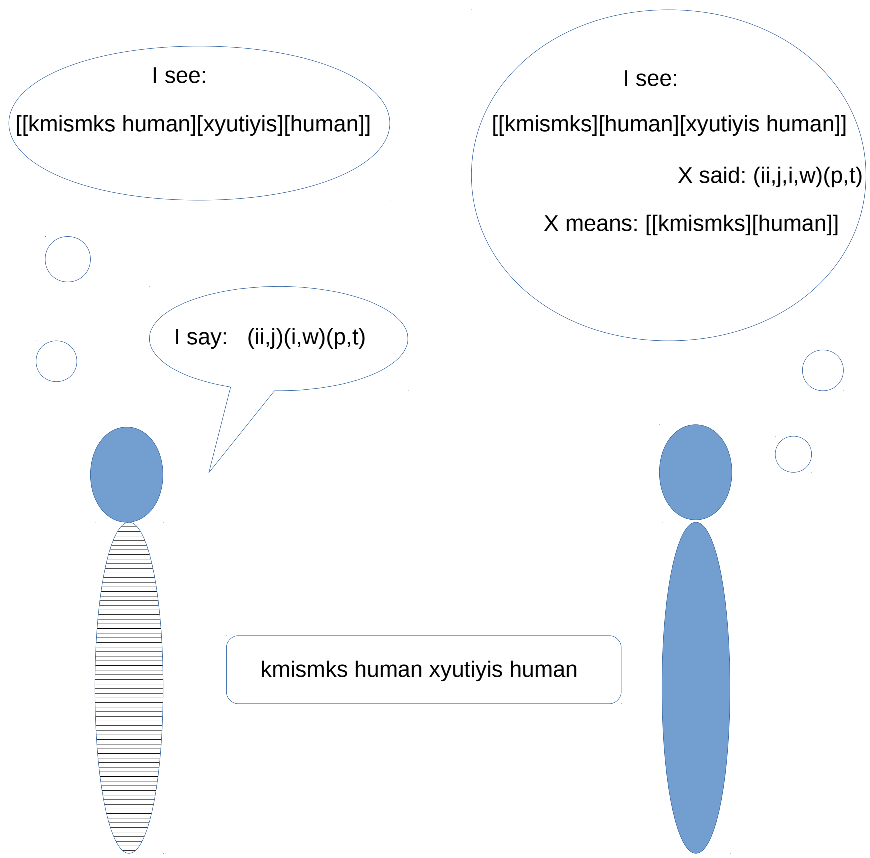


Fig. 1.1 A simulation of two agents inventing a linguistic system in the context of an event. For the purposes of the figure, the agent on the left has a name, X, but agent naming is not hard-coded into the system: "I see", "X said", etc. are all meta-level labels that characterize the role of the other objects.

An instance of the interaction scenario just described is illustrated in Figure 1.1. The scenario depicts an event that has been generated; within the model this takes the form of a sequence of atomic symbols, as in (1.16).[15]

$$\texttt{kmismks human xyutiyis human}. \tag{1.16}$$

$$\texttt{[kmismks human][xyutiyis][human]} \tag{1.16a}$$
$$\texttt{[kmismks][human][xyutiyis][human]} \tag{1.16b}$$

Models of events of this form are each finite but have arbitrary complexity because of the possibility for instances of event types being embedded within instances of event types.[16] The capacity for complex embedding events provides for an arbitrarily large space of possible meanings. As noted above, agents in the model witness exactly the same event, but potentially with distinct perceptions of what has happened. Alternative bracketing of sub-sequences of event atoms are used to model distinct perspectives; for example, (1.161.16a) and (1.161.16b) represent distinct perspectives on the event of (1.16). Speakers utter phoneme sequences using one sub-sequence for each component that they individuate within their construal of the shared event, as in the overall utterance and individual segment-meaning pairs of (1.17) in the example depicted in Figure 1.1.[17]

$$\text{“}\texttt{(ii, j)(i,w),(p,t)}\text{''} \tag{1.17}$$

$$[\![\texttt{(ii,j)}]\!] = \texttt{[kmismks human]} \tag{1.17a}$$
$$[\![\texttt{(i,w)}]\!] = \texttt{[xyutiyis]} \tag{1.17b}$$
$$[\![\texttt{(p,t)}]\!] = \texttt{[human]} \tag{1.17c}$$

$$\text{“}\texttt{(ii,j,i,w),(p,t)}\text{''} \tag{1.18}$$

$$[\![\texttt{(ii,j,i,w)}]\!] = \texttt{[kmismks]} \tag{1.18a}$$
$$[\![\texttt{(p,t)}]\!] = \texttt{[human]} \tag{1.18b}$$

[15] The implementation environment for the LEW platform is SICStus Prolog, so these are Prolog atoms.

