



# Experimental Semiotics: A Systematic Categorization of Experimental Studies on the Bootstrapping of Communication Systems

Angelo Delliponti<sup>1</sup> · Renato Raia<sup>2</sup> · Giulia Sanguedolce<sup>2</sup> · Adam Gutowski<sup>1</sup> · Michael Pleyer<sup>1</sup> · Marta Sibierska<sup>1</sup> · Marek Placiński<sup>1</sup> · Przemysław Żywczyński<sup>1</sup> · Sławomir Wacewicz<sup>1</sup>

Received: 19 October 2022 / Accepted: 1 May 2023  
© The Author(s) 2023

## Abstract

*Experimental Semiotics* (ES) is the study of novel forms of communication that communicators develop in laboratory tasks whose designs prevent them from using language. Thus, ES relates to pragmatics in a “pure,” radical sense, capturing the process of creating the relation between signs and their interpreters as biological, psychological, and social agents. Since such a creation of meaning-making from scratch is of central importance to language evolution research, ES has become the most prolific experimental approach in this field of research. In our paper, we report the results of a study on the scope of recent ES and evaluate the ways in which it is relevant to the study of language origins. We coded for multiple levels across 13 dimensions related to the properties of the emergent communication systems or properties of the study designs, such as type of goal (*coordination versus referential*), *modality* of communication, absence or presence of *turn-taking*, or the presence of *vertical vs. horizontal transmission*. We discuss our findings and our classification, focusing on the advantages and limitations of those trends in ES, and in particular their ecological validity in the context of bootstrapping communication and the evolution of language.

**Keywords** Linguistics · Experimental Semiotics · Language Origins · Communication Systems · Semiotic Game · Language Evolution

## Introduction

*Experimental Semiotics* (ES) “focuses on the experimental investigation of novel forms of human communication [...] which people develop when they cannot use pre-established communication systems” (Galantucci & Garrod, 2011, p. 1). A typical semiotic experiment is a game in which participants attempt to communicate without using conventional signs. As a familiar example, think of the popular party game of charades, where the actor (in ES, often technically referred to as the director) must convey without using language a movie title to an audience of guessers (matchers), using to that end non-linguistic resources such as gestures, facial expressions and other bodily movements. Of course, the repertoire of ES games is much broader, involving, for example, the use of drawings, abstract shapes or non-linguistic vocalizations.

A key advantage of ES is that it provides researchers with tools to investigate communication in very general terms, distinguishing the mechanisms of different communication systems (Galantucci et al., 2012a, b). This makes ES relevant for a range of fields in cognitive science and the study of communication and makes its results particularly valuable for several subfields of linguistic research. For example, this approach shows that different factors that operate during communication (such as the structure of meanings or biases for alignment between interlocutors) and during transmission (including population dynamics and constraints on learning) contribute to structures that emerge under different circumstances and shape patterns of variation in languages (Tamariz, 2017).

ES is a successor of an earlier tradition of laboratory studies on language acquisition and change. In a recent review, Nölle and Galantucci (2021) list several 20th century laboratory experiments on communication whose paradigms resemble those used by ES, albeit they have a different focus<sup>1</sup>. The other close intellectual cousin of ES is Experimental Pragmatics (see esp. Galantucci & Garrod, 2011). Experimental Pragmatics investigates how humans use pre-established forms of communication, such as spoken language in dialogic form, with a specific focus on pragmatic aspects of communication like relevance, degree of specificity, and structure. However, the main difference is that ES focuses on the emergence of novel forms of communication so it prevents participants from using natural language, while Experimental Pragmatics studies already existing forms of communication in various contextual situations.

Particularly interesting is the role that Experimental Semiotics has come to play in studies on the evolutionary origins and development of language. At its core, ES investigates the process of creating the relation between signs and their interpreters as biological, psychological, and social agents (see e.g., Morris, 1938 or the notion of “intermediary pragmatics” in Bar-On, 2021). The relationship between symbols and referents, which makes symbolic thought distinct from mere association-making

<sup>1</sup> These include the experiments done by Bartlett (1932) and Bavelas (1950, 1952), who studied how information transmission leads to the creation of ‘stereotypical’ memory patterns or the emergence of leadership roles, and the early attempts of using artificial miniature languages to study language change (e.g., Esper 1925; Wolfle, 1933).

processes, was seen as the hallmark of human reasoning and language, and essential for their phylogenetic development. It has been of interest since ancient times, with the key question being whether it was an innate and universal ability or the outcome of complex social processes. Several unverified stories speak to this interest in recreating the process of how a communication system is born: namely through children being raised in linguistic isolation without any human interaction (see e.g., Żywiczyński, 2018). Such a cruel procedure, known as the “forbidden experiment”, was allegedly performed by Psamtik I of Egypt, Frederick II (Hohenstaufen) of Sicily, and James IV of Scotland (Campbell & Grieve, 1981). The only conclusion to be drawn is that language is not entirely innate: if a child is completely deprived of any linguistic input, they will not speak any language (Galantucci, 2017). In a way, ES could be seen as a descendant of these stories about the “forbidden experiment” that is, however, following modern-day ethical standards. Its goals consist in studying how communication systems are brought into existence and how sets of conventional relations are created and then shaped through repeated use, which in turn helps to understand the processes that underlie the development of language in ontogeny, cultural-historical change and phylogeny (Galantucci, 2005).

Over the years, ES has grown into an extensive field and, as such, has become the object of several overviews (e.g., Galantucci et al., 2012a, b; Galantucci & Roberts, 2012; Galantucci 2017). In an early summary of ES research, Galantucci et al. (2012a, b) outlined the basic research problems tackled in the field, described the main study paradigms, and explained their implications for linguistics. They described three types of experimental paradigms: semiotic referential games, semiotic coordination games, and semiotic matching games. The authors also recognised five main research themes of ES: “the emergence of linguistic structure, the role of interaction in communication, the role of inter- and intragenerational processes in the evolution of language, the study of sociolinguistic processes in the laboratory, and the bootstrapping of communication” (Galantucci et al., 2012a, b, p. 581). To demonstrate the potential of ES as a major complement to linguistic research, Galantucci and colleague Galantucci et al. (2012a, b) specified three reasons: enabling the study of novel communication systems; providing full access to the history of their development; and the potential for easily controlling the conditions of this development. Another, similar overview of ES recognises three main themes: linguistic properties as the consequence of communication, social factors in communication, and the bootstrapping of communication (Galantucci & Roberts, 2012).

