NORMS FOR CONSTRUCTING LANGUAGE IN HUMANS AND ANIMALS

Fachbereich Erziehungswissenschaft und Psychologie
Arbeitsbereich Vergleichende Entwicklungspsychologie
der Freien Universität Berlin

kumulative Dissertation

zur Erlangung des akademischen Grades

Doktor der Philosophie (Dr. phil.)

Doctor of Philosophy (Ph.D.)

von

MSc

Ullrich, Robert

Berlin im April 2018

Tag der Disputation: 12.07.2018

1. Gutachterin: Prof. Dr. Katja Liebal

2. Gutachter: Prof. Dr. Markus Wild

Danksagung

Dank all jenen, die mich die letzten dreieinhalb Jahre unterstützt, beraten und bei Laune gehalten haben. Allen voran Prof. Dr. Katja Liebal, die die Dissertation primär betreute, nahezu alle Abstracts und Manuskripte als Erste las, kommentierte und positiv vorantrieb; die finanzielle Ressourcen bereitstellte, Konferenzreisen ermöglichte und mich auch in Momenten, an denen ich dem Aufgeben sehr nah war, weiter vorantrieb. Dank an meinen Zweitbetreuer Prof. Dr. Markus Wild für die Ermöglichung eines Forschungsaufenthaltes in Basel am Fachbereich Philosophie in den ersten Monaten und für die finanzielle Unterstützung in schwierigen Zeiten. Ohne sein Feedback wäre das erste Drittel der Dissertation nicht zustande gekommen. Dank auch an Prof. Dr. Norbert Müller für die Lektion, dass meine Gutgläubigkeit manchmal zu weit geht.

Darüber hinaus großen Dank an Dr. Kathrin S. Kopp,
Philipp Norton und Dr. Moritz Mittelbach für
kontinuierliches, kritisches Feedback, bzw.
für das unerbittliche Produzieren von
Programmiercode in R. Dank an
Prof. Constance Scharff.
Dank an Freunde.
Danke Familie.

Publication Notes

At the time of submitting the current dissertation, three studies were accepted for publication in peer-reviewed media. One study was submitted for publication, though not finally accepted.

Publication 1

Ullrich, R. (2016). From ‘speech’ to ‘gesture’: The ‘oral’ as norm in ‘language’ research. In G. Hartung & M. Herrgen (Eds.), Interdisziplinäre Anthropologie (pp. 179–208). Wiesbaden: Springer. https://doi.org/10.1007/978-3-658-14264-3_15

Publication 2

Ullrich, R., Mittelbach, M., & Liebal, K. (2018). Scala Naturae: The Impact of Historical Values on Current “Evolution of Language” Discourse. *Journal of Language Evolution*, *3*(1), 1–12. https://doi.org/10.1093/jole/lzx017

Publication 3

Ullrich, R., & Liebal, K. (in print). Times change, values change: Criteria for attributing *language* in species comparative research [expected summer 2018]. In E. Luef & M. M. Marin (Eds.), *The talking species: Perspectives on the evolutionary, neuronal, and cultural foundations of language*. Graz: Unipress Graz-Verlag.

Publication 4

Ullrich, R., Mittelbach, M., & Liebal, K. (submitted). Introducing *intention*: How an idea has spread within species-comparative science. *Animal Cognition*

Table of Contents

List of Figures

List of Tables

General abstract

When investigating the evolution of language, scientists often approach one prominent question: What makes humans human? While researchers might share a common question that motivates them to investigate the origin of language, they do not share a common definition of that original term. The current dissertation hypothesises that definitions of *language* are ever-changing, temporary constructions which are implicitly informed by historical and social values. By utilising a mixed-methods approach, which combines socio-historical research with quantitative strategies, three examples are investigated: (i) Language defined as *speech* (oral norm); (ii) Language defined as highest evolutionary achievement (*Scala Naturae*); (iii) Language defined from its external structure (*Behaviourism*). The dissertation reveals and illustrates that each of the historical constructions has its problems. Regarding (i), a historical argument from analogy to the deaf discourse reveals prejudices towards nonverbal forms of communication. Addressing (ii), a quantitative text analysis on 915 articles from a time span of 10 years testifies the use of value-laden adjectives in some publications. Analysing (iii), a citation network on 653 articles, published over the time of 69 years, illustrates how the term *intention* enters a discourse that originates from a behaviouristic era and turned into a cognitive one. The quantitative evidence revealed by the dissertation demonstrates: Science is not, never was and likely never will be free from social and historical influences. That is not a problem. It is a problem, however, to neglect or ignore those influences. The current meta-analysis points to them in order to enable the reader to develop a critical standpoint in relation to the current and past language origin discourse. Hence, providing evidence by systematic investigation of these values is an active contribution to scientific self-correction.

General introduction

1. On the author

1.1 Background

Admittedly this dissertation has an untypical, since personal beginning. The incentive to comment on the personal background and motivation of the author of the dissertation is an insight of the past three and a half years of work: Transparency about motivation, assumptions, and personal background plays a key role for reliable and good science.

An article from *The New Yorker* on the so-called *replication crisis* in science (see *General introduction*; section 3; p.10), concludes with a simple, though far reaching, sentence: "When the experiments are done, we still have to choose what to believe" (Lehrer, 2010). In fact, this is what I realised when I was finishing my master's project in 2010, preparing to write up the results for publication.

At that time, I was enrolled in an *Animal Behaviour and Neurobiology* programme. Prior to that, during my undergraduate studies in Biology, I was additionally attending courses in the Philosophy Department. Together with my former supervisor Constance Scharff we interpreted the results of my observations on courtship behaviour of male domestic zebra finches. The video recordings showed males approaching a female while *singing* and *dancing*. We felt, the multimodal combination of song and dance during courtship was worth to report, since it reveals that the zebra finch's signalling integrates more than one modality, possibly pointing towards a common biological substrate to human communication. We struggled, however, in finding the *appropriate* vocabulary: Do we describe *body movements* or *gestures*? Should we call it an act of *communication*, or *signalling*? Do we observe *dance* or *stereotypic movements*? Can we compare *human language* to *animal communication* or is that comparison a category error? What did we *observe* and what do we *believe*? We decided to write a mostly, as I would characterise it today, behaviouristic article, with some use of cognitive vocabulary. What was denoted as "stereotypic movement" at one point was taken up later as "gesture" whose further investigation "could extend the numerous parallels between human speech and birdsong" (Ullrich, Norton, & Scharff, 2016, p. 293).

I describe this personal experience, because the scene captures much of what should become the focus of the current dissertation project. Especially the influences from my current place of work (primate research in a cognitive-ethology tradition) as compared to my former place of work (birdsong research in a behaviouristic-ethology tradition) stimulated many questions which I later started to investigate.

1.2 Motivation

Yet again, I will annotate something that is not uncommon, but not commonly mentioned: The motivation of conducting the dissertation project changed in the course of project noticeable.

When I wrote the initial proposal to that project, I was primarily focussing on the question on how to provide a sound theoretical framework for studies on cross-species comparative language evolution. I believed that the struggles I experienced personally were caused by insecurities about knowledge of existing philosophical *schools of thought*. I believed that a more thorough analysis of the philosophical concept of language would help to improve theoretical substance and clarity in any empirical study.

When I realised that there is no philosophical *concept* of language, but myriads of *concepts*; and when I realised that the discussions – both in empirical as well as in theoretical studies – somehow repeat themselves with only little progress, I increasingly started focussing on the implicit background assumptions of researchers.

At this time I got in touch with a field called *meta-research*. Soon I understood that science is done by human beings and since every human being is part of a society it is – not deliberately – prone to bias, to manipulation, and mistakes. The author of those lines is no exclusion. It seems trivial to acknowledge that, but indeed it is not. Instead of coming to grips with theoretical concepts about *language*, I started to investigate research articles as empirical source material for empirical meta-analyses. My motivation shifted from the idea to qualify differences in the scaffold of a theory to the idea to quantify hidden norms and values instead. I believe that the current dissertation is not free from bias and mistakes. Obviously neither dropped in on purpose. The unusual personal introduction to that dissertation might help the reader to reveal possible biases right from the beginning.

2. The dissertation title

The title of this dissertation—*Norms for constructing language in humans and animals*—might need an explanation. Possible questions are briefly addressed hereinafter.

2.1 Why species comparison?

It might sound plain and simple, but it needs the very act of cross-species comparison to distinguish the human from non-human animals. For instance, as one compares humans to – say – other primates one will find that bipedalism stands out. As one compares humans to – say – birds one will find that bipedalism is not that exceptional. The same holds true for traits like human hairlessness, or the possession of an appendix. All of them are present *in a slightly different form* in other species as well. As of today all *physiological* traits, can be traced back in their evolutionary history to other species. They cannot serve as a characteristic to distinguish humans from non-human animals *qualitatively*. Instead they confirm a common insight from evolutionary theory that differences between species are

of degree and not of kind. That phrasing was originally used by Charles Darwin who literally wrote in *The Descent of Man*: "Nevertheless the difference in *mind* between man and the higher animals, great as it is, certainly is one of degree and not of kind" (Darwin, 1871/1963, p. 85, italics by R.U.)

Darwin went one step further when he wrote that sentence. He did not only predict gradual differences on a physiological, but also on a *psychological* level. It is exactly that prediction that is still discussed nowadays. As Marc Hauser once wrote: "[…] we tend to be much more possessive about some traits than others […]" (Hauser, 1997, p. 31). In fact, those traits researchers are currently *possessed* are of *psychological* nature. *If* there is any trait that separates human from non-human species *qualitatively*, it is assumed to be a *psychological* rather than physiological trait.

Consequently, asking for unique psychological traits was the origin of *comparative psychology* as it exists today: "The field of comparative psychology, […] investigates the evolutionary origins of human cognition […] comparative approaches have therefore been critical in generating and testing hypotheses regarding the origins of human-unique cognition." (Rosati, Wobber, Hughes, & Santos, 2014, p. 449).