[16] An example of such described in English might involve a seeing of a helping of someone to do something.

[17] The demonstrations in (1.17) and (1.18 use a symbol from denotational semantics ($[\![.]\!]$) – the goal of natural language semantics is to identify the right compositional specification of this function, into which one plugs arbitrarily complex expressions and is delivered their meanings. Here, it is used to demonstrate meanings of components as understood by speaker (1.17) and hearer (1.18).

Since this is a model of *ab initio* language invention, speakers have no choice at the outset but to innovate (potentially then re-usable) mappings between sequences of phonemes and denotations. The system encodes a propensity to re-deploy mappings that have been used previously over innovating new mappings at every stage. This propensity can be analyzed as containing an ability to identify copies and re-construct copies, and to that extent embodies a sensitivity to symmetry (see Section 1.3). Hearers are able to perceive the stream of phonemes produced by the speaker, but may group those phonemes into meaning bearing units into segments that are distinct to those presumed by the speaker, as in (1.18), corresponding to the example from Figure 1.1. Hearers interpret the utterance according to their segmentation and according to the hearer's perception of the event, finding a denotation for each segment perceived within the utterance. Hearers are also inclined to copy prior mappings between sequences of phoneme and their meanings, but may also create novel mappings as they interpret the utterances, innovating during utterance interpretation, in symmetry with speakers innovating during utterance production. As indicated before, as determined by the values used for the various simulation input parameters, the interlocutors may be able to determine whether they have been successful in communicating, but do not learn the locus of miscommunication or success in any of their utterances. Exploring models provided by configuration of the LEW with distinctive input parameters, even where the starting point agent pool is large, it seems that a shared "copy imperative" is adequate in enabling relatively speedy emergence of simulated languages that an external observer can determine to be successful in the sense that the agents may be seen to have significant levels of success in their communications. Even when extremely pessimistic assumptions are adopted, simulations using the LEW can achieve interesting levels of understanding and systematicity in the resulting linguistic systems.

Results of simulations show that adding more structure to the world, essentially by varying the likelihood of different event types being instantiated, using a kind of Zipfian distribution to determine the likelihoods, so that some event types are twice as likely to occur than other event types, and so on, leads to the emergence of communication systems that are stable with relatively greater speed than otherwise, and these simulations exhibit greater levels of agreement on mappings between phoneme sequences and meanings than when each event type is equally as likely as every other event type [75]. Similarly, when speakers and hearers have a tendency agree on the segmentation of phonemes in utterances, greater levels of agreement on the meanings of phoneme sequences may be noticed [69]. In simulations in which agents participate in groups, then greater levels of coordination of meaning tends to be visible in small groups than when the same number of agents communicate with each other without internal sub-communities [3].[18] When agents are provided limited feedback on their level of communicative success, as described above, then more coordination of meaning can be externally observed than in simulations where

[18] This outcome is consistent with "the Data Processing Theorem" according to which no information processing system is able to offer as output more information than is available in the signal provided as input to the system. When reasoning with the data processing theorem, one would predict information loss to correlate positively with group size. See [25, 35].

the agents do not receive feedback [4]. In simulations with both groups and limited feedback, then high levels of meaning coordination tend become established more quickly than otherwise. [5]. The value of the LEW and systems like it is in making it possible, through simulation, to provide a testing ground for theories about what parameters, what settings of those parameters, and what combinations of parameters and their settings are relevant to the to the emergence of systems like natural language. To be "like natural language" is to share coding properties, like synonymy avoidance [17], or to share properties of use, for example the manner in which communications conventions are observed to emerge among humans [36] and how features of the overall system depend on the degree to which each agent participates [38].