Importantly from the perspective of this paper, all the existing overviews of ES follow the traditional format of the review paper, without applying the tools of a systematic literature review (SLR). Although these existing overviews describe particular paradigms in considerable detail, their narratives and coverage of the literature are necessarily subjective and selective. In this paper, we propose a different, bottom-up, approach to characterizing ES, inspired by the systematic literature review approach (Xiao & Watson, 2019). Our main goal is to create a comprehensive resource of ES studies relevant to the earliest stages of establishing a communication system, which is categorised by a broad range of design parameters. In doing so we aim to create a resource that will inform future ES works, but also to understand how to conduct research and which paradigms have been ignored. Therefore, we conducted

a systematic review of 60 ES studies, published over the period of 20 years, from 2002 to 2021. We coded the papers for multiple levels across dimensions related to the properties of the emergent communication system and the properties of the study design. The results of the coding were subjected to statistical data analyses as categorical variables. Other dimensions were textual and reflected the more general and qualitative aspects of the papers, such as their main findings. All the coding was compiled into a single, interoperable and reusable dataset, as is described in detail below. Thanks to these efforts, we are able to provide a novel, systematic approach to characterize the properties of ES studies.

## Dataset

### Inclusion Criteria, Acquisition of Articles, Coding Procedure

#### Inclusion Criteria

Although a classic understanding of ES restricts its meaning to “controlled studies in which human adults develop novel communication systems” (Galantucci et al., 2012a, b, p. 477), this definition is occasionally extended onto controlled studies in which adults “impose novel structure on systems provided to them” (Galantucci et al., 2012a, b, p. 477). On this broader definition, ES also subsumes studies where sign-meaning pairings are already provided by the experimenters rather than emerging naturally in the game, as in most “alien language” studies using the iterated learning paradigm (e.g., Cuskley, 2019). In line with our interest in language origins, particularly the early bootstrapping phase of communication, we adopted this first – classic and narrower – definition of ES as an inclusion criterion. That is, we included studies on the emergence of novel communication systems and excluded studies in which participants began learning the meanings already assigned by fiat to a set of signs. For practical reasons, we limited articles to those that had been published in peer-reviewed journals, thus excluding experiments reported in chapters in edited volumes or proceedings papers.

#### Acquisition of Articles

Articles that matched the inclusion criteria were identified and acquired through a three-step procedure. First, an initial list of ES studies consistent with our criteria was compiled bottom-up. In the second step, the coders went through the references of the articles in the initial list, as well as references in review articles ((Galantucci et al., 2012a, b; Galantucci & Roberts, 2012; Galantucci, 2017; Galantucci & Garrod, 2011; Nölle & Galantucci, 2021) to identify articles containing further studies eligible for inclusion. Finally, the coders did a series of targeted searches on Google Scholar and Connected Papers for keywords such as “experimental semiotics,” “semiotic game,” or “laboratory languages,” in order to extend the search to all studies linked by similar topics. The completeness of the list created in steps one through three was later approved by a leading expert in ES external to the coding team.

## Coding Procedure

The coders were first trained to apply the coding dimensions to an initial set of two papers each. Their coding was then discussed and refined by all researchers. All papers on the final list were distributed among eight coders. First, they worked independently, each coding the assigned papers for the 13 dimensions described below (Sect. 2.2), and marking potentially difficult classificatory decisions. These were then resolved consensually through discussing such unclear cases in the coding in a group.

## Coding Dimensions

The papers were coded for three types of dimensions: (1) basic bibliographic and scientometric information (the year of publication, the total number of citations on Google Scholar as of April 22, 2022, as well as citations per year), which gave us an idea of the popularity of each paper in the field; (2) general information: the paper's main themes or topics, a brief summary of the main findings, the number of participants, their age range, and the experimental setting (laboratory or online); (3) study design properties, which were treated as categorical variables, coded as numerical values assigned to category labels. For example, the variable "type of game" had two values, "1" for referential games and "2" for coordination games. These values were then statistically analysed. The dimensions included in (3) are described in detail below. The coding dimensions are based on descriptions of ES paradigms in the literature as well as the key differences evident between the studies that can be related to overarching type differences.

### Type of Games: Referential vs. Coordination

Despite its recent origin, ES has developed two main paradigms: *referential games* and *coordination games*. The referential framework of ES is derived from standard referential communication tasks that were employed in Experimental Pragmatics (see e.g., Krauss & Weinheimer, 1966), in which participants had to converse about novel shapes using natural language. In the ES version, the use of natural language is forbidden, so players must communicate about a predetermined stimulus (e.g., a piece of music or a concept) using other means. In standard referential games, the set of signals used for communication is open, whereas the set of referents to communicate about is closed and pre-established by the experimenters (Galantucci et al., 2012a, b) (see Sect. 2.2.3). The purpose of the communicative act is communication itself; the goal of the director is to have the matcher correctly guess the intended meaning. A paradigmatic model of referential games is the "Pictionary" set-up employed by Garrod and colleagues (2007), in which the director has to graphically depict various concepts and communicate them to the matcher(s).

In coordination games, the communicative act is instrumental for the purpose of the game, which is succeeding in a specific task that usually involves moving an agent in a virtual space and coordinating the moves with the partner. In these games, successful communication can be supported by different sets of referents, therefore players must agree not only on the set of signals but also on the set of referents used

to make communication successful ((Galantucci et al., 2012a, b). One model of coordination games has been dubbed the “tacit communication game” (TCG; Galantucci 2017): in TCG, each player of a dyad controls one virtual agent (a geometric figure, the “token,” which can be moved and rotated) over a  $3 \times 3$  grid, and their goal is to place their tokens in the correct positions, established by the experimenters. Only one of the players, the sender, knows the correct position and has to communicate it to the other using only moves over the board. The moves of the sender thus serve a double function of, first, moving the player token into the correct position and, second, communicating to the other player their correct position. The sender has to find a way to clarify which moves have just an instrumental purpose and which have a communicative purpose.

### Vertical Transmission

In most ES studies, a communication task is performed within a dyad or a larger group of participants whose composition remains constant throughout the ES game. However, there is an interesting minority of studies with a dynamic group composition, such that some players leave, and others join the group within the timeframe of the game. Such a design enables the *vertical transmission* of information, which occurs when the communicative output of one generation (e.g., a set of signs they have converged on) becomes the input to which the next generation is exposed. One example are replacement microsociey studies (e.g., Caldwell & Smith, 2012), where the interacting group is composed of a director and a small number of receivers; at the end of each turn, the director is removed from the game and the most experienced matcher becomes the new director, while a new player enters the group as the least experienced matcher. These studies simulate a natural aspect of human society: the communicative conventions created at a given time are passed onto the next generations, which have to learn and inevitably modify them. Inserting the vertical transmission of established conventions into ES designs hence offers a way to study the cultural evolution of sign systems. However, studies that do not feature vertical transmission focus on the emergence of novel communication systems in the interaction of agents engaged in a particular activity, either reference or coordination (e.g., Galantucci et al., 2012a, b).