The motivation to identify human unique traits is not only inherent to comparative psychology, but to a variety of other cross-species comparative projects, including those in anthropology, linguistics, biology or philosophy. For that reason the dissertation frequently refers to 'cross-species comparative science' in general, instead of 'comparative psychology' specifically.

Even at the psychological level, there are some traits that researchers are more interested in than others. To put it less formally, the possession on the human origin of language is a long-running issue in species comparative sciences. It needs further elaboration to understand what exactly *language* makes special as compared to other psychological states.

2.2 Why language evolution?

The intuition that humans are *exceptional* as compared to other non-human animals is widespread in the cross-species comparative literature. However, from a biological perspective, any species is exceptional, since the establishment of taxa requires uniqueness.

If one looks back in the history of cross-species comparative science one will find numerous claims concerning supposedly human unique features as illustrated in **Tab. 1**. More examples are reviewed in Hauser (1997, p. 31). The debate on any of the listed psychological traits in **Tab. 1** often follows a *similar structure*: A trait – for example, tool use and tool production – is believed to be unique to humans (Oakley, 1956). Then someone observes, as in the tool-case, that chimpanzees can *use* tools (Goodall, 1964). It will not take long that critics will agree that *tool-use* is not unique, but *tool-production* still is. Soon after, another scientist will make an observation on that issue and will find evidence for tool-*production* as well (Boesch & Boesch, 1990). The more evidence piles up for apes, the more observations for *tool-use* and *tool-production* will '*trickle down'* to non-primate species such

as birds (Weir, Chappell, & Kacelnik, 2002) or fish (Brown, 2012). At the end of a long debate, one often finds two possible outcomes. Either, scientists will accept that the trait at issue is not human unique, or researchers will choose to redefine it (Frederick, 2015, p. 300). As in the case of *tools* the question changed to: What exactly constitutes *making*? For some researchers the production of tools rests on "cumulative culture", "diagnostic learning", and a "conceptual system" (Vaesen, 2012; see Bentley-Condit & Smith, 2010 for Review). From this perspective the trait remains human unique. Others criticise that anthropocentric perspective and use the available evidence to conclude that the trait differs only in degree (e.g. Brown, 2012; McGrew, 2013; Pepperberg, 2018). Basically, such a debate as reproduced for tool-use and tool-making exists for any of the given examples in **Tab. 1**.

Table 1 Short list collecting few examples of traits once supposed to be human unique and nowadays under debate.

Claims about human uniqueness	**Doubts**
Tool production (Oakley, 1956)	Tool production in chimpanzees (Boesch & Boesch, 1990; Goodall, 1964) and birds (Bluff, Weir, & Rutz, 2007; Weir et al., 2002)
Episodic memory (Tulving, 1983) & **Mental time travel** (Suddendorf & Busby, 2003)	Episodic memory and mental time travel in birds (Allen & Fortin, 2013; Dally, Emery, & Clayton, 2006; Güntürkün & Bugnyar, 2016)
Recursion (Hauser, Chomsky, & Fitch, 2002)	Doubts about the pivotal role of recursion (Martins, Muršič, Oh, & Fitch, 2015) and first counter-experiments with birds (Abe & Watanabe, 2011)
Rhythm (Patel, 2006)	Rhythm perception found in a bird (Patel, Iversen, Bregman, & Schulz, 2009)

Language, however, still plays a special role to that discussion. The linguist Dereck Bickerton once wrote: "Uniqueness isn't the issue. Unlikeness is the issue." (Bickerton, 2009, p. 21). For researchers like Bickerton, the evolution of language appears very unlikely, because *language* seems to exist without *appropriate* comparison. He continues that "For every other 'unique' thing that's evolved, you can see what was there before it, what evolution had to work on in order to produce it. Not with language." (Bickerton, 2009, p. 21). While many claims about qualitative differences have fallen short, claims relating to the uniqueness of human language are pervasive. **Table 2** is an addition to a similar collection from Marc Hauser (1997, p. 33) and lists statements as they can be found frequently in the literature. Indeed, one could go so far as Hauser and say "We have, and probably always will have, an obsession about our uniqueness." (Hauser, 1997, p. 30). Many of the quotations from **Tab. 2** stress the qualitative difference between humans and other species.

To ask whether language differs qualitatively or quantitatively to animal communication is one of the main issues of the discourse. That is shown in an exemplary way by reference to a recent exchange between Noam Chomsky and Tecumseh Fitch, where the former insists that "language capacities" are truly "unique to humans in essentials" (Chomsky, 2017) and the latter contradicts that by writing "many human conceptual structures have clear animal homologs" (Fitch, 2017b).

Table 2 Collection of quotes from different scholarly fields, research groups and years, mentioning the uniqueness of human language

Reference	Claims concerning the uniqueness of human language
Maynard-Smith & Harper, 2003, p. 130	"Language is the crucial difference between humans and other animals"
Christiansen & Kirby, 2003	"Language is one of the hallmarks of the human species – an important part of what makes us human. Yet, despite a staggering growth in our scientific knowledge […] we know comparatively little about how our unique ability for language originated and evolved into the complex linguistic systems we use today."
Zuberbühler, 2003	"Language is without doubt one of the most intricate and complex behaviors known to date, and among the few that clearly distinguishes humans from the rest of the living world."
Bickerton, 2009, p. 4	"Language is what makes us human. [...] It's also the greatest problem in science."
Chater, Reali, & Christiansen, 2009	"Indeed, our unique and nearly universal capacity to acquire and use language has even been cited as one of eight key transitions in the evolution of life"
Jackendoff, 2011	"The human language faculty is a cognitive capacity shared by all normal humans but no other species on the planet."
Hobaiter & Byrne, 2011	"It is a truth universally acknowledged that the greatest cognitive difference between humans and other animals lies in the use of language; thus, it is no surprise that vigorous research attention has been paid to communication among primates."
Taglialatela et al., 2015	"Human language is unique within the animal kingdom."
Smit, 2016	"Humans are unique because they are language-using creatures."
Oesch, 2016	"Language may be one of most important attributes which separates humans from other animal species"
Townsend et al., 2016	"Language is considered to be one of the pinnacles of human biological evolution. Its emergence in the *Homo* lineage was presumably enabled by the presence of a set of cognitive abilities and ecological conditions not shared by other species. "

Basically, this argument exists as long as Charles Darwin put out his prediction of the human-non-human-continuity. It is worth reading the Chomsky and Fitch articles and to compare them with the first chapter from Gregory Radick's book *The Simian Tongue*. Radick portrays a very similar debate from the 19th century between the linguist Friedrich Max Müller ("Language is our Rubincon, and no brute will dare to cross it"; Müller, 1885, p. 354) and Charles Darwin (there is a "continuity across the language barrier"; Radick, 2007, p. 36). It seems that not much has changed ever since.

If a scholarly issue is contested over such a long period of time, than it suggests the assumption that the *real* dispute is on something else. The linguist Talbot Taylor once observed: "the narrative way of satisfying a people's feeling that they are special […] is the main distinguishing feature between humanity and the animal kingdom: that which separates 'us' from 'them'" (Taylor, 1997, pp. 67–68). As Taylor suggests in his essay, *language* may serve as famous *narrative* to human speciality. This is the case because among most other psychological traits, *language* stands out for its often attributed importance. Indeed, for many people language is *the* most important or primary trait as compared to

less important or secondary traits such as *tool production*, *mental time travel,* or *syntax.* While secondary traits can be explained by widely shared cognitive processes in many species, primary traits, like language, make possible all other human accomplishments (cf. Snowdon, 2004, p. 131). Sometimes the distinction is also made between a 'narrow' and 'broad faculty of language' - with a similar result (Hauser et al., 2002). In philosophy the importance of language is covered by the 'language of thought hypothesis', which holds that "thinking" – as prerequisite for many secondary traits – "takes places in a mental language" (Murat, 2010). Taken together these discussions explain why *language* is often perceived as key narrative to explain the supposed 'special status' of humans.

To sum up the chapter, the inquiry of *language* and the questions about its origin serve as example for the current dissertation project for four reasons: First, the debate has a very long historical tradition and dates back to at least ancient Greece (see Modrak, 2001). That allows analysing large time frames. Second, many animals do communicate, which is why comparisons to human language are so frequent and relate to a large collection of diverse species (see Naguib, 2006). Third, human language is often described as one of the crucial traits to distinguish humans from non-human species (see **Tab. 2**). Finally, researchers are more *possessive* about *language*, than about most other psychological traits because of the narrative function. That possession is assumed to inform the debate not only by actual evidence, but to influence also the subsequent *interpretation*. Interpretations are governed not only by facts, but by a variety of value-laden influences. The next paragraph will explain in more detail what that means.

2.3 Why norms and values?

> *Science is built up of facts, as a house is built of stones; but an accumulation of facts is no more science than a heap of stone is a house.* (Poincaré, 1905, p. 141)

The term 'norm' might evoke the most reaction to the dissertation's title. Norms, here, are broadly defined. They encompass any decision in scientific practice that is not based on evidence. Admittedly, the definition of norms was narrower in the beginning of the dissertation (Ullrich, 2016) as compared to the studies that followed. The reason for adjusting the notion was due to the *subtle*, *implicit*, and sometimes *unwanted* character of norms. Hence, the term *norm* got increasingly replaced in later studies by alternative notations like 'value' or 'social and historical influence'. As mentioned in *publication 2* (Ullrich, Mittelbach, & Liebal, 2018), the scholarly discussion usually distinguishes between *epistemological* and *non-epistemological* values (e.g. Dorato, 2004; Steel, 2010). The former are accepted values, because they encompass *good* practices like a scientist's conviction for transparency or reproducibility. The latter are perceived suspiciously, since they encompass moral judgments or political convictions. Hence, when it is said that *norms encompass any decisions in scientific practice that are not based on evidence*, it refers to non-epistemological forms, only. For the sake of convenience many authors call those non-epistemological norms simply *values* (cf. Elliott & Willmes, 2014).