1.5 Related Work

The ideas that we have presented here from the perspective of understanding human cognitive architecture are compatible with the approaches taken by those who present their work as driven by the application-end of cognitive architectures, such as those who seek to develop effective robots. For example, research has addressed the ontology of interaction such as underpins robotic supports in ambient assisted living [34]. In attempting to build effective mobile robots, researchers have depended on reflection on possible cognitive architectures, sometimes identifying a need for models best described as hybrids of those available [19]. Some have addressed the fact that where robots or other technological advances become part of human existence, then this results in an entanglement [8, 77], and this, in turn entails that the "human cognitive architecture" will always be an evolving construct, as addressed by [72, 73].[19] The notion of "mathability" has been explored as another function that is a likely component of the human cognitive architecture, studied from theoretical, empirical and application oriented perspectives [10, 65, 63, 18]. In the domain of virtual reality, closely aligned with robotics, researchers explore rotation as a cognitive construct – this depends on three dimensional copy recognition, as opposed to sequence copy recognition; researchers [20, 21] seek to provide assessment of individual differences in ability. In any case, the point of this section is not to provide a comprehensive review of related work, but to demonstrate that among those immersed in robotics and closely related fields, other research is visible which we see as sympathetic to the approach we have described here.

[19] The quotation marks there, scare quotes, are meant to acknowledge that the human cognitive architecture may assimilate modules that arise from technology, and one might be averse, at first encounter, to thinking of the human cognitive architecture having components due to external artifacts, even human created artifacts.

1.6 Conclusions

In section 1.3 we offered an example approach to the advancement and validation of human cognitive architecture models. The idea is to individuate robust processes as possible cognitive faculties and to inspect whether the process has a role in a number of domains of cognitive processing, even if that role is more or less prominent, depending on the domain. This is a means of faculty individuation that is aligned directly with neither the horizontal faculties nor the vertical faculties as characterized by Fodor, as we recall from Section 1.2, where described them as diagonals. The approach begins with a focus on the role of the proposed process within one domain and examines the extent of evident availability to other domains. We demonstrated this with the *expect a copy* function. In section 1.4 we noted that simulation systems that depend on abstract embodiment models can be used to model agents and their interactions to explore what emerges and how the emergent systems relate to theoretical predictions. In our focus on abstract embodiment, we do not denigrate approaches that argue the importance of physical embodiment [59, 60, 61, 62], but rather we emphasize our view that the most fundamental cognitive activity is abstraction. Our discussion of simulations had a focus upon the emergence of language, but simulations that make it possible to model aspects of human interaction and cognition, more generally, for example non-linguistic behavioural interaction and emotion, also have enormous potential for increasing understanding of the parameters that influence those aspects of human existence. In one approach we suggest exploring the role specific processes that may be re-used throughout the cognitive architecture and in the other we suggest the use of simulation systems, also attuned to those processes. We have paid particular attention to symmetry sensitivity. We suggested that the *expect a copy* and its specialization as *expect a complemented copy* are fundamental to habituation and persistence, respectively. One could speculate that the expectation of a complemented copy is the basis of hope, as an emotion. The idea is that the nature of hope is the eager anticipation of a situation that complements some situation closer to hand, and thereby amounts to an expectation of a complemented copy. In support of this hypothesis, one might note that evidence suggests that specific neural substrate is implicated by optimism [57], and in the contrapositive direction, also a neural basis for bias against change of state in the case of depression [24].[20] We do not think that the copy imperative, even if it is rooted in neural substrate such as mirror neurons is the explanation of all human behaviours and attitudes, but rather that it underlies a human ability that is not completely mapped in the current state of knowledge at present. We have proposed here a method for exploring the role of this function and comparable functions that may be proposed as components of the human cognitive architecture. The analysis of robust process models and use of simulations are both useful methods for increasing understanding about human cognitive architecture.

---

[20] The contrapositive suggestion is that where there is a persisting absence of hope, in the form of depression, there is no invocation of an *expect a complemented copy* operation, perhaps due to an unavailability of that function.