### Signals and Referents

Two dimensions in our coding scheme concern the type of signals adopted in each experiment. The first is about the kind of *medium* employed in communication. A large majority of experiments use either vocalizations, bodily-visual signals (i.e., communicative bodily movements, such as gesture, pantomime, facial expression or gaze), or graphical signals (drawing, symbols, lines, colours, etc.). There was also a small minority of studies whose medium of communication did not fall under any of these three possibilities (e.g., Iizuka et al., 2013).

The category *signal space* was further subdivided into discrete and continuous. In a *discrete signal space*, senders chose the signal from a set of specific, predefined possibilities, often limited in number, effectively making signal production a multiple

alternative forced choice task. An example of a discrete signal space is to choose an Arabic numeral from a set of 1 through 10 as the signal to be sent to the receiver. TCGs are games that usually employ a discrete signal space: the possible configurations of the tokens which the sender has to use for communicative purposes are inherently limited because the token can only move on a small grid (e.g., Blokpoel et al., 2012). Conversely, in a *continuous signal space*, senders could produce any signal form possible within the constraints of the communication medium; an example would be pen-and-paper (or digital) drawings, which are not limited to a number of distinct variants but instead can take on any shape. An interesting but much less frequent possibility is that the signal space is continuous but not unlimited; in this case, the director must choose within a spectrum of possibilities, for example, shades of colour (e.g., Roberts & Clark, 2020).

We also coded for what we dubbed the *meaning space* and identified the types of referents used for communication. The referents can be common concepts (objects or actions which are easily verbalizable, like “house,” “dog” or “giving a kiss”) or more abstract entities (unfamiliar geometric shapes, pieces of music, configurations). We decided to create another level for this category, which mostly applies to coordination games, that is when the referents are a particular position or disposition of the tokens. In Zlatev and colleagues (2017), we have an example of a study with a referential game in which referents made up of meaningful concepts are taken as meaning space. In this case, pantomime was used to express concepts such as a father kissing his daughter, a person hugging another person, etc. On the other hand, in Stevens and Roberts (2019), we have a coordination game, as the sender and the receiver had to coordinate in order to find the best way to communicate and interpret the expressed signs. The meaning space was composed of lines inside the cells; therefore, it was the position of the signs that was communicated. In this sense, we claim that the meaning space refers to a location.

## Interaction

We also examined the parameters of interactions between players and the general setup of the game. One of the categories we used for this was related to the *feedback*, which is information about the outcome of the communicative interaction process. As a simple example, if the director produces a clenched-fist gesture, to which the matcher responds “war,” in most studies this will be followed by feedback in the form of “correct” (if “war” was indeed the intended meaning) or “wrong” (if the intended meaning was something else). We were interested in the source of this information: in some experiments, feedback comes from the other player(s), in others from the experimenter themselves, and in still others there is no feedback. Sometimes (as in the Pictionary-like game in Fay et al., 2017), the presence of feedback is itself one of the studied variables, as its presence or absence can alter communicative success and other important properties of the exchange.

Related to feedback is the category of *turn-taking*, which describes the turn-order of the players’ actions. As one option, there could be no turn structure, with players being free to take their actions at any time and in any order, even simultaneously. However, there could also be a fixed turn structure governing the exchange of turns; a

frequent pattern is the director acting first by sending the signal, and then the matcher selecting a possible referent. Such a structure could either be pre-specified, that is, imposed by the mechanics of the experiment or may emerge spontaneously during the game, despite not being formally determined by the experimenters. A different category, *interchangeability*, captured whether the roles of directors and matchers were assigned to particular players for the entire duration of the game, or if players could change their roles. That is, if a player could be the director at one point of the game and then change their role to that of the matcher at another point, interchangeability was present. Conversely, if one player was always the director and the other was always the matcher, interchangeability was absent.

Two further categories are related to the interacting group. *Group size* was the total number of players in a group, whereas *communication type* referred to how many players took part in an individual interaction act. For example, if a study had groups consisting of seven players but each communicative act always happened between two players, *group size* and *communication type* would be classified respectively as “larger groups” and “dyadic communication.”

Finally, we examined whether the interaction between senders and receivers in the experimental setup was, or not, *simultaneous*. A simultaneous interaction is when the reception of the signal occurs immediately after its creation by the director, which is a characteristic of live interaction. If the matcher is looking at a stimulus recorded at an earlier time, the interaction is considered non-simultaneous.

## Alignment of Interest

The category of alignment of interest was introduced to study one of our greatest conceptual interests: whether the origins of language were marked by a competitive or cooperative use of our communicative means (e.g. Tomasello, 2008; Scott-Phillips, 2014; Ferretti, 2022). Language is traditionally believed to be born out of a cooperative attitude among humans: after all, if signals were used mostly for deceptive purposes, no one would have reason to trust them and language would become useless and disappear. Note that models of animal communication, inspired by Krebs and Dawkins (1984; also Dawkins & Krebs, 1978), mostly see communication as a means to influence and manipulate the behavior of others to one’s own advantage: the cooperative presupposition would, in fact, imply an evolutionarily unlikely altruism by the senders of signals or a similarly unlikely gullibility by receivers. Some current models of language function and evolution consider it to be characterized by a mixture of competitiveness and cooperativeness (Sperber et al., 2010; Lee & Pinker, 2010). ES is a particularly well-suited means for studying the emergence of the early properties of human communication, such as compositionality and combinatoriality, under the influence of humans’ pragmatic abilities. It would be interesting to know if this development can also occur when there are differing interests among the people involved in communication. An important distinction must be made here: one thing is competition *among* interacting groups (which is sometimes employed as an incentive for players; the group with higher communication success receives more points – sometimes associated with a monetary prize); another is competition *inside* each interacting group, that is, the existence of a conflict of interest between senders and



receivers. This latter kind of competition is the one we are interested in as it allows us to study the competitive or cooperative nature of early communication.

## Summary

This is a short summary of the dimensions included in the statistical analysis. The numbers associated with each value (which are followed by an explanation of those values) are the actual numbers that were subjected to a cluster analysis. In all cases, a further value “other” was added for papers that did not fit into any of the pre-established levels. In these cases, additional specifications were included (Table 1).