Whatever they are called, the presence of *values* or *non-epistemological norms* is more likely perceived as an *error* in science, or as an indicator for *bad scientific practice*. As such it should not have to play a role: "Subjective accounts of the world are the remit of art, not of science." (Mogie, 2000). In fact, not only scientists themselves are sceptical about values and norms, but the general public likewise. In a study that examined how citizens view scientists when they publicly acknowledge values, Elliott and colleagues found preliminary evidence of reduced credibility (Elliott, McCright, Allen, & Dietz, 2017). Consequently, many scholars defend a *value-free ideal* of science (Betz, 2013; Hudson, 2016; for an overview, see Reiss & Sprenger, 2014).

One inspiration for the value-free ideal dates back to the year 1786, where the philosopher Immanuel Kant wrote that the only sound scientific investigation rests on mathematics[1] (Kant, 1786/2003, p. 15). According to that, non-mathematical influences – such as narratives, values and the like – are problematic, because they are simply unscientific. Roald Hoffmann explains: "Because narrative is not reducible to mathematics, it is not given its due in our scientific world." and he continues: "Too bad; storytelling is both ancient and deeply human. It is a shared treasure between science and the arts and humanities." (Hoffmann, 2014).

The second part of the Hoffmann quotation includes the central issue to the value-free ideal. Again, one can explain this with Kant, who mentions a few lines before his earlier quoted statement that scientific inquiry inevitably rests on metaphysical assumptions, because the terms and concepts needed for scientific description cannot be empirically measured or mathematically deduced[2] (Kant, 1786/2003, p. 13). Indeed, there exists a growing body of literature that suggests that values are inevitably intertwined with scientific practice (Allchin, 1999; Douglas, 2009, 2016; Elliott, 2017; Longino, 1990; McKaughan & Elliott, 2015). The following cartoon (**Fig.1**), which is famous in the internet and sometimes shown at conferences on animal cognition, might illustrate that proposal.

As the person on the desk asks the recipients to perform a "fair" task, many questions arise to the viewer: On which basis did the researcher select the species and why? On which basis did the scientist *define* the task? On which basis did the examiner formulate the expected outcome? The dissertation argues that the answers to these questions are not based on evidence, but on values. The cartoon reveals the omnipresence of them in scientific practice.

[1] original: „Ich behaupte aber, daß in jeder besonderen Naturlehre *nur* so viel eigentliche *Wissenschaft* angetroffen werden könne, als darin *Mathematik* anzutreffen ist" (Kant, 1786/2003, p. 15, italics in original)

[2] original: „Eigentlich so zu nennende Naturwissenschaft setzt zuerst Metaphysik der Natur voraus; denn Gesetze, d.i. Prinzipien der Notwendigkeit dessen, was zum Dasein eines Dinges gehört, beschäftigen sich mit einem Begriffe, der sich nicht konstruieren lässt, weil das Dasein in keiner Anschauung a priori dargestellt werden kann." (Kant, 1786/2003, p. 13)

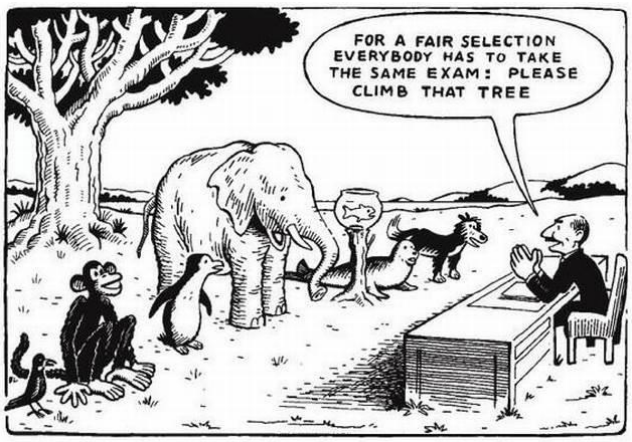

Figure 1 Cartoon from unknown origin illustrating the influence of values on scientific practice. See text for details. (Retrieved from: goo.gl/o5fErW; 27.03.2018)

Values govern the use of terminology, the types of models regarded as most compelling, the experimental design, the strength of evidence required, the conclusions drawn from an observation, the explanations regarded as appropriative, and the story-telling in publications following the experiments (Douglas, 2016; McKaughan & Elliott, 2015). Just the very perspective on a topic – top down versus bottom up – turns out to be a value judgement (de Waal & Ferrari, 2010). That is to say, when researchers decide to investigate the "pinnacles of [human] mental evolution" (de Waal & Ferrari, 2010) such as linguistic abilities (= top down), they are more likely to find human unique traits, than focussing otherwise on vocal learning and associative learning (=bottom up). Once more, a citation from Marc Hauser's book *Animal Communication* will suggest a reason for that: "The history of our species shows that every time a discovery has been made that challenges our domination of the animal kingdom, we are disbelieving at first and, once convinced, unleash all of our intellectual horsepower and search for something else that will set *us* apart from *them*." (Hauser, 1997, p. 30 italics in original).

As reasons where presented above, the current dissertation assumes that values are inextricably linked to cross-species comparative research on language evolution. Scientists are humans and hence part of a value-based and value-laden society. Consequently, it is the idea of the dissertation to reveal and to monitor those values. The idea is further illustrated by Kevin Elliott in his recent book *A Tapestry of Values*. The quote summarises the purpose of the dissertation:

> *Attempting to exclude values is a bit like claiming that knives should no longer be allowed in kitchens because people could be injured by them. Values can cause serious problems in science, just like knives can cause significant injuries, but the fact that they can be used unwisely or inappropriately in some cases does not mean that they are problematic under all circumstances. Values have important roles to play in scientific reasoning; the key is to recognize the variety of ways in which they can exert their influences and to figure out when those influences are appropriate and when they are not.* (Elliott, 2017, p. 8)

3. *The advent of Meta-Research*

Research produces new knowledge which is communicated in form of publications. Publications and their associated citations constitute the scientific discourse. Consequently quantifying the number of publications or citations often serves as measurement for scientific activity. Apart from the obvious difficulties to measure the real numbers, estimations find that scientific activity is constantly growing (Ioannidis, Boyack, & Klavans, 2014; National Science Board, 2018, pp. 17–20; Tenopir & King, 2014). A recent bibliometric analysis estimates that research output increases by 8-9% each year (Bornmann & Mutz, 2015). This would result in the doubling the scientific output every nine years (ibid.). The estimation was trialled by the author of these lines, by performing a quick search at the (fragmentary) scientific database *Web of Science* that indexes 2,929,480 publications for the year 2016 but approximately half of that number in 1996 (=1,497,089).

However, the reason to bring up these numbers is the simple fact that someone has to read all the published knowledge. As should be apparent at first glance, it is already impossible to keep track of just one sub-discipline, such as comparative psychology. In most cases researcher simply have to *trust* the knowledge published therein. The establishment of peer-review processes is one way to empower trust. Journals serve as gate keepers, whereas more prestigious journals receive the most trust because of their strict review processes – so much for the theory. In practice, trust in science started fading in the course of the last years.

3.1 Fading trust in science

> *Everyone knows that plumbers, carpenters, electricians cannot always be trusted and that it is wise to keep an eye on them. […] The same applies to the so called 'higher' professions"* (Feyerabend, 1978, p. 97)

The year 2005 marks an important turning-point in research on science. Back than the professor of medicine and statistics John Ioannidis published an article titled *Why most published research findings are false* (Ioannidis, 2005). According to Google Scholar (March 20th 2018), the article is referenced by 5,619 academic publications. That might come as a surprise for three reasons: First, the Ioannidis (2005) paper is not the only one that criticises current scientific practice as deficient. For instance, direct evidence for defective statistics and research procedures was already gathered in the year 1966, where the authors Schor and Karten (1966) systematically investigated articles from medicine in relation to their statistical soundness[3]. Second, the Ioannidis essay is written in a fairly technical language, including mathematical formulas and theoretical corollaries. Beside the catchy title, it is not easy to comprehend. Third, the essay is not an original meta-research article, with an actual quantitative analysis performed on a dataset. Rather, it is the introduction of a model, based on mathematical logic which predicts that well-validated large statistical effects are very rare and

[3] From 295 articles they checked, they found in only about half of the cases (52%) a sound statistical validity (Schor & Karten, 1966).

uncommon. Indeed, they are so rare that they should be observed in everyday science much less, than they are reported in the current scientific discourse. For that reason, Ioannidis suspects that most of the findings must be wrong (Ioannidis, 2005). That does not only relate to psychology, but to any scientific discipline.

However, apparently Ioannidis was in the right spot at the right time. The essay was published when the academic world experienced a couple of fraud-scandals. One of them became famous as 'Schön-scandal' (Brumfiel, 2002). In 2001, the physicist Jan Hendrik Schön published on average one peer-reviewed article on nanotechnology every eight days (ibid.). The outstanding results he measured in the Bell-laboratories could not be replicated by other research groups, because they were fabricated. However, the scientific protection measures such as peer-review and the need for successful replication did not reveal his fraud- case. He *himself* gave reasons to doubt as he used one and the same figure for different publications. After that, some physicists noticed that the depicted random noise in different figures was identical, whereupon investigations on the possibility of scientific misconduct started (Beasley, Datta, Kogelnik, Kroemer, & Monroe, 2002). As a result, eight *Science*-articles and another six *Nature*-publications were withdrawn, all released in 2000 and 2001. According to the blog *retraction watch,* a total number of 31 publications were red-flagged as fraud (retraction watch, 2018). The scandal raised questions about the role of co-authors and the role of 'prestigious' scientific journals as gate-keepers.