References

1. Aho, A.V.: Indexed grammars—an extension to context-free grammars. Journal of the Association for Computing Machinery **15**(4), 647–671 (1968)
2. Bach, E., Brown, C., Marslen-Wilson, W.: Cross and nested dependencies in german and dutch: A psycholinguistic study. Language and Cognitive Processes **1**(4), 249–262 (1986)
3. Bachwerk, M., Vogel, C.: Modelling social structures and hierarchies in language evolution. In: M. Bramer, M. Petridis, A. Hopgood (eds.) Research and Development in Intelligent Systems XXVII, pp. 49–62. London: Springer (2010)
4. Bachwerk, M., Vogel, C.: Establishing linguistic conventions in task-oriented primeval dialogue. In: A. Esposito, A. Vinciarelli, K. Vicsi, C. Pelauchaud, A. Nijholt (eds.) Analysis of Verbal and Nonverbal Communication and Enactment: The Processing Issues, pp. 48–55. Heidelberg: Springer (2011)
5. Bachwerk, M., Vogel, C.: Language and friendships: A co-evolution model of social and linguistic conventions. In: T.C. Scott-Phillips, M. Tamariz, E.A. Cartmill, J.R. Hurford (eds.) The Evolution of Language, pp. 34–41. Singapore: World Scientific (2012). Proceedings of the 9th International Conference (EvoLang9)
6. Baranyi, P., Csapo, A.: Cognitive infocommunications: Coginfocom. In: 2010 11th International Symposium on Computational Intelligence and Informatics (CINTI), pp. 141–146 (2010). DOI 10.1109/CINTI.2010.5672257
7. Baranyi, P., Csapo, A.: Definition and synergies of cognitive infocommunications. Acta Polytechnica Hungarica **9**(1), 67–83 (2012). URL https://uni-obuda.hu/journal/Baranyi_Csapo_33.pdf
8. Baranyi, P., Csapo, A.: Revisiting the concept of generation ce - generation of cognitive entities. In: 2015 6th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), pp. 583–586 (2015). DOI 10.1109/CogInfoCom.2015.7390659
9. Baranyi, P., Csapo, A., Varlaki, P.: An overview of research trends in CogInfoCom. In: IEEE 18th International Conference on Intelligent Engineering Systems, INES, pp. 181–186 (2014). DOI 10.1109/ines.2014.6909365
10. Baranyi, P., Gilányi, A.: Mathability: Emulating and enhancing human mathematical capabilities. In: IEEE 4th International Conference on Cognitive Infocommunications (CogInfoCom), pp. 555–558 (2013). DOI 10.1109/CogInfoCom.2013.6719309
11. Bourbakis, N., Esposito, A., Kvaraki, D.: Extracting and associating meta-features for understanding people's emotional behaviour: Face and speech. Cognitive Computation **3**(3), 436–448 (2011). DOI 10.1007/s12559-010-9072-1
12. Bresnan, J., Kaplan, R., Peters, S., Zaenen, A.: Cross-serial dependencies in dutch. Linguistic Inquiry **13**(4), 613–35 (1982)
13. Brooks, R.A.: How to build complete creatures rather than isolated cognitive simulators. Architectures for intelligence pp. 225–239 (1991)
14. Bryll, R., Quek, F., Esposito, A.: Automatic hand hold detection in natural conversation. In: IEEE Workshop on Cues in Communication, pp. 1–6 (2001)
15. Cangelosi, A.: Evolution of communication and language using signals, symbols, and words. IEEE Transactions on Evolutionary Computation **5**(2), 93–101 (2001)
16. Cangelosi, A., Parisi, D. (eds.): Simulating the Evolution of Language. London: Springer (2002)
17. Carstairs-McCarthy, A.: Synonymy avoidance, phonology and the origin of syntax. In: J.R. Hurford, M. Studdert-Kennedy, C. Knight (eds.) Approaches to the evolution of language. Social and cognitive bases, chap. 17, pp. 279–296. Cambridge University Press (1998)
18. Chmielewska, K., Gilányi, A., Łukasiewicz, A.: Mathability and mathematical cognition. In: 2016 7th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), pp. 245–250 (2016). DOI 10.1109/CogInfoCom.2016.7804556
19. Csaba, S.: Symbolic cognitive abilities implementation on the ni-9631 pervasive mobile robot. In: 2015 6th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), pp. 11–15 (2015). DOI 10.1109/CogInfoCom.2015.7390555