To follow up on our first example, according to these coding dimensions, a typical game of charades would be classified as referential (goal: to be understood, to convey a concept), involving no vertical transmission (or only marginally so, if successive players adopted some gestures and pantomimes used by the previous players, see e.g. Christiansen & Chater, 2022), a bodily-visual medium of communication, an open and continuous signal space (no predefined set of gestures – any bodily configuration can be used), a meaning space of meaningful concepts (such as movie titles), feedback that comes from the director, no turn taking (directors and matchers do not need to wait their turn, can send signals / provide responses at any time and in any order), the presence of interchangeability (people change roles of doing the pantomimes and guessing their meaning), group size and communication type that depend on the number of matchers in the audience; interaction that is simultaneous (unless the pantomimes are recorded and later shown to the matchers), and the interests that are aligned (the director wants the matchers to guess correctly, and so do the matchers).

## Applications

The database presented in the above sections is intended as a multipurpose resource with a broad spectrum of diverse applications in Experimental Semiotics research. Here, we limit ourselves to pointing to three avenues in which this resource can be put to use.

### Informing Reviews and Designs

Firstly, a basic application of the database is in informing literature reviews on the field – both those intended to provide theoretical overviews and those underlying experimental studies – to facilitate a more systematic and comprehensive coverage of the relevant literature. For example, researchers planning to address their research question through a *coordination game* design will be in a position to instantly identify and access an exhaustive set of previous studies using this particular paradigm. Furthermore, the proposed classifications may help scaffold new experiments in ways that facilitate rigor and productivity. Since the development of experimental designs involves many decisions that are usually taken implicitly, the dimensions used in our database may serve as a guide to reviewing such decisions in an informed manner. For example, planning the design of the said novel coordination game study involves

**Table 1** Coding dimensions used in the study

Dimension	Description	Values
Game type	If the goal of the game is to be understood (referential) or if communication is a means of achieving something else (coordination).	1. Referential 2. Coordination
Vertical Transmission	If the study incorporates replacing “generations” of players.	1. Vertical transmission is present 2. There is no vertical transmission
Medium of communication	The medium of the signals.	1. Vocalizations 2. Bodily-visual 3. Graphical
Signal Space	Properties of the set of usable signals as means of communication.	1. Limited and Discrete 2. Open and Continuous 3. Continuous but limited
Meaning Space	Type of meaning of the referents.	1. Meaningful concepts / words 2. Abstract shapes / symbols / configurations 3. Location
Feedback	Source from which participants receive information about the outcome of the game.	1. No feedback 2. Feedback comes from other participants 3. Feedback comes from the experimenters or emerges from the experimental setup itself 4. Feedback is provided by both the system and the participants 5. Other
Turn-taking	Does the study include roles of directors and matchers or not?	1. The roles are pre-established by the experimenters 2. The roles are absent in the study 3. The roles are not formally established by the experimenters, but they emerge spontaneously
Interchangeability	Do the players exchange their roles as directors and matchers?	1. Yes, players alternate between the roles 2. No, players remain in their roles for all the length of the experiment
Group size	Number of people in each interacting group.	1. Pairs 2. Small groups (three to five people) 3. Large groups (more than five people)
Communication type	Number of people involved in each interaction.	1. Dyadic 2. Triadic 3. Four-way
Simultaneity	Is the interaction simultaneous or is the matcher watching a recording?	1. Simultaneous 2. Non simultaneous
Alignment of interest	Is there a conflict of interest inside the interacting group’s members?	1. Common interests 2. Conflict of interest

deciding on the medium of the communication, community size, openness of the signal space, and so forth, which can be readily compared against such decisions in existing studies.

It is worth noting that these points generalise beyond the area of Experimental Semiotics, extending into studies on the origins of communication more broadly. This is particularly relevant to agent-based modelling, where building the model

itself involves taking explicit decisions on several dimensions, such as signal space, turn-taking, or alignment of interests (e.g. Zubek et al., 2023).

## Identifying Patterns in ES Research

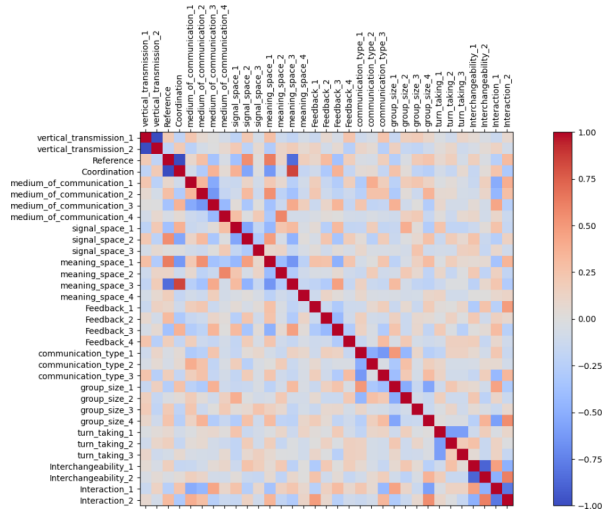
Another application with direct implications for research practice is searching the multidimensional space of possible design configurations to identify over- as well as underrepresented designs. These can be either choices in a single dimension or, more interestingly, choices along two or more dimensions that are highly correlated with one another (e.g. *game type: coordination* almost invariably involves *medium of representation: graphical*). By extension, this also allows us to point at alternative design configurations – i.e. ones that are possible in principle but not actually implemented in existing studies – thus showing us unexplored or underexplored possibilities.

As a simple example, consider the dimension *alignment of interests*. Almost all studies conform to the default “cooperative” setting of making the interests of the communicators aligned with each other: all parties of the communicative situation share the same goal of converging on the same referents or locations. There are only two exceptions (dos Santos et al., 2012; Inoue & Morita, 2021), which introduce some degree of conflict of interest (thus, rivalry) between the communicators, who are incentivised to pursue one’s own communicative goals even when this might be at the expense of their partners. This seems to be consistent with the theory: that it is difficult to imagine the bootstrapping of a communication system without cooperation. Of the two studies considered, at least in one case, competition had a positive effect on the consolidation of communication. However, this happened when the competition was on a global scale, and not on a local one: the result is that “humans change their level of cooperation as a function of the scale of competition (...), highlighting the importance of considering the scale of competition in studies of cooperation and communication” (dos Santos et al., 2012). Thus, a question could be posed on whether we need more studies with some kind of conflict of interest to investigate the possible role of competition in communication, which would be in line with some recent theoretical proposals on the role that persuasion may have played in the evolution of language (Ferretti & Adornetti, 2021).