Although the dimension of the Schön-scandal in physics was exceptional, it was not the only scandal at that time. In late 2003, Anders Pape Møller, an expert in behavioural ecology, was found guilty of scientific dishonesty where "good lab practices hadn't been followed" (Odling-Smee, Giles, Fuyuno, Cyranoski, & Marris, 2007). In 2004 and 2005, the stem-cell researcher Woo Suk Hwang published two papers in *Science*, whose claims were verified as fabricated (Odling-Smee et al., 2007). In 2005, the MIT-based immunologist Luk van Parijs admitted that he altered data in at least one published paper and in various grant applications (Odling-Smee et al., 2007). In 2001, the Harvard psychologist Karen Ruggiero admitted that she fabricated five experiments, published in two peer-review articles. She was further convicted for manipulating results in a third publication (Price, 2010). Even more attention received another Harvard psychologist and biologist six years later. Marc Hauser, was an established figure in the field of cross-species comparative cognition, with various articles in *Science* and *Nature*. In 2011, the University of Harvard found him guilty of scientific misconduct in eight cases, involving a lack of data integrity and missing raw data (Samuel Reich, 2012).

There are more fraud-cases for the years *following* 2005, after Ioannidis published his essay. However, the main message from those early scandals was that there was no institution (Harvard, MIT, Bell Labs), no journal (*Science, Nature*), no subject (Medicine, Ecology, Physics, Psychology) and no reputation that could prevent scientific misconduct. The overall fading trust in science initiated a number of meta-research activities and the Ioannidis (2005) article just marked the advent of it.

3.2 Questionable research practices

Scientific misconduct is the most extreme and *most rare* case of bad science as compared to the more prevalent *questionable research practices* (henceforth QRP) (Martinson, Anderson, & de Vries, 2005). QPRs were first defined in 1992 by the *Panel on Scientific Responsibility and Conduct of Research* as "actions that violate traditional values of the research enterprise and that may be detrimental to the research process" (Commitee on Science Engineering and Public Policy, 1992, p. 5). The panel listed activities such as "inadequately supervising research subordinates or exploiting them" or "refuse to give peers reasonable access to unique research material" (ibid., p.6). In 1992 little empirical validation was done on QRPs, which is why the Committee concluded that they "do not directly damage the integrity of the research process" (ibid., p.5).

This perception changed significantly the following 25 years. In 2017 Munafò and colleagues write that QRPs "will serve to undermine the robustness of published research, and may also impact on the ability of science to self-correct" (Munafò et al., 2017, p. 2). In their article they provide a collection of recent meta-research studies that revealed and quantified various QRPs as summarised in **Tab. 3.**

Table 3 List summarising prominent *questionable research practices* (QRPs) as identified by recent meta-research.

Impediments	**Selected References**
'HARKing' (hypothesising after results are known)	Kerr, 1998
'selective reporting' (incomplete and biased); relates to the 'Chrysalis Effect' (an increase in Type I errors and a suppression of null effects)	Chan, Hróbjartsson, Haahr, Gøtzsche, & Altman, 2004; O'Boyle, Banks, & Gonzalez-Mulé, 2017
poor data availability	Wicherts, Borsboom, Kats, & Molenaar, 2006
'publication bias' (aka 'positivity bias': the publication of positive results, only)	Fanelli, 2010
'analytical flexibility' (choosing among dependent variables, sample size, covariates, and reporting subsets of experimental conditions)	Simmons, Nelson, & Simonsohn, 2011
'Voodoo correlations' (setups for producing idealised, inflated effects)	Fiedler, 2011
lack of replication	Makel, Plucker, & Hegarty, 2012; Open Science Collaboration, 2015
'power failure' (too low statistical power)	Button et al., 2013
use of overtly positive words (e.g. 'novel' 'groundbreaking', 'innovative', etc.) relating to the interpretation of data	Vinkers, Tijdink, & Otte, 2015
poor statistical reporting	Nuijten, Hartgerink, Assen, Epskamp, & Wicherts, 2015
'p-hacking' (data dredging for statistical significance)	Head, Holman, Lanfear, Kahn, & Jennions, 2015

Indeed, nowadays QPRs are perceived as a prevailing, large 'grey area' of scientific practice. They undermine the trustworthiness of scientific results not only in psychology, but in many disciplines (de Vries, Anderson, & Martinson, 2006; John, Loewenstein, & Prelec, 2012).

In 2009, Danielle Fanelli published a meta-analysis on 21 surveys that asked scientists anonymously concerning their QRP-behaviours. He found that on average about 2% of all respondents fabricated or modified data or results at least once (=scientific misconduct). About 34% of the respondents admitted deploying QRPs (Fanelli, 2009). When the same researchers were asked what they thought about the behaviour of their *colleagues* in other institutions, they estimated that peers would manipulate data in 14% of all cases and apply QRPs to 72% (Fanelli, 2009). Some researcher criticise the methodology of these surveys and suggest that the effect of QPRs might be overestimated by them (Fiedler & Schwarz, 2016).

Still, the prevalence of QPRs might explain other findings, such as the complications to replicate scientific results. As the psychologist Brian Nosek once formulated: "To show that 'A' is true, you don't do 'B'. You do 'A' again." (Yong, 2012, p. 300). However, the fundamental need to replicate studies does find its limits. When a large collaboration of psychologists attempted to replicate 100 experimental and correlational studies from three journals of psychology, they revealed that the mean effect size of their replications was about half the size as compared to the published material (Open Science Collaboration, 2015). While 79% of the original studies could report statistically significant results, that was only the case in 36% of the replications (ibid.).

The result of the large meta-analysis supports those who years earlier diagnosed a *replication crisis* (Pashler & Harris, 2012; recent review: Begley & Ioannidis, 2015). The *replication crisis* was first described in psychology, but concerns fields outside psychology as well. After the publication from the Open Science Collaboration, *Nature* released a survey where 1,576 scientists were asked about their opinion on current scientific practice. The survey found that 52% of the respondents believe that there is a "significant crisis" involving not only psychology, but various subjects of science such as medicine (Baker, 2016).

For the latter case – in biomedical research – an amount of US$100 billion is invested per year globally. Chalmers and Glasziou estimate that 85% of that investment is "wasted" because the research does not produce meaningful results (Chalmers & Glasziou, 2009). Such estimations might be difficult to prove. However, there exist indeed cases in the history of science where effort and money was used in order to produce unfounded evidence. John Ioannidis recounts three such examples: (i) neuropsychological research on 'Phrenology' in the 19th century; (ii) empirical eugenic research in Germany (1933-45); and (iii) research by the tobacco industry in the 20th century on 'unclear' consequences of smoking (Ioannidis, 2012). Ioannidis concludes that "self-correction does not always happen to scientific evidence by default" (Ioannidis, 2012). For that reason the following chapter presents some ideas that are discussed to improve scientific self-correction mechanisms.

3.3 Self-correcting science

In 2012, the former president of the *Association for Psychological Science* (APS) Henry Roediger wrote a comment including a sentence therein that is accepted as common-sense since centuries: "Science is a self-correcting process, although sometimes the correction is slow in coming." (Roediger, 2012). The reason why Roediger found himself constrained to write that sentence was the revealed fraud case of the well-established social psychologist Diederik Stapel shortly before. By 2015, *retraction watch* announced that 58 peer-reviewed publications had to been flagged as fraudulent, including one paper withdrawn by *Science* (Palus, 2015). Interestingly Stapel's misconduct was not detected by peer-review processes or by unsuccessful replications, but by whistle-blowers inside Stapel's department. Like in earlier cases, critics asked why nobody (e.g. funding agencies, journals, or co-authors) became aware of his fraud. Stapel himself answers that question laconically in an radio interview: "nobody noticed it because I was doing things everybody liked" – and benefited (Knetsch, 2017).

The case raised doubts on the mechanisms of scientific self-correction: "Unfortunately, however, there is every reason to believe that the great majority of errors that do enter the literature will persist uncorrected indefinitely, given current practices." (Pashler & Harris, 2012, p. 535).

On top of that, the pessimists do not see much changing behaviour by the scientists themselves. There is for instance the 'file drawer problem', which encompasses that non-significant studies remain often unpublished. The problem was first described as *positivity bias* decades ago by Theodore Sterling (1959) and Robert Rosenthal (1979). In 1959, Sterling could show that 97% of the publications in four major psychological journals were reporting statistically significant results only. When he reassessed the same study in 1995, the results were still the same (Sterling, Rosenbaum, & Weinkam, 1995). Despite the description of the 'file drawer problem' and its negative consequences, nothing had changed over a period of 30 years. Similar results come from a study that investigated the prevalence of reporting inconsistent p-values. Despite calls for refined methodological competence the poor reporting remained stable over a period of 28 years (Nuijten et al., 2015; see also a case study on unchanged low statistical power in: Smaldino & McElreath, 2016). From that perspective, it might seem as if scientific self-correction mechanisms fall short, because well-known errors are not removed from the literature.

However, the current dissertation does not share an overtly pessimistic view. To the contrary, it perceives the advent of meta-research as vivid example of science as self-correcting: "Accumulating evidence [through meta-research, R.U.] is the scientific community's method of self-correction and is the best available option for achieving that ultimate goal: truth" (Open Science Collaboration, 2015).

Meta-research revealed that the actual problems are not to be found in singular cases of intentional fraud, but arise from the large 'grey area' of *questionable research practices.* In a next step,

researchers started questioning the established scientific infrastructure and began to re-organise it: "In order to improve the culture of science, a shift must be made away from correcting misunderstandings and towards rewarding understanding" (Smaldino & McElreath, 2016). Both, the identification of problems and the subsequent correction appears immanent to meta-research and to scientific self-correction.