20. Csapó, Á.B., Sudár, A., Berki, B., Gergely, B., Berényi, B., Czabán, C.: Measuring virtual rotation skills in maxwhere. In: 2020 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), pp. 587–590 (2020). DOI 10.1109/CogInfoCom50765.2020.9237889
21. Csapó, Á.B., Sudár, A., Berki, B., Gergely, B., Berényi, B., Czabán, C.: A test space for virtual rotation measurement in maxwhere. In: 2020 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), pp. 585–586 (2020). DOI 10.1109/CogInfoCom50765.2020.9237865
22. Culy, C.: The complexity of the vocabulary of Bambara. Linguistics and Philosophy **8**(3), 345–51 (1985)
23. Dahl, V., McCord, M.: Treating coordination in logic grammars. American Journal of Computational Linguistics **9**(2), 69–91 (1983)
24. Elliott, R., Sahakian, B.J.: The neural basis of mood-congruent processing biases in depression. Archives of General Psychiatry **59**, 597–604 (2002)
25. Esposito, A.: The amount of information on emotional states conveyed by the verbal and nonverbal channels: Some perceptual data. In: Y. Stylianou, M. Faundez-Zanuy, A. Esposito (eds.) Progress in Nonlinear Speech Processing, LNCS 4391, pp. 249–268. Springer-Verlag: Berlin (2007). DOI 10.1007/978-3-540-71505-4_13
26. Esposito, A.: The perceptual and cognitive role of visual and auditory channels in conveying emotional information. Cognitive Computation **1**(3), 268–273 (2009). DOI 10.1007/s12559-009-9017-8
27. Esposito, A., Campbell, N., Vogel, C., Hussain, A., Nijholt, A. (eds.): Development of Multimodal Interfaces: Active Listening and Synchrony. Berlin: Springer, LNCS 5967 (2010)
28. Esposito, A., Esposito, A.M., Vogel, C.: Needs and challenges in human computer interaction for processing social emotional information. Pattern Recognition Letters **66**, 41–51 (2015). DOI http://dx.doi.org/10.1016/j.patrec.2015.02.013
29. Esposito, A., Jain, L.C.: Modeling social signals and contexts in robotic socially believable behaving systems. In: A. Esposito, L.C. Jain (eds.) Toward Robotic Socially Believable Behaving Systems Volume II - "Modeling Social Signals", pp. 5–13. Cham, Switzerland: Springer (2016). DOI 10.1007/978-3-319-31053-4_2
30. Esposito, A., Marinaro, M.: What pauses can tell us about speech and gesture partnership. In: A. Esposito, E. Keller, M. Marinaro, M. Bratanic (eds.) The Fundamentals of Verbal and Non-Verbal Communication and the Biometrical Issue, pp. 45–57. Amsterdam: IOS Press (2007). URL http://ebooks.iospress.nl/volumearticle/8732
31. Esposito, A., Riviello, M.T., Bourbakis, N.: Cultural specific effects on the recognition of basic emotions: A study on Italian subjects. In: A. Holzinger, K. Miesenberger (eds.) HCI and Usability for e-Inclusion: 5th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society, USAB 2009, Linz, Austria, November 9-10, 2009 Proceedings, pp. 135–148. Springer Berlin Heidelberg (2009). DOI 10.1007/978-3-642-10308-7_9
32. Fitch, W.T., Friederici, A.D.: Artificial grammar learning meets formal language theory: an overview. Philosophical Transactions of the Royal Society of London B: Biological Sciences **367**(1598), 1933–1955 (2012). DOI 10.1098/rstb.2012.0103. URL http://rstb.royalsocietypublishing.org/content/367/1598/ 1933
33. Fodor, J.A.: The Modularity of Mind. Cambridge: MIT Press (1983)
34. Fülöp, I.M., Csapó, Á., Baranyi, P.: Construction of a coginfocom ontology. In: 2013 IEEE 4th International Conference on Cognitive Infocommunications (CogInfoCom), pp. 811–816 (2013)
35. Gallager, R.G.: Information Theory and Reliable Communication. New York: John Wiley & Son (1968)
36. Garrod, S., Doherty, G.: Conversation, co-ordination and convention: an empirical investigation of how groups establish linguistic conventions. Cognition **53**, 181–215 (1994)
37. Healey, P.: Communication as a special case of misunderstanding: Semantic coordination in dialogue. Ph.D. thesis, Centre for Cognitive Science, University of Edinburgh (1995). URL https://era.ed.ac.uk/handle/1842/20559