To provide a more complex example, we conducted an analysis of correlations between our dimension values. To this end, the data frame was transformed in such a way that each factor level could be encoded as either 0 (= not present in the study) or 1 (= present in the study). For instance, the category “game type” was divided into “referential” and “coordination”, each of which were subsequently marked with 0 or 1 s. Several more technical variables, such as the year of publication or population age, were not included in the analysis; others were removed due to being a single-level variable (e.g. “alignment of interests”). Here, we present only a sample of our analysis with the strongest correlations between the encoded categories, and the whole data frame of correlations can be accessed under [https://osf.io/ad7b4/?view\\_only=0590ad2c505840dd8cceb1d8f890cb4](https://osf.io/ad7b4/?view_only=0590ad2c505840dd8cceb1d8f890cb4) (Fig. 1).

Our analysis suggests a strong correlation ( $r = 0.85$ ) between *meaning\_space\_3*, i.e. communicating about a location, and *Coordination*, i.e. studies that investigate

**Fig. 1** Correlation matrix of features, where values closer to 1 suggest a positive correlation between two variables, whereas values closer to -1 a negative one



communicative coordination between participants. Referential games, on the other hand, have a strong negative correlation ( $r = -0.85$ ) with this type of meaning space, and are more frequently used ( $r = 0.63$ ) in studies where participants have to communicate about meaningful concepts (meaning\_space\_1). Another strong correlation ( $r = 0.61$ ) exists between interchangeability\_2 (i.e. experiments in which participants switch between roles) and interaction\_2 (i.e. communication in dyads).

These results inform us about the limits of particular study designs and those of communication itself. Communicating about location is a complex process that typically requires coordination between participants, whereby they incrementally update their state of knowledge. It would be rather difficult to communicate about the location of an object in a referential game; doing so would perhaps be possible but would require an innovative design. The other strong correlation – between dyadic communication and interchangeability – can reflect a concern for interpreting the results of the study and removing factors related to the number of interlocutors involved in a conversation.

## Agglomerative Hierarchical Clustering

Agglomerative hierarchical clustering is an approach adopted in the exploratory analysis of multi-dimensional data (Nielsen, 2016). This method involves building a hierarchical tree from “leaves” - the most basic units - and iteratively builds a hierarchical structure. The leaves of the tree are merged on the basis of the smallest distance between them, then those merged leaves are aggregated into bigger units until the root of the tree is reached (Manning et al., 2008). The output of agglomerative hierarchical clustering is usually depicted in the form of a dendrogram; the dendrogram resulting from the analysis of our dataset can be accessed here: <https://cles.umk.pl/evolang-network/dendrogram/>.

The dendrogram analysis resulted in a tree structure that can be divided into three broad categories. The first division between papers occurs between a single paper

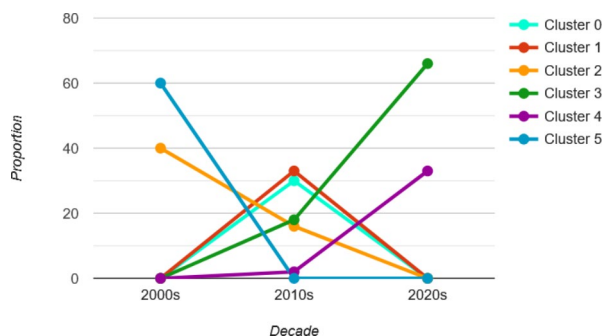
(Perlman & Lupyan, 2018) and the rest of the database. This divide occurs due to the fact that the study reported in that paper involved relatively rare design choices along several dimensions, such as feedback, group type and turn-taking. The second category in the dendrogram occurs between a sub-branch represented by such studies as Raviv et al. (2019), Garrod et al. (2007), Selten and Warglien (2007). What these studies have in common is their medium of communication (primarily drawings), referential game type, alignment of interests and the rigid assignment of roles (e.g. director and guesser). The last major subgroup consists of such studies as Żywicznyński et al. (2021), and Motamedi et al. (2018; 2019). In this group, a majority of studies were conducted with the use of referential games and a non-verbal bodily medium of communication (i.e. gestures or full-body pantomime).

## Experimental Semiotics over the Years

Finally, a meta-level application of the database is in identifying trends across time, to help achieve a deeper understanding of the historical development of Experimental Semiotics. To properly understand the evolution of ES studies, we have conducted a cluster analysis based on 11 features described in Sect. 2: “Presence vs. absence of Vertical transmission?”, “Referential vs. Coordination”, “Medium of communication,” “Signal space,” “Meaning space,” “Feedback,” “Communication type,” “Group size,” “Participants of the main study: Age,” “Turn-taking,” “Interchangeability of the signaller/receiver roles”. Using the Python programming language (van Rossum & Drake 1995) and the Scikit-learn library (Pedregosa et al., 2011), we ran a *k*-means clustering algorithm that classified papers into six clusters based on these features. The optimal number of clusters was determined using the Elbow Method. Figure 2 shows how the clusters are distributed over the considered decades, with each color corresponding to a specific cluster<sup>2</sup>.

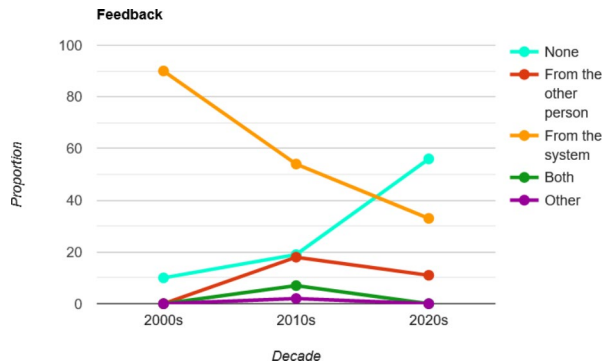
One thing that stood out when looking at the coded dimensions is the presence of static categories, that is, those dimensions that varied little or very little in most of the studies analyzed throughout the entire period of time considered. Among the static dimensions, we have “Communication type,” which was almost exclusively dyadic (although in recent years, there has been a greater presence of non-dyadic communication); “Presence or absence of Vertical transmission?,” which reports only four

**Fig. 2** A graphic representation of the result of clustering, with the x-axis representing the decades, and the y-axis the predominance (in proportion) of a cluster in the considered decade

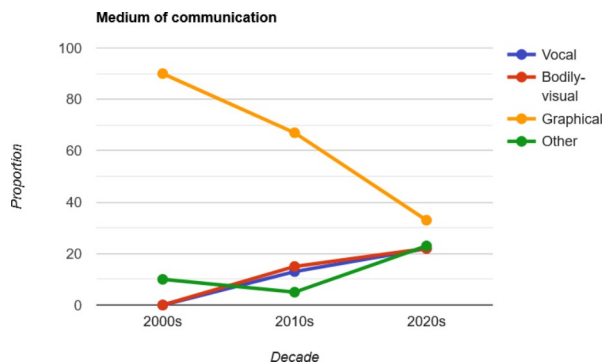


<sup>2</sup> All visualisations were produced using RapidTables (n.d.).