The suggestions from meta-research to alter the scientific *infrastructure* include:

i. promote an *open research culture* (open data, open methods, open workflow, open publishing) (Nosek et al., 2015; Nosek & Bar-Anan, 2012)
ii. create "*badges* to acknowledge open-science practices" (Munafò et al., 2017)
iii. *pre-registration* of studies (Wagenmakers, Wetzels, Borsboom, von der Maas, & Kievit, 2012)
iv. alter the *incentives to publish* "promote truth over publishability" (Nosek, Spies, & Motyl, 2012) and also alter "the *incentives for career success* in science" (Smaldino & McElreath, 2016)
v. reward *replications* through 'Registered Replication Reports' (Ioannidis, Boyack, et al., 2014)
vi. "*abandon journals* altogether, in favour of a library-based scholarly communication system" (Brembs, Button, & Munafò, 2013)
vii. develop *reporting checklists* for appropriate and transparent descriptions of statistical methods (Simmons et al., 2011)
viii. build a technical *open source infrastructure* to enable open collaboration and data storage (e.g. Open Science Framework) (Nosek et al., 2015)
ix. acquire *evidence through meta-research* to encourage change (Schooler, 2016)

Some of the suggestions might appear 'revolutionary'. However, a defining quality of any *crisis* is the need for change. 'The correction is slow in coming' was Roediger cited above. Indeed. As a first success one can regard the implementation of a new meta-research section in Plos Biology in 2016 (Kousta, Ferguson, & Ganley, 2016). The same journal announced in 2018 that "scooped" manuscripts will be considered for publication in order to encourage "organic" replication ("better than any post-hoc") (Plos Biology, 2018). In 2015, a first multidisciplinary journal (*Royal Society Open Science*) launched a registered reports option. As of November 2016, over 40 more specialist journals followed (Munafò et al., 2017). Also, journals retract erroneous papers faster. Steen and colleagues found that the number of retracted papers increased after 2002 and that the time-to-retract decreased as compared to before 2002 (Steen, Casadevall, & Fang, 2013).

The *replication crisis* and developments in meta-research are still developing. It remains to be seen, if the suggested and implemented changes in the scientific infrastructure will gain acceptance. In spite of that, meta-research is sometimes described as "hottest field to emerge" (Schooler, 2016) which "could rescue the 'replication crisis'" (Schooler, 2014). As meta-research is still *emerging*, the next chapter will explain in what way the dissertation project contributes to that development.

3.4 Why mixed-methods?

Research that focusses on scientific method, on values in science, or the sociology of scientific knowledge, is often denoted as *qualitative* (Alasuutari, 2010). Qualitative research on science can be

manifold, as expressed in philosophy of science (e.g. Karl Popper, Heather Douglas), sociology of science (e.g. Bruno Latour), history of science (e.g. Gregory Radick), or socio-history of science (e.g. Thomas Kuhn, Paul Feyerabend). Meta-research has its roots in these qualitative research traditions. It is distinguished from them by the reliance on quantitative analysis (Schooler, 2014).

For instance, while qualitative research might want to describe and reveal *research traditions* and *values*, it is the aim of meta-research to reveal and quantify structural biases like 'selective reporting bias' and 'positivity bias'. In other words, while the former focusses on narratives, the latter focusses on numbers. Contemporary meta-research on cognitive science is mostly focussed on arbitrary statistical analyses, post-hoc data-mining, selective analysis and other *questionable research practices* (Ioannidis, Munafò, Fusar-Poli, Nosek, & David, 2014).

As the current dissertation is interested in research on language evolution, it must be acknowledged that most previous work in that field was predominantly qualitative. A reason might be that many strategical decisions in research on language evolution rely on conceptual convictions, which are difficult to quantify. Meta-research, by contrast, takes another perspective:

> *Evaluating biases in each single study is attractive, but most difficult, because the data are usually limited, unless designs and analysis plans are registered a priori. It is easier to evaluate bias across multiple studies performed on the same question. When tests of bias are applied to a wider scientific corpus, it is difficult to pinpoint which single studies in this corpus of evidence are affected more by bias. The tractable goal is to gain insight into the average bias in the field.*(Ioannidis, Munafò, et al., 2014, p. 236)

The aim of this dissertation is to use the perspective and the tools from meta-research and to extend it to a field where historical and social influences might play a large role. The idea is to broaden the perspective of meta-research towards socio-historical influences on science. While contemporary meta-research investigates incentives such as *for what reason* (e.g. to make a career) or *how* (e.g. selective) researchers publish their findings, the dissertation investigates incentives such as *what* is investigated (e.g. modality vs. cognition) or *why* (e.g. promote human-non-human differences or similarities).

Hence, the dissertation is a combination of both mentioned strategies: qualitative and quantitative methods. The quantitative methods are used to investigate how researchers inquire and work on language evolution. The qualitative methods are used to identify influential articles and people, as well as *schools of thought* and social values. That pluralism of strategies makes it an example for mixed-methods research. The strategy is employed in order to compensate for limitations resulting from each respective methodology such as historical amnesia in quantitative and a poverty of a large amount of data in qualitative research traditions (Pluye & Hong, 2014).

Summary of the main results

Each of the four articles published as part of the current dissertation were conceptualised, designed, coordinated, interpreted, and drafted by the author of these lines. The two more quantitative articles (*publication 2* and *4*) received substantial contributions in terms of data analysis and programming of *R*-code by Moritz Mittelbach. All four articles received intellectual content from Katja Liebal and were also revised by her.

The four resulting publications interrelate in terms of their scope. While the two more quantitative publications (*publication 2* and *4*) focus on the collection of empirical evidence, the two more qualitative publications (*publication 1* and *3*) take a more socio-historical perspective and supplement the other two.

The common thread of all the studies is to investigate how cross-species comparative research proceeded and proceeds on language evolution. More precisely the articles investigate three defining properties and constructions of language as present in the past discourse. Reasons for the specific interest in language and its evolution are listed in the *General introduction* (section 2.2; p.4).

One overall finding from all four publications is that cross-species comparative research on language evolution is not, never was and likely never will be free of social and historical influences. All of the three chosen and investigated values are grounded in the past. The values are analysed from a retrospective standpoint. Despite their historical origin, aftermaths up to the present day were revealed in various subtle ways. The three norms and values encompass:

(i) the value of *directed progress* (e.g. Language defined as highest evolutionary achievement; aka *Scala Naturae*)
(ii) the *oral norm* (e.g. Language defined as *speech*)
(iii) the *behaviouristic norm* (e.g. Language as defined from its structural, physical expression, as compared to its cognitive dimension)

The investigations can hardly be treated separately from each other. For instance, no matter what focus an article puts on a certain investigation, a primate bias will always be apparent. Consequently evidence for each of the listed values is spread in all four articles. Due to the scattered nature of the results, the summary does not follow the chronology of their publication, but their order of numeric listing above.

It makes sense to begin with one ‘major cultural value’ that ‘has dominated the field’ for decades (Ruse, 1996, p. 14). The value of *directed progress* was qualitatively investigated in detail by various authors (e.g. Ghiselin, 2005; Hodos & Campbell, 1969; Ruse, 1996). The idea that evolution follows *directed progress* so that species can be charted in ascending order (on a *Scala Naturae)*, left their marks on many constructions of *language*. Often, *language* is perceived as ‘pinnacle of human evolution’ (Townsend et al., 2016, p. 2). In that case, *language* is constructed as something *high*,

which has consequences for its investigation in cross-species comparative research. In order to quantify the value of *directed progress* in the current literature a corpus of 915 peer-reviewed articles was analysed using a text-mining strategy (*publication 2*; Ullrich, Mittelbach, & Liebal, 2018). It was the aim of *publication 2* to compare the use of 'progressionist attributes', the range of study species, and the frequency of the term 'unique' between two subsets (*language* and *communication*).

The investigated articles mostly avoid direct linkage of 'high' and 'low' to various species groups. By contrast, until the middle of the 20th century, articles often contain terms like 'lower invertebrates' or 'higher primates' – a direct expression of *Scala Naturae*-thought (cf. Rigato & Minelli, 2013). However, as *publication 2* shows, the value of progress continuous to exists in a more subtle and implicit form. The use of 'directional vocabulary' depends on study species, and article type (e.g. review vs. research article), but not on membership to one of the subsets (communication vs language). Differences between the subsets exist in terms of 'narratives'. An indirect measurement for that claim is the extraction of the 80 most common adjectives for each respective corpus. While articles from corpus 'language' often make use of adjectives like 'human', 'cognitive', 'cultural' or 'semantic', articles from corpus 'communication' use instead 'male / female', 'sexual', 'aggressive', or 'territorial'. The different use of adjectives may trigger biased perceptions in readers who might associate *language* with more 'sophisticated' adjectives than communication. That, again, can influence the construction of *language* in multiple ways.

Another interesting finding from this study is the focus on modality as compared between the subsets. While in subset 'language' primates dominate the discourse, in subset 'communication' insects and birds take that role. Although subset 'language' is dominated by primates - which are often perceived as "unsatisfactory springboard for vocal learning" (Nottebohm, 1972, p. 133) - the focus on oral modality is more pervasive (58.5%) as compared to the subset 'communication' (41.4%). **Figure S1** (*publication 2*, Supplementary Results, p. 62) illustrates that researchers aim to investigate the evolution of language primarily by focussing on the oral modality. Hence, **Figure S1** provides additional evidence for the investigation of the *oral norm* in *publication 1* (Ullrich, 2016).

Publication 1 investigates the historical idea that the oral modality would embrace the whole concept of language. It is the idea of that publication to highlight similar prejudices in research on human and non-human forms of communication, regarding modality. The study reveals that in older articles oral forms of communication are usually more valuated than non-oral or multimodal alternatives. The publication is meant to conceptualise the impact and function of norms in everyday scientific practice. The socio-historical analysis explains the shift of deaf people's sign language, once perceived as 'deficient pantomime', towards an acknowledged natural *language*. To allow that change, the normative component of the oral construction had to be analysed. From the 1960s onwards constructions of *language* increasingly reject the idea that defining properties of language must rest

on mere physical or structural features. Instead inner mental capacities become subject of investigation (again) – both, in human and non-human species.