38. Healey, P.G.T., Mills, G.J.: Participation, precedence and co-ordination in dialogue. In: Proceedings of the 28th Annual Conference of the Cognitive Science Society, pp. 1470–1475 (2006)
39. Hopcroft, J.E., Ullman, J.D.: Introduction to Automata Theory, Languages, and Computation. Addison-Wesley Publishing Co., Reading MA. (1979)
40. Hurford, J.: Biological evolution of the saussurean sign as a component of the language acquisition device. Lingua **77**, 187–222 (1989). DOI 10.1016/0024-3841(89)90015-6
41. Hurford, J.: The evolution of the critical period for language acquisition. Cognition **40**, 159–201 (1991)
42. Hurford, J.R., Kirby, S.: Co-evolution of language-size and the critical period. In: D. Birdsong (ed.) New Perspectives on the Critical Period Hypothesis and Second Language Acquisition, pp. 39–64. Mahwah, NJ: Lawrence Erlbaum (1998)
43. Kaplan, R., Bresnan, J.: Lexical-functional grammar: A formal system for grammatical representation. In: J. Bresnan (ed.) The Mental Representation of Grammatical Relations, chap. 4, pp. 173–281. MIT Press, Cambridge, MA. (1982)
44. Kirby, S.: Syntax without natural selection: How compositionality emerges from vocabulary in a population of learners. In: C. Knight, M. Studdert-Kennedy, J. Hurford (eds.) The Evolutionary Emergence of Language, pp. 303–323. Cambridge University Press (2000)
45. Krichmar, J.L.: Design principles for biologically inspired cognitive robotics. Biologically Inspired Cognitive Architectures **1**, 73–81 (2012)
46. Leshinskaya, A., Caramazza, A.: Abstract categories of functions in anterior parietal lobe. Neuropsychologia **76**, 27–40 (2015)
47. Leshinskaya, A., Caramazza, A.: For a cognitive neuroscience of concepts: Moving beyond the grounding issue. Psychonomic Bulletin & Review **23**(4), 991–1001 (2016)
48. Lewis, H., Papadimitriou, C.: Elements of the Theory of Computation. New Jersey: Prentice Hall (1981)
49. Milward, D.: Non-constituent coordination: Theory and practice. In: Proceedings of the 15th Annual Conference on Computational Linguistics, pp. 935–41 (1994)
50. Parisi, D., Cangelosi, A.: A unified simulation scenario for language development, evolution and historical change. In: A. Cangelosi, D. Parisi (eds.) Simulating the Evolution of Language, pp. 255–275. Springer (2002)
51. Pinker, S.: Language Learnability and Language Development. Harvard University Press (1996 (1984))
52. Pullum, G., Gazdar, G.: Natural languages and context-free languages. Linguistics and Philosophy **4**, 471–504 (1982)
53. Reitter, D., Keller, F., Moore, J.: Computational modeling of structural priming in dialogue. In: Proceedings of the Human Language Technology Conference of the North American Chapter of the ACL, pp. 121–124. Association for Computational Linguistics (2006)
54. Reitter, D., Moore, J.: Predicting success in dialogue. In: Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics, pp. 808–815. Association for Computational Linguistics (2007)
55. Riviello, M.T., Esposito, A., Vicsi, K.: A cross-cultural study on the perception of emotions: How Hungarian subjects evaluate American and Italian emotional expressions. In: A. Esposito, A.M. Esposito, A. Vinciarelli, R. Hoffmann, V.C. Müller (eds.) Cognitive Behavioural Systems, pp. 424–448. Springer, LNCS 7403 (2012). DOI 10.1007/978-3-642-34584-5_38
56. Scherer, K., Banse, R., Wallbott, H.: Emotion inferences from vocal expression correlate across languages and cultures. Journal of Cross-Cultural Psychology **32**(1), 76–92 (2001)
57. Sharot, T., Riccardi, A.M., Raio, C.M., Phelps, E.A.: Neural mechanisms mediating optimism bias. Nature **450**, 102–105 (2007). DOI 10.1038/nature06280
58. Shieber, S.: Evidence against the context-freeness of natural language. Linguistics and Philosophy **8**(3), 333–43 (1985)
59. Steels, L.: The spontaneous self-organization of an adaptive language. In: K. Furukawa, S. Muggleton, D. Michie (eds.) Machine Intelligence, vol. 15, pp. 205–226. Oxford University Press, Oxford (1996)