**Fig. 3** The “Feedback” dimension (over considered decades)



**Fig. 4** The “Medium of communication” dimension (over considered decades)



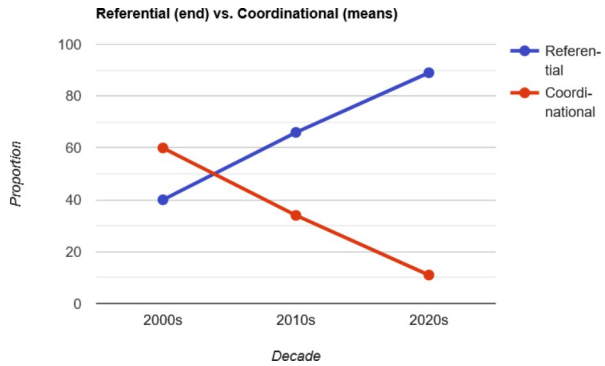
studies containing some transmission of the result of the communication task from one group to another; “Lab or online,” with only two studies that were conducted online; “Simultaneous interaction,” of which only six studies were not characterized by a contextual interaction; and “Alignment of interests,” with only two studies that included conflict of interest.

Regarding dynamic categories, on the other hand, an example is found in the “Feedback” dimension, which for the 2002–2007 period was characterized by a predominance of information received from the experimenter, or more generally, from the system. Only one of these studies reported different values. In Fig. 3, this phenomenon can be observed, with a progressive decrease in the paradigm over the considered decades.

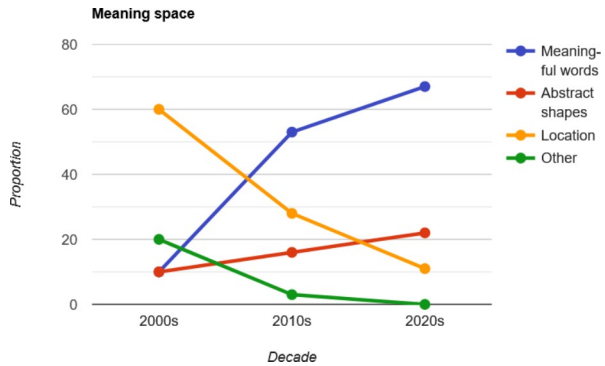
Another salient dimension is “Medium of communication”. In the period between 2009 and 2014 this dimension was characterized by the fact that it was made up almost exclusively of studies that used a graphical medium (except for one). In Fig. 4, it is possible to note that there has been a decrease in the number of studies that used a graphical type of communication over the decades.

Two closely related dimensions are “Referential vs. coordination” and “Meaning space.” Indeed, starting from 2014, it is possible to observe an almost exclusive use of referential games, which corresponds to an equally preponderant use of meaningful words/concepts of the relative “Meaning space” dimension (See Figs. 5 and 6).

**Fig. 5** The “Referential vs. coordination” dimension (over considered decades)



**Fig. 6** The “Meaning space” dimension (over considered decades)



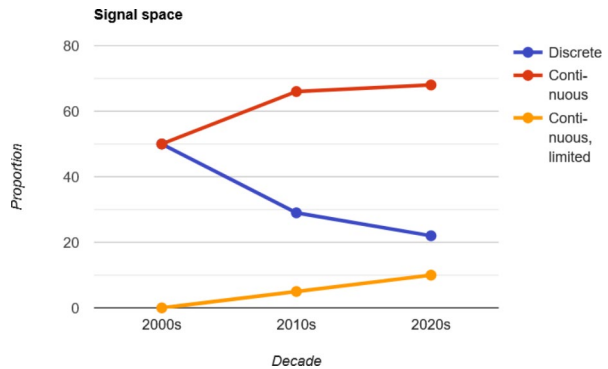
In both categories, it is possible to observe a progressively predominant use of the aforementioned values. This is directly related to the goal of the game: in referential tasks, where the goal of the game is to be understood, it is easier for the relative meaning to be made up of meaningful concepts (or, at most, abstract shapes); in coordination games, where communication is only a means for accomplishing the goal of the game, it is easier to have location as a meaning space.

Another interesting dimension is “Signal space”, which starting from 2014 becomes almost exclusively open and continuous (in only one paper this is not the case, (Fig. 7)).

**ES over the Years: Discussion**

One of the research questions we mentioned was whether there is any reason for the presence of static categories, that is, why most ES studies show largely identical values for specific dimensions. While the reasons for the lack of diversity in research designs under certain dimensions can be investigated in more detail in future studies, we offer some preliminary answers. It seems intuitive that dyadic communication is more suitable for the observation of communicative interaction according to the classic sender/receiver model, despite the recent increase in interest in non-dyadic studies. The relative absence of vertical transmission may be in large part due to our inclusion criteria being limited to studies on the creation of new communicative sys-

**Fig. 7** The “Signal space” dimension (over considered decades)



tems “from scratch,” while vertical transmission is typically studied in the lab with artificial language designs, where the initial signal-meaning pairings are given to the participants (e.g., Kirby et al., 2008). The paucity of studies carried out online could also be because the laboratory is perhaps more suitable for building experimental settings that are ecologically realistic. A similar argument may be valid for studies that do not involve simultaneous interactions.

It would be interesting to explore why there has recently been a drop in the number of studies that make use of the coordination task paradigm. One answer could be that it is more logistically challenging. However, studies that use coordination games are potentially of considerable interest for ES and, more generally for the analysis of the bootstrapping of communication systems, as some kind of alignment of interests is necessary to achieve coordination. Results of such studies could potentially highlight some of the social and cognitive dynamics that underlie communication and language. One potential reason for coordination tasks to be used less frequently is that it is more difficult to establish novel form-meaning pairings for purposes of coordination than for the sole purpose of identifying referents. This is because for referential communication games the potential meaning space is generally prespecified and limited, as opposed to the meaning space required for coordination games which is potentially open-ended.

Starting from 2014, the *signal space* dimension became almost exclusively open and continuous (Fig. 6.) This could be explained by the fact that ES studies seem to become increasingly ecologically realistic over time. The use of open and continuous signals is consistent with important threads in the literature on the evolution of language, for example, related to iconicity and holistic nature of early signs (Perlman et al., 2015). In a study by Nölle and colleagues (2018) gestural communication was used in order to express meanings represented by drawings of characters who belonged to categories delimited by, among other things, shared colours. This is an example of an open and continuous signal space, which was also the case with Zlatev and colleagues (2017).