The epistemological change that demanded such a cognitive perspective was investigated in the socio-historical *publication 3* (Ullrich & Liebal, 2018). That publication, serves as supplement to the quantitative *publication 2* (*Scala Naturae*) and *publication 4* ('Intention'). In both cases it provides important background information. It explains, for instance, how ethology and cognitive science mutually interchanged when the cognitive shift took place in the course of the 20th century. The cognitive shift directed the behaviourist focus on *language* away from external, structural properties, and instead towards more inner mental (psychological) aspects.

The use of the term *intention* is an example for the shifting epistemological perspective. The change of perspective was not a result of a 'scientific revolution'. As *publication 4* (Ullrich, Mittelbach, & Liebal, submitted) shows, the introduction of the cognitive term *intention* provoked surprisingly little objection. To validate that claim a corpus was investigated, consisting of 653 peer-reviewed journal articles from the years 1948 to 2017. All 653 articles use the term 'intention'. The further investigation of that corpus shows that articles use the term increasingly more often over time. Interestingly, experiments empirically validating the phenomena of intentionality are lacking until the 1990s (**Fig. 3**, *publication 4*, p.93). From the meta-analysis it appears that most researchers do not challenge the original idea that non-human animals have intentions. Rather the debate is on questions to what order intentions exist outside the human species.

A main focus of *publication 4* is the analysis of all cited references within the 653 articles. The idea was to visualise which authors and articles are the most prominent. The identification of hub articles and prominent authors allowed for further socio-historical investigations in *publication 3*. Personal influences and the academic background were investigated for the most prominent authors of the corpus. As shown by the analysis in *publication 3* & *4* the term *intention* was introduced from two *schools of thought* independently (behaviourism and cognitive science), but received influences from both research traditions. Exactly that exchange of ideas between the different *schools of thought* makes it necessary to reflect the different uses and origins of *intention* to prevent confusions in further discussions.

In a last step, the categorisation of all corpus articles according to their focus on primate or non-primate research could reveal a pronounced primate-bias for the past discourse. One possible explanation could lie in the norm of directed progress, which refers back to the impacts of Scala Naturae- thinking. Constructions of language are ever changing, but some hidden assumptions (e.g. the value of directed progress) are identified more frequently and appear more persistent over time than others (e.g. oral norm).

Publications

Publication 1: From 'speech' to 'gesture'

Ullrich, R. (2016). From 'speech' to 'gesture': The 'oral' as norm in 'language' research. In G. Hartung & M. Herrgen (Eds.), Interdisziplinäre Anthropologie (pp. 179–208). Wiesbaden: Springer. https://doi.org/10.1007/978-3-658-14264-3_15

Interdisziplinäre Anthropologie, 2016, https://doi.org/10.1007/978-3-658-14264-3_15

From ‘speech’ to ‘gesture’: The ‘oral’ as norm in ‘language’ research

Robert Ullrich

Freie Universität Berlin, Comparative Developmental Psychology

Abstract:

The term ‘language’ is used ambiguously by scientists. As a consequence, up until now no attempt to define ‘language’ as a clear-cut ‘faculty’ remains uncontested. This text investigates the term ‘language’ as a putative social construct, based on social norms. Here it is proposed that the existence of a specific social norm – the oral norm – led scientists to the idea that one aspect of ‘language’ could embrace the whole concept. Until the middle of the 20th century, an overly narrow construction of ‘language’ delimited the ascription of the term to certain populations within the human species itself. For instance, deaf people’s use of non-oral communication was considered insufficient in constituting ‘language’. The present study aims to track the form and function of the oral norm historically and its aftermaths in recent scientific discourse. A comparative approach applies the findings of this examination to current research of animal communication. As a result, it will be shown how the oral norm of the past and its remaining manifestations played and play a part in contributing to the construction of a concept of ‘language’ that is unique to humans.

Interdisziplinäre Anthropologie, 2016, https://doi.org/10.1007/978-3-658-14264-3_15

1. Introduction

'Language': A plain and an ambiguous definition.

To avoid misunderstandings from the onset, it seems worthwhile clarifying the use of vocabulary in this text. The term 'language' is used in a twofold sense: (i) 'A language' – a noun with an article and a plural form - denotes a system by which individuals communicate (e.g. 'Russian' or 'American Sign Language' (ASL)). The use of this term is not controversial, but well-defined[4]. (ii) By contrast, 'language' – a mass noun without article and plural - is generally used to make assertions about 'language' in general. For most researchers, 'language' in general is a type of behaviour[5]. But it is behaviour, so the agreement goes, of a very particular sort. Thus it is often supplemented by attributions like: "cornerstone"[6], "most complex"[7], "extraordinary"[8], or "without parallel"[9]. 'Language' stands out, because it appears as *the* defining feature that grants humans the competence for "cultural change"[10], "cooperation about common future goals"[11], "abstract thought"[12], "explicit norms and allocation rules"[13], or even being "the dominant species on the planet today"[14]. After all, there exists no consensus about what specifically identifies 'language' in general[15]. Some scientists use the mass noun 'language' to refer to signal coding that enables communication[16], while others refer to the underlying cognitive abilities which make communication as behaviour possible[17]. Still others claim that the use of the term only makes sense in metalinguistic discourse[18]. This brief overview makes it clear that there is no common-sense use of the term 'language' within the scientific discourse. That is why the term is set in inverted commas throughout the paper. Given the ambiguous use of terminology, there are in fact scientists who doubt the existence of an entity labelled 'language'[19]. Noam Chomsky labelled this position the 'nonexistence approach'[20]. For its exponents, 'language' is first of all a concept; hence a result of normative construction[21]. This approach will be

[4] Simpson: 'Language', 1994, p. 1894.
[5] e.g. Zuberbühler: 'Linguistic Capacity of Non-Human Animals', 2015, p. 315.
[6] Berwick et al: 'Evolution, Brain, and the Nature of Language', 2013.
[7] Gómez, Gerken: 'Infant Artificial Language Learning and Language Acquisition.', 2000.
[8] Tattersall: 'A Putative Role for Language in the Origin of Human Consciousness', 2010, p. 193.
[9] Fisher, Marcus: 'The Eloquent Ape', 2006, p. 93.
[10] Maynard-Smith, Harper: 'Animal Signals', 2003, p. 130.
[11] Gärdenfors: 'The Role of Cooperation in the Evolution of Protolanguage and Language', 2013, p. 193.
[12] Bickerton: Adam's Tongue, 2009, p. 5.
[13] Smith: 'Communication and Collective Action', 2010, p. 242.
[14] Fitch: The Evolution of Language, 2010, p. 1.
[15] Simpson: 'Language', 1994, p. 1894.
[16] Jackendoff: 'What Is the Human Language Faculty?', 2011; Lieberman: 'Language Did Not Spring Forth 100,000 Years Ago.', 2015; Zuberbühler: 'Primate Communication', 2013.
[17] Christiansen, Chater: 'The Language Faculty That Wasn't, 2015; Scott-Phillips, Kirby: 'Information, Influence and Inference in Language Evolution', 2013; Tomasello: Origins of Human Communication, 2008; Sperber, Origgi: 'A Pragmatic Perspective on the Evolution of Language', 2010.
[18] Taylor: 'Folk-Linguistic Fictions and the Explananda of the Language Sciences', 2015.
[19] Taylor: 'The Origin of Language: Why It Never Happened.', 1997.
[20] Chomsky: 'Language and Other Cognitive Systems', 2011, p. 264.
[21] Harris: The Language Myth, 1981; Love: 'Cognition and the Language Myth', 2004.

used as a starting point for the current text. It receives support from studies investigating the role of value-driven social norms in science in general[22]. Various researchers have investigated to which extent unuttered background assumptions construct narratives in empirical research[23]. While a number of publications report biases[24], attitudes[25], values[26], ideologies[27] and paradigms[28] of 'language' research, the discourse analysis employed here aims to subsume those approaches under the heading of social norms in science, which are hereafter referred to as 'scientific norms'. Scientific norms differ from the above-mentioned alternative concepts by their form and function; therefore, a theory of scientific norms might depict more precisely actual scientific practice.

The form and function a scientific norm.

In contrast to values, biases and attitudes, which encompass only partial aspects, scientific norms are sufficiently characterised by five criteria:

(i) Scientific norms are implicitly, rarely explicitly, shared by a group of scientists[29]. (ii) They are followed by normative attitudes, which imply that one can distinguish something as right/appropriate or wrong/inappropriate[30]. (iii) They exist independently of one's own desire, since a significant proportion of a group expects or enjoins group members to conform to the scientific norm[31]. (iv) The violation of a scientific norm may be followed by sanctions. (v) The core function of a scientific norm is to express shared values of what is commonly approved[32]. Such knowledge creates the group's self-concept[33], where the element of normativity can constitute hierarchical, exclusive, and unjust relationships between this group and another[34]. Consequently, the social functions of a scientific norm range from maintaining group cohesion[35] and social order[36] to promoting cooperative behaviour[37].

[22] Mulkay: 'Norms and Ideology in Science', 1976; Allchin: 'Values in Science', 1999; Longino: Science as Social Knowledge, 1990.
[23] e.g. the role of 'Chimpocentrism' to (re-)construct human evolution: Vaesen: 'Chimpocentrism and Reconstructions of Human Evolution (a Timely Reminder)', 2014; e.g. the role of the Aristotelian idea of a 'scala naturae' in neuroscientific research: Northcutt: 'Changing Views of Brain Evolution.', 2001; e.g. the role of motives, norms, and values of scientists in general in relationship to empirical results: Fisch: 'Psychology of Science', 1977.
[24] e.g. 'sampling bias': Evans: 'Language Diversity as a Resource for Understanding Cultural Evolution', 2013.
[25] e.g.: Bauman: 'Audism: Exploring the Metaphysics of Oppression.', 2004.
[26] e.g.: Jackendoff: 'Your Theory of Language Evolution Depends on Your Theory of Language', 2010.
[27] Bauman, Briggs: Voices of Modernity. Language Ideologies and the Politics of Inequality, 2003; Woolard, Schiefflein: 'Language Ideology', 1994; Joseph, Taylor: Ideologies of Language, 1990.
[28] Hintikka: 'Paradigms for Language Theory', 1998.
[29] Bicchieri, Muldoon: 'Social Norms', 2014.
[30] Southwood, Eriksson: 'Norms and Conventions', 2011.
[31] Bicchieri: The Grammar of Society, 2006, p. 11.
[32] Southwood, Eriksson: 'Norms and Conventions', 2011, p. 208.
[33] Bicchieri, Muldoon: 'Social Norms', 2014.
[34] Southwood, Eriksson: 'Norms and Conventions', 2011, p. 211.
[35] Tajfel: 'Social Categorization, Social Identity and Social Comparison', 1978.
[36] Elster: The Cement of Society, 1989.
[37] Axelrod: 'An Evolutionary Approach to Norms', 1986.