60. Steels, L.: Synthesizing the origins of language and meaning using co-evolution, self-organization and level formation. In: J. Hurford, M. Studdert-Kennedy, C. Knight (eds.) Evolution of Human Language, pp. 384–494. Cambridge University Press (1997)
61. Steels, L., Kaplan, F.: Bootstrapping grounded word semantics. In: T. Briscoe (ed.) Linguistic Evolution through Language Acquisition: Formal and Computational Models, pp. 53–73. Cambridge University Press (2000)
62. Steels, L., Kaplan, F., McIntyre, A., Looveren, J.V.: Crucial factors in the origins of word-meaning. In: A. Wray (ed.) The Transition to Language, chap. 12, pp. 252–271. Oxford University Press (2002)
63. Szi, B., Csapo, A.: An outline of some human factors contributing to mathability research. In: 2014 5th IEEE Conference on Cognitive Infocommunications (CogInfoCom), pp. 583–586 (2014). DOI 10.1109/CogInfoCom.2014.7020411
64. Taylor, T.J.: Mutual Misunderstanding: Scepticism and the Theorizing of Language and Interpretation. Duke University Press (1992). DOI 10.1215/9780822383000
65. Tørøk, M., Tóth, M.J., Szøllősi, A.: Foundations and perspectives of mathability in relation to the coginfocom domain. In: 2013 IEEE 4th International Conference on Cognitive Infocommunications (CogInfoCom), pp. 869–872 (2013). DOI 10.1109/CogInfoCom.2013.6719218
66. Tyler, C.W. (ed.): Human Symmetry Perception and Its Computational Analysis. VSP International Science Publishers (1996)
67. Vernon, D.: Reconciling constitutive and behavioral autonomy. the challenge of modelling development in enactive cognition. Intellectica **1**(65), 63–79 (2016)
68. Vernon, D., von Hofsten, C., Fadiga, L.: Desiderata for developmental cognitive architectures. Biologically Inspired Cognitive Architecures **18**, 116–127 (2016)
69. Vogel, C.: Group cohesion, cooperation and synchrony in a social model of language evolution. In: A. Esposito, N. Campbell, C. Vogel, A. Hussain, A. Nijholt (eds.) Development of Multimodal Interfaces: Active Listening and Synchrony, pp. 16–32. Springer, LNCS 5967 (2010)
70. Vogel, C.: Attribution of mutual understanding. Journal of Law & Policy **21**(2), 377–420 (2013). URL https://brooklynworks.brooklaw.edu/jlp/vol21/iss2/6
71. Vogel, C., Esposito, A.: Advancing and validating models of cognitive architecture. In: 10th IEEE Conference on Cognitive Infocommunications, pp. 61–66 (2019). DOI 10.1109/CogInfoCom47531.2019.9089951
72. Vogel, C., Esposito, A.: Linguistic and behavior interaction analysis within cognitive infocommunications. In: 10th IEEE Conference on Cognitive Infocommunications, pp. 47–52 (2019). DOI 10.1109/CogInfoCom47531.2019.9089904
73. Vogel, C., Esposito, A.: Interaction analysis and cognitive infocommunications. Infocommunications Journal **12**(1), 2–9 (2020). DOI 10.36244/ICJ.2020.1.1
74. Vogel, C., Woods, J.: Simulation of evolving linguistic communication among fallible communicators. In: J. Hurford, T. Fitch (eds.) Proceedings of the Fourth International Conference on the Evolution of Language, p. 116. Harvard University, Cambrige, MA (2002)
75. Vogel, C., Woods, J.: A platform for simulating language evolution. In: M. Bramer, F. Coenen, A. Tuson (eds.) Research and Development in Intelligent Systems XXIII, pp. 360–373. London: Springer (2006)
76. Vogel, C.M., Hahn, U., Branigan, H.: Cross-serial dependencies are not hard to process. In: Proceedings of the 16th International Conference on Computational Linguistics, pp. 157–62 (1996). COLING‘96, Copenhagen, Denmark.
77. Waldbuesser, P., Komlósi, L.: Empirical findings in cognitive entity management: A challenge in the digital era. In: 2015 6th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), pp. 433–437 (2015). DOI 10.1109/CogInfoCom.2015.7390632
78. Westphal-Fitch, G., Huber, L., Gómez, J.C., Fitch, W.T.: Production and perception rules underlying visual patterns: effects of symmetry and hierarchy. Philosophical Transactions of the Royal Society B: Biological Sciences **367**(1598), 2007–2022 (2012). DOI 10.1098/rstb.2012.0098. URL http://rstb.royalsocietypublishing.org/content/367/1598/ 2007