Some of the research questions we asked relate to the problem of how ES studies have evolved over time: if there are particular trends in specific periods over the examined decades (although experimental semiotics is a rather new research field); if the categories can be related to each other in some way; if there is an explanation we can provide for the observed trends; or why some paradigms are systematically



ignored at the expense of others (e.g., why there are so few studies with vertical transmission). Other questions could refer to the results of the studies analysed, for example, whether similar results correspond to similar experimental paradigms or if there is evidence for specific empirical results that correspond to a coherent global picture, whether they are in line with the theoretical proposals, and so on.

## Conclusion

In this paper, we present a novel, systematic approach to the characterisation of ES studies according to the coding dimensions of the type of communication game, the presence of vertical transmission, the properties of the signaling and meaning spaces, the type of interaction, and the presence of the alignment of interest. This resulted in a dataset of 60 studies that were coded for these dimensions. In an exploratory analysis, we showed several potential applications of this dataset, including demonstrating how it can be used to examine changes in ES through a cluster analysis of the distribution of coding dimensions over time. This approach, along with the generated annotated dataset, has several potential applications. For example, it allows for a more fine-grained analysis of similarities and differences in the development of novel communication systems depending on the design features of ES studies. It also allows us to measure which dimensions cluster to provide more information about which experimental design is best suited for investigating particular research questions. Overall, an approach that systematically compares the underlying design properties of ES studies can help to specify the different mechanisms that influence the properties of novel, emerging communication systems.

**Author Contributions** A.D. contributed to conceiving and designing the analysis, collecting and coding the data, conceptualizing and writing the manuscript, and doing the cluster analysis. R.R. contributed to conceiving and designing the analysis, collecting and coding the data, and writing the manuscript. G.S. contributed to conceiving and designing the analysis, collecting and coding the data, and conceptualizing and writing the manuscript. A.G. contributed to conceiving and designing the analysis, collecting and coding the data, and conceptualizing and writing the manuscript, did the cluster analysis, and provided figures. M. Pleyer contributed to conceiving and designing the analysis, collecting and coding the data, and conceptualizing and writing the manuscript. M.S. contributed to the coding and to conceptualizing and writing the manuscript. M. Placiński contributed to the coding and to conceptualizing and writing the manuscript, did the cluster analysis, and prepared the figures. P.Z. contributed to the coding and to conceptualizing and writing the manuscript. S.W. contributed to conceiving and designing the analysis, collecting and coding the data, and conceptualizing and writing the manuscript. All authors reviewed the manuscript.

**Funding** This research was supported by the Polish National Science Centre under grant agreement UMO-2019/34/E/HS2/00248. M. Pleyer was supported under grant agreement UMO-2021/43/P/HS2/02729 co-funded by the National Science Centre and the European Union Framework Programme for Research and Innovation Horizon 2020 under the Marie Skłodowska-Curie grant agreement No. 945339. For the purpose of Open Access, the authors have applied a CC-BY public copyright licence to any Author Accepted Manuscript (AAM) version arising from this submission.

## Declaration

**Conflict of interest** The authors have no competing interests as defined by Springer, or other interests that might be perceived to influence the results and/or discussion reported in this paper.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Bar-On, D. (2021). How to do things with nonwords: Pragmatics, biosemantics, and origins of language in animal communication. *Biology & Philosophy*, 36(6), 1–25.
- Bartlett, F. C. (1932). *Remembering. A study in experimental and social psychology*. Cambridge University Press.
- Bavelas, A. (1950). Communication patterns in task-oriented groups. *The Journal of the Acoustical Society of America*, 22(6), 725–730.
- Bavelas, A. (1952). Communication patterns in problem-solving groups. In H. von Foerster (Ed.), *Cybernetics. Circular causal and feedback mechanisms in biological and social systems. Transactions of the ninth conference* (pp. 1–44). Josiah Macy Jr. Foundation New York, NY.
- Blokpoel, M., van Kesteren, M., Stolk, A., Haselager, P., Toni, I., & Van Rooij, I. (2012). Recipient design in human communication: Simple heuristics or perspective taking? *Frontiers in human neuroscience*, 6, 253.
- Caldwell, C. A., & Smith, K. (2012). Cultural evolution and perpetuation of arbitrary communicative conventions in experimental microsocieties. *Plos One*, 7(8), e43807. <https://doi.org/10.1371/journal.pone.0043807>.
- Campbell, R., & Grieve, R. (1981). Royal Investigations of the Origin of Language. *Historiographia Linguistica*, 9, 43–74. <https://doi.org/10.1075/hl.9.1-2.04cam>.
- Christiansen, M. H., & Chater, N. (2022). *The Language game: How Improvisation created Language and changed the World*. Penguin Press.
- Cuskley, C. (2019). Alien forms for alien language: Investigating novel form spaces in cultural evolution. *Palgrave Communications*, 5(1), 1–15.
- Dawkins, R., & Krebs, J. R. (1978). Animal signals: Information or manipulation. *Behavioural ecology: An evolutionary approach*, 2, 282–309.
- dos Santos, M., Rodrigues, J. F. M., Wedekind, C., & Rankin, D. J. (2012). The establishment of communication systems depends on the scale of competition. *Evolution and Human Behavior*, 33(3), 232–240.
- Esper, E. A. (1925). *A technique for the experimental investigation of associative interference in artificial linguistic material*. Linguistic Society of America.
- Fay, N., Walker, B., & Swoboda, N. (2017). Deconstructing Social Interaction: The Complimentary Roles of Behaviour Alignment and Partner Feedback to the Creation of Shared Symbols. In *CogSci*.
- Ferretti, F. (2022). *Narrative persuasion. A cognitive perspective on Language Evolution* (7 vol.). Springer Nature.
- Ferretti, F., & Adornetti, I. (2021). Persuasive conversation as a new form of communication in Homo sapiens. *Philosophical Transactions of the Royal Society B*, 376(1824), 20200196.
- Galantucci, B. (2005). An experimental study of the emergence of human communication systems. *Cognitive Science*, 29(5), 737–767.
- Galantucci, B. (2017). Experimental semiotics. In *Oxford Research Encyclopedia of Linguistics*. Retrieved 2 March 2023 from <https://oxfordre.com/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-210>. <https://doi.org/10.1093/acrefore/9780199384655.013.210>
- Galantucci, B., & Garrod, S. (2011). Experimental semiotics: A review. *Frontiers in human neuroscience*, 5, 11.
- Galantucci, B., & Roberts, G. (2012). Experimental Semiotics: An engine of discovery for understanding human communication. *Advances in Complex Systems*, 15 (03n04).
- Galantucci, B., Garrod, S., & Roberts, G. (2012a). Experimental semiotics. *Language and Linguistics Compass*, 6(8), 477–493.