Scientific norms may have a bad reputation since they are mostly mentioned – as in the present case – when they influence scientific progress with values rather than by empirical facts. Consequently, the ideal scientific practice is often described as value-free[38]. Yet in some cases scientific norms are very conducive to further progress; e.g. by producing alternative possible answers to questions where empirical evidence is scarce. In this case, competing norms and their respective proponents critically examine opposing results and push forward new ideas[39]. Seen in this light, a value-free science is not only impossible, but not even an ideal[40].

The 'oral norm' as an example for a scientific norm.

In order to demonstrate the validity of the opening premise of this text ('language as construction), it must be verified to what extent scientific norms assist in constructing a certain concept of 'language'. Here it is claimed that historically the oral norm has contributed to constructing an oral-oriented concept of 'language' which has led to the marginalisation and non-recognition of non-oral forms of communication. The hypothesis is: The oral norm did not just exclude deaf people's sign languages from recognition as a 'natural language'; it also imposed a restricted framework on the concept of 'language' beyond, that which would be merited by empirical actualities. The analysis of the history of this discourse will highlight some parallels to more current discourse regarding signal coding in non-human animals. By uncovering the aftermaths of the oral norm in current discourse regarding animal communication, this paper aims to show to which extent value-based narratives still help to construct a concept of 'language' as something unique to humans.

2. The structure of the oral norm in recent history.

Norm compliance is often not deliberate, but unreflective.

The mass noun 'language' was, and still is, defined from a broad range of different perspectives. 'The Encyclopaedia Britannica' – to pick a 'layman's position' - defined 'language' as follows (Emphasis added.):

1951: "By language in the widest sense of the word is meant any means of communication between living beings. […] We may distinguish two kinds of language […] *ear-language* and *eye-language*, of which the former is by far the more important."[41]

1962: "Language may be defined as an arbitrary system of *vocal symbols* by means of which human beings, as members of a social group and participants in a culture, interact and communicate."[42]

1977: "Language, the chief means of human communication. As conventionally defined, language consists of *vocal sounds* to which meanings have been assigned by cultural convention; it is often supplemented by various gestures."[43]

[38] Reiss, Sprenger: 'Scientific Objectivity', 2014.
[39] Vinck: The Sociology of Scientific Work, 2010, p. 46.
[40] Douglas: Science, Policy, and the Value-Free Ideal, 2009, p. 87.
[41] Jespersen: 'Language', 1951, p. 696.
[42] Trager: 'Language', 1962, p. 696.
[43] Robins: 'Language', 1977, p. 32.

2002: "Language, a system of conventional *spoken* or *written symbols* by means of which human beings, as members of a social group and participants in its culture, communicate. Language so defined is the peculiar possession of humans."[44]

2015: "Language, a system of conventional *spoken*, *manual*, or *written symbols* by means of which human beings […] express themselves."[45]

What is evident in those definitions is a change over time from a purely oral picture of 'language' towards the openers for other modalities. The fact that the oral modality was set as a defining feature was most likely not always registered. Already the very term 'language' stems from the Latin expression 'linguia', which can also be translated as 'tongue'[46]. In his famous article 'The Origin of Speech', Charles Hockett wrote that the "vocal-auditory channel" is the most obvious design feature of 'language', which "appear[s] so trivial that no one looking just at language would bother to note"[47]. Hence, the oral modality was implicitly assumed as a kind of naturally given default modus. Without any critical reflection, 'speech' and 'language' were interpreted as being inextricably linked with each other[48] and responsible for the creation of mind[49]. There was no reflection or discussion about the constitutive role of the oral modality. This might have happened because those defining the concept of 'language' were already immersed in this modality, though this was not true for all human beings. The simple fallacy behind a purely oral picture of 'language' was pointed out by Brenda Brueggemann, who summarised thus: "Language is human; speech is language; therefore deaf people are inhuman."[50]. Although one clearly identifies the fallacy of the statement today, experts from the 19th century did in fact partially internalise it[51]. This is exemplified by Thomas Huxleys ("A man born dumb [as a "result from congenital deafness", R.U.] […] would be capable of few higher intellectual manifestations than an Orang or a Chimpanzee […]."[52]) and Mary McCowen[53] who thought that *verbal* 'language' is "one of the distinguishing characteristics between man and the lower order of animals"[54]. Lewis Dudley from the Clarke School for the Deaf (1866) confirmed that people using a sign language "felt themselves to be less than human"[55]. Referring to the use of a sign language, the Swiss deaf teacher Johan Conrad Amman wrote: "How little do they differ from animals!"[56].

These drastic statements were an outcome of the implicit presumption that 'language' must be oral. With these illustrations one arrives at the first backbone constituting the oral norm: The norm was stated implicitly. As such it escaped critical reflection or even mere detection for a long time. Its

[44] Robins: 'Language', 2002, p. 147.
[45] Crystal: 'Language', 2015.
[46] pointed out by Bauman: 'Audism: Exploring the Metaphysics of Oppression.', 2004.
[47] Hockett: 'The Origin of Speech', 1960, p. 6.
[48] Armstrong, Karchmer: 'William C. Stokoe and the Study of Signed Languages', 2009, p. 389.
[49] DeMatteo: 'Visual Imagery and Visual Analogues in American Sign Language', 1977, p. 109.
[50] Brueggemann: 'Lend Me Your Ear: Rhetorical Constructions of Deafness', 1999, p. 11.
[51] Gessinger: 'Der Ursprung Der Sprache Aus Der Taubstummheit', 1989, p. 347.
[52] Huxley: 'Man's Place in Nature', 1863, p. 121.
[53] founder of the Oral School for deaf individuals in Chicago (1883)
[54] quoted by Baynton: 'Forbidden Signs', 1996, p. 52
[55] quoted by Baynton: 'Forbidden Signs', 1996, p. 52.
[56] Amman: 'A Dissertation on Speech', 1873, p. 2.

omnipresence looms large, as shown in the excerpts from 'The Encyclopaedia Britannica'. The sometimes implicit character of norms directly leads to expectations on the part of scientists. Those expectations constitute the second backbone of the oral norm.

Expectations can be based on collective decisions.

It is expectations and motives that make a collective behaviour a norm[57]. The belief of a scientist that the majority of a scientific community conforms to the conviction that 'language' must be oral is called "empirical expectation"[58]. The belief of a scientist that the majority of a scientific community *expects* to conform to the conviction that 'language' must be oral is called "normative expectation"[59]. The oral norm in 'language' discussions holds true for both. The omnipresence of a purely oral picture of 'language' is a good example of the former. 'The Second International Congress on Education of the Deaf' taking place in 1880 in Milan is a good example for the latter. In this case, 158 educators of deaf students (out of 164, including 1 deaf person) formed a joint commitment concerning the method of instructing deaf students[60]. The collective decision to abandon instructions via a sign language and to replace them by oral education, such as lip reading, is an example of how normative expectations for actual and subsequent teachers of the deaf were installed. After all, the majority of experts followed the decision until the late 1960s[61], in Germany even to the late 1980s[62]. These 'dark ages'[63], as the deaf community dubbed that period, lead structurally to a dramatic deterioration in the quality of deaf pupils' educations[64]. This brings the discussion to the third backbone of a norm in general and the oral norm specifically: sanctions.

Sanctions range from gossip to open censure.

The biography of William Stokoe – one of the first linguists[65] who assumed American Sign Language (ASL) to be a 'natural language' and who carried out a structural analysis of it – offers two generic examples of how sanctions follow the violation of a scientific norm. When Stokoe accepted a job offer for Gallaudet University – the oldest and largest university for deaf students in the United States – he concurrently took a job of a deaf candidate (Bob Panara) who already taught at Gallaudet and had received major support from the deaf students. Clearly, in 1955, oral education was the dominant orientation of Gallaudet University and Stokoe profited from what McDonnell and Saunders describe as "internal strategies against signing"[66]. In the narrow sense of the word, appointment procedures of

[57] Bicchieri: The Grammar of Society, 2006, p. 29.
[58] Bicchieri: The Grammar of Society, 2006, p. 11.
[59] Bicchieri: The Grammar of Society, 2006, p. 11.
[60] Van Cleve: A Place of Their Own, 1989, p. 108 ff.
[61] McDonnell, Saunders: 'Sit on Your Hands', 1993.
[62] Mally: 'The Long Road to Self-Confidence of the Deaf in Germany', 1993.
[63] Chough, Dobyns: 'How Is Asian Deaf Culture Different from American Deaf Culture', 2006, p. 227.
[64] Sacks: Seeing Voices, 1990, p. 28.
[65] Bernard T. Tervoort was similar intentioned. He also concluded that manual communication can be described as behaviour in a linguistic sense; see: Tervoort: Developmental Features of Visual Communication, 1975.
[66] McDonnell and Saunders: 'Sit on Your Hands', 1993, p. 259.

this kind were sanctions. Sanctions could entail seating students with their hands restrained in order to prevent signing[67], or to fill a position with staff conforming to the oral orientation of the school. Bob Panara retrospectively described the situation when he was replaced as chairman by William Stokoe thus: "we were much like the Negro at that time"[68]. The statement indicates the existence of various sanctions that existed and created a negative image of the use of signs in general. The attempt to suppress sign languages "created a stigma […] and a negative, guilty attitude about its use."[69] "Even educated deaf people were ashamed of it though among themselves, and in secret, they signed."[70].