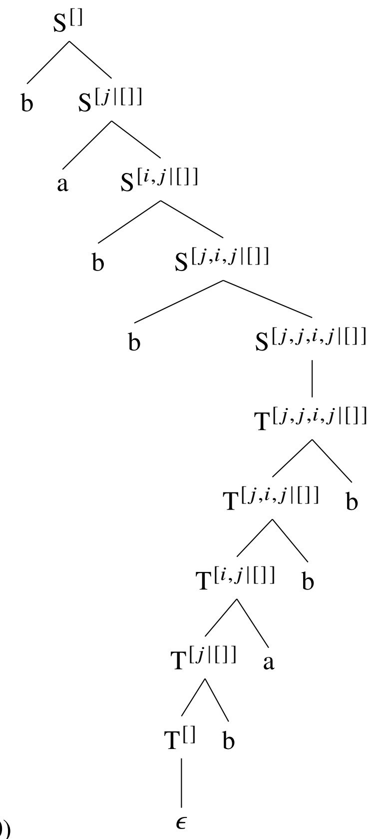
Acknowledgements This research is supported by the Science Foundation Ireland (12/CE/I2267 and 13/RC/2106) through the CNGL programme and the ADAPT Centre.

Appendix

A linear indexed grammar that generates the language of (1.6), $\{ww|w \in \{a, b\}^*\}$, is provided below. Assume a terminal alphabet as the set $\{a, b\}$, and non-terminals $\{S, T\}$, with $\{S\}$ as the set of possible re-write start or tree-root symbols. Let ϵ be the empty symbol. Production rules from a linear indexed grammar have the same constraints as those of context free grammars. Linear indexed grammars permit associating a stack of indices with a production rule's left-hand side non-terminal symbol and at most a single non-terminal symbol from the production's right-hand side. Allowing the stacks of indices to associate with non-terminals within productions entails that productions may embody the stack copy operation, pushing an index onto the stack, and popping an element off of the stack. The functional consequence of this expressivity is the ability to inter-constrain the number of times that productions are used within derivations, something that is not possible with properly regular grammars or context free grammars. Let [] represent the empty stack, and let $[H|R]$ represent a non-empty stack that has H as its first element and R as the rest of the stack. Let $I = \{i, j\}$ be a set of indices.

(1.19)Assume the following productions, letting M be a variable over stacks:

a. $S^M \rightarrow aS^{[i|M]}$
b. $S^M \rightarrow bS^{[j|M]}$
c. $S^M \rightarrow T^M$
d. $T^{[j|M]} \rightarrow T^M b$
e. $T^{[i|M]} \rightarrow T^M a$
f. $T^{[]} \rightarrow \epsilon$



(1.20)

Under those assumptions, the productions in (1.19) generate the copy language of (1.6), and a parse tree that results for the sentence of (1.7) using those productions is provided in (1.20).[21]

[21] A fully indexed grammar allows more than one non-terminal symbol on the right hand side of productions to participate in the stack sharing operations. From the tree that derives the sentence of interest ($babbbabb$) through the sequence of nodes at its terminal fringe, shown in (1.20), it may be seen that there is a single line through the tree with stack annotations; fully indexed grammars, in contrast to linear indexed grammars, would allow branching paths of non-terminal nodes annotated with stacks through trees that they license.