- Galantucci, B., Theisen, C., Gutierrez, E. D., Kroos, C., & Rhodes, T. (2012b). The diffusion of novel signs beyond the dyad. *Language Sciences*, 34(5), 583–590.
- Garrod, S., Fay, N., Lee, J., Oberlander, J., & MacLeod, T. (2007). Foundations of representation: Where might graphical symbol systems come from? *Cognitive Science*, 31(6), 961–987.
- Iizuka, H., Marocco, D., Ando, H., & Maeda, T. (2013). Experimental study on co-evolution of categorical perception and communication systems in humans. *Psychological research*, 77(1), 53–63.
- Inoue, N., & Morita, J. (2021). A behavioral task for exploring dynamics of communication system in dilemma situations. *Artificial Life and Robotics*, 26, 329–337.
- Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*, 105(31), 10681–10686.
- Krauss, R. M., & Weinheimer, S. (1966). Concurrent feedback, confirmation, and the encoding of referents in verbal communication. *Journal of personality and social psychology*, 4(3), 343.
- Krebs, J. R., & Dawkins, R. (1984). Animal signals: Mind-reading and Manipulation. *Behavioral Ecology: An Evolutionary Approach* (pp. 380–402). Blackwell.
- Lee, J. J., & Pinker, S. (2010). Rationales for indirect speech: The theory of the strategic speaker. *Psychological Review*, 117(3), 785.
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge University Press.
- Morris, C. W. (1938). Foundations of the theory of Signs. *International encyclopedia of unified science* (pp. 1–59). Chicago University Press.
- Motamedi, Y., Smith, K., Schouwstra, M., Culbertson, J., & Kirby, S. (2018). The emergence of systematic argument distinctions in artificial sign languages.
- Motamedi, Y., Schouwstra, M., Smith, K., Culbertson, J., & Kirby, S. (2019). Evolving artificial sign languages in the lab: From improvised gesture to systematic sign. *Cognition*, 192, 103964.
- Nielsen, F., & Nielsen, F. (2016). Hierarchical clustering. *Introduction to HPC with MPI for Data Science*, 195–211.
- Nölle, J., & Galantucci, B. (2021). Experimental semiotics: Past, present, and future. In A. M. García, & A. Ibáñez (Eds.), *The Routledge Handbook of Semiosis and the brain* (pp. 66–81). Routledge.
- Nölle, J., Staib, M., Fusaroli, R., & Tylén, K. (2018). The emergence of systematicity: How environmental and communicative factors shape a novel communication system. *Cognition*, 181, 93–104.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., & Vanderplas, J. (2011). Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research*, 12, 2825–2830.
- Perlman, M., & Lupyán, G. (2018). People can create iconic vocalizations to communicate various meanings to naïve listeners. *Scientific reports*, 8(1), 2634.
- Perlman, M., Dale, R., & Lupyán, G. (2015). Iconicity can ground the creation of vocal symbols. *Royal Society open science*, 2(8), 150152.
- RapidTables. (n.d.). <https://www.rapidtables.com/>
- Raviv, L., Meyer, A., & Lev-Ari, S. (2019). Compositional structure can emerge without generational transmission. *Cognition*, 182, 151–164.
- Roberts, G., & Clark, R. (2020). Dispersion, communication, and alignment: An experimental study of the emergence of structure in combinatorial phonology. *Journal of Language Evolution*, 5(2), 121–139.
- Scott-Phillips, T. (2014). *Speaking our minds: Why human communication is different, and how language evolved to make it special*. Bloomsbury Publishing.
- Selten, R., & Warglien, M. (2007). The emergence of simple languages in an experimental coordination game. *Proceedings of the National Academy of Sciences*, 104(18), 7361–7366.
- Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origg, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language*, 25(4), 359–393.
- Stevens, J. S., & Roberts, G. (2019). Noise, economy, and the emergence of information structure in a laboratory language. *Cognitive Science*, 43(2), e12717.
- Tamariz, M. (2017). Experimental studies on the cultural evolution of language. *Annual Review of Linguistics*, 3, 389–407.
- Tomasello, M. (2008). *Origins of Human Communication*. MIT Press.
- Van Rossum, G., & Drake, F. L. Jr. (1995). *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam.
- Wolfe, D. L. (1933). The relative Stability of First and Second Syllables in an Artificial Language. *Language*, 9(4), 313–315. <https://doi.org/10.2307/409418>.

- Xiao, Y., & Watson, M. (2019). Guidance on conducting a systematic literature review. *Journal of planning education and research*, 39(1), 93–112.
- Zlatev, J., Waciewicz, S., Zywczyński, P., & van de Weijer, J. (2017). Multimodal-first or pantomime-first?: Communicating events through pantomime with and without vocalization. *Interaction Studies*, 18(3), 465–488.
- Zubek, J., Korbak, T., & Rączaszek-Leonardi, J. (2023). Models of symbol emergence in communication: a conceptual review and a guide for avoiding local minima. arXiv:2303.04544. <https://doi.org/10.48550/arXiv.2303.04544>
- Żywczyński, P. (2018). *Language Origins: From mythology to science*. Peter Lang.
- Żywczyński, P., Sibierska, M., Waciewicz, S., van de Weijer, J., Ferretti, F., Adornetti, I., & Deriu, V. (2021). Evolution of conventional communication. A cross-cultural study of pantomimic re-enactments of transitive events. *Language & Communication*, 80, 191–203.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

## Authors and Affiliations

**Angelo Delliponti<sup>1</sup> · Renato Raia<sup>2</sup> · Giulia Sanguedolce<sup>2</sup> · Adam Gutowski<sup>1</sup> · Michael Pleyer<sup>1</sup> · Marta Sibierska<sup>1</sup> · Marek Placiński<sup>1</sup> · Przemysław Żywczyński<sup>1</sup> · Sławomir Waciewicz<sup>1</sup>**

✉ Angelo Delliponti  
503302@doktorant.umk.pl

Renato Raia  
raiazenato@gmail.com

Giulia Sanguedolce  
sangu.giulia@gmail.com

Adam Gutowski  
adgut1509@gmail.com

Michael Pleyer  
pleyer@umk.pl

Marta Sibierska  
sibier@umk.pl

Marek Placiński  
marpla@umk.pl

Przemysław Żywczyński  
Przemyslaw.Zywczyński@umk.pl

Sławomir Waciewicz  
waciewicz@umk.pl

<sup>1</sup> Nicolaus Copernicus University in Toruń, Toruń, Poland

<sup>2</sup> Roma Tre University, Rome, Italy