After Stokoe had observed Bob Panara's and various deaf pupils' use of manual signs, he proposed to systematically investigate the structure of ASL. In 1960, he published his first results, claiming that sign languages were fully-fledged natural languages containing syntax, morphology, and grammar[71]. Sanctions to this violation of the oral norm (that did not allow for a non-oral 'language' definition) followed quickly. Maher writes: "[…] his paper was like that of Martin Luther's Ninety-five Theses"[72]. Gilbert Eastman, who worked at Gallaudet, recalls that "my colleagues and I laughed at Dr. Stokoe and his crazy project"[73]. Lou Fant, also a teacher at Gallaudet at that time, added, "oralism was the dominant philosophy of education. […] The quickest path to becoming a nonentity was to downgrade oralism."[74]. Accordingly, Stokoe "was reviled and ridiculed"[75]. Few people found his work interesting or devoted much attention to it[76] and people who read it (deaf people included) reacted with "much resistance"[77]. These sanctions were a result of what people saw as inappropriate research conclusions. They were a consequence of the violation of shared values, which in fact constitute the fourth backbone of the oral norm.

Norms can be valued.

Norms are accompanied by normative attitudes, which allow evaluating a behaviour or concept as 'right'/'appropriate' or 'wrong'/'inappropriate'[78]. Deaf people were victims of normative attitudes for most of history of their institutional education. To be sure, mediated through personal experience, early educators of the deaf like George W. Veditz (1890: "to make the eye take the place of the ear […] is a violation of the laws of nature"[79]), Edward M. Gallaudet (1898: "gestural expression is in no

[67] Simms: 'Deaf Education: Whose "Way" Is It?', 2006, p. 84.
[68] Maher: Seeing Language in Sign, 1996, p. 49.
[69] Gannon: Deaf Heritage, 2012, p. 362.
[70] Maher: Seeing Language in Sign, 1996, p. 55.
[71] Stokoe: 'Sign Language Structure, [1960], 2002.
[72] Maher: Seeing Language in Sign, 1996, p. 71.
[73] Maher: Seeing Language in Sign, 1996, p. 71.
[74] Maher: Seeing Language in Sign, 1996, p. 75.
[75] Padden, Humphries: 'Inside Deaf Culture', 2006, p. 125.
[76] Gannon: Deaf Heritage, 2012, p. 365.
[77] Nomeland, Nomeland: The Deaf Community in America, 2012, p. 107.
[78] Southwood and Eriksson: 'Norms and Conventions', 2011, p. 199.
[79] Veditz: 'Notices of Publications: The Deaf-Mute and His Language.', 1890, p. 272.

respect inferior"[80]), Wilhelm Wundt (1904: "sign languages are independent and arise spontaneously"[81]), and Arnold Hill Payne (1911: "Speech […] is only one way of expressing […] language"[82]) presented alternative accounts for the use of signs years earlier. However, those scientists did not share the values of the majority. The majority of scientists were able to hear and had barely any contact with deaf persons. Their values were defined from the perspective of members of hearing society. Here follow a few statements from more recent history which, though representing untested claims, exercised a great impact on the prevailing scientific discourse:

1873: "[…] how inadequate and defective is the language of gestures and signs which they must use!"[83]
1933: "Even when gestures are symbolic, they go little beyond the obvious […] Some communities have a *gesture language* […] It seems certain that these gesture languages are merely developments of ordinary gestures […] language always ran ahead of gesture."[84]
1947: "How tragic it is to see a group of deaf adults talking to each other by grotesque and inefficient finger acrobatics!"[85]
1950: "Beside speech there is no other generally used universal sign system. […] Other sign systems, like the deaf-mute language […] are either transposed, restricted or parasitic."[86]
1958: "The sign language, like writing, is a substitute for speech, not an independent or original method of communication."[87]
1961: "Speech is the highest and most developed of all forms of communication, alongside which are to be found, even in humans, more rudimentary language systems based on gesture, sign, and acoustic signal."[88]
1964: "Ideographic language systems, in comparison with verbal symbol systems, lack precision, subtlety, and flexibility. […] Comparatively, a verbal language is more abstract."[89]
1970: "It is generally agreed that sign language is bound to the concrete and is rather limited with respect to abstraction, humour, and subtleties such as figures of speech with rich expression"[90]

It is not surprising that the values related to the oral norm also constitute their own empirical evidence. From today's perspective, it seems odd what kinds of attributes were ascribed to sign languages in the late 19th and early 20th century: Garrick Mallery observed that, for instance, "an English deaf-mute had no difficulty in conversing with Laplanders"[91]. He concluded that there exists only one universal sign language. Hans Furth reviewed 39 studies comparing the cognitive

[80] Gallaudet: The Deaf and Their Possibilities, 1898, p. 211.
[81] Wundt: Völkerpsychologie. Erster Band. Die Sprache, 1904, p. 138.
[82] "If we apply the same test to English that is applied to signs by those who would rule out any which they suppose cannot come under the head of natural gesture or pantomime, what fraction of our so-called natural language should we have left?"; Payne: 'Deaf and Dumb', 1911, pp. 884–885.
[83] Amman: A Dissertation on Speech, 1873, p. 2.
[84] Bloomfield: Language [1933], 1984, pp. 39 – 40 [emphasis in the original].
[85] Goodhill: 'The Educational Treatment of the Pre-School Deaf Child.', 1947, p. 559.
[86] Lotz: 'Speech and Language', 1950, p. 712.
[87] Kroeber: 'Sign Language Inquiry', 1958, p. 13.
[88] Kainz: Die 'Sprache' der Tiere, 1961, p. 3. (translated by RU. Original reading: „Die Lautsprache ist das höchste und vollendetste aller Kommunikationsmittel, neben dem sogar noch beim Menschen unvollkommenere Systeme dieser Art (Gebärden-, Zeichen-, Ton- und Schallsignalsprachen) stehen.")
[89] Myklebust: The Psychology of Deafness., 1964, p. 241.
[90] Davis, Silverman: Hearing and Deafness, 1970, p. 390; cited in: Battison, 'American Sign Language Linguistics 19701980, 2000, p. 6.
[91] Mallery: Sign Language among North American Indians, 1881, p. 307.

performance of deaf students to that of hearing ones[92]. He found that deaf students are "linguistically deficient"[93]. If there was a study showing that deaf students outperformed the hearing group, Furth hypothesised, it would be "because of their less sophisticated approach"[94]. When hearing experts did research on a sign language used by deaf people, they usually translated expressions 'sign-by-sign' into English language. The result read as the early attempts of machine translation. Hence, they concluded that the signs of deaf people would lack syntax, meaning, complex grammar, and would suffer a paucity of vocabulary[95]. All these putative shortcomings would ultimately lead to a deficiency in abstraction[96] and would result in the development of behaviourally impulsive, immature beings[97]. This short overview simply aims to show the manner in which the oral norm constituted its own evidence. As long as the oral was set as the default modus, terms and concepts were defined according to it. The modality-independent nature of 'language' was ignored or not even perceived. The oral-associated structure (i.e. syntax, morphology) was considered the core of 'language'. Gesture, expression, and context were marginalised as supplementary to it. Here it is necessary to turn to the core function of norms. The question arises: What are norms good for and why they are impossible to remove from science?

Norms express shared values of what is commonly approved.

Norms in science serve the same functions they serve in other parts of human life. They coordinate expectations, facilitate cooperation, create and galvanise meanings and identities, generate accountability[98], constitute repressive, hierarchical, exclusive, or unjust relationships; in short, they express shared values[99]. The human preference for similar, rather than dissimilar, interlocutors has been widely investigated under the heading of homophilic behaviour[100]. Investigations of homophilic behaviour include prominent lines of research such as in-group favouritism[101] and biases to conformity[102]. The results show that people who belong to salient groups tend to evaluate the attributes of their own (in-) group more positively than those of other (out-) groups[103]. One of the driving forces behind this seems to be the well-studied human tendency to homophilic social

[92] Furth: 'Linguistic Deficiency and Thinking: Research with Deaf Subjects 1964-1969', 1971.
[93] cf. Bornstein, Roy: 'Comment on "Linguistic Deficiency and Thinking: Research with Deaf Subjects 1964-1969"', 1973.
[94] Furth: 'The Influence of Language on the Development of Concept Formation in Deaf Children', 1961, p. 388.
[95] Schlesinger, Meadow: Sound and Sign, 1982, p. 39; Klima, Bellugi: The Signs of Language, 1979, p. vii.
[96] Myklebust, Brutten: 'A Study of the Visual Perception of Deaf Children', 1953.
[97] Schlesinger, Meadow: Sound and Sign, 1982, p. 2; Lane: 'The Medicalization of Cultural Deafness in Historical Perspective', 1993, p. 484.
[98] Elster: The Cement of Society, 1989, p. 97.
[99] Southwood, Eriksson: 'Norms and Conventions', 2011, p. 211.
[100] Haun, Over: 'Like Me', 2013.
[101] Tajfel et al: 'Social Categorization and Intergroup Behaviour', 1971.
[102] Levine: 'Solomon Asch's Legacy for Group Research.', 1999.
[103] Fu et al: 'Evolution of in-Group Favoritism', 2012.

preference[104]. Such preference appears ubiquitous in all human and some non-human populations[105]. The psychologist Harlan Lane wrote, with specific reference to deaf culture and the use of sign languages, "on the face of it, people are quite afraid of [human] diversity and look to social institutions to limit or eradicate it."[106]. The institutionalised oral education of deaf students is an example of one such 'social institution'. Scientific practice itself, with its highly ritualised research procedures, is another one. Given the concrete example discussed here, there is no need for explicit awareness of the creation of such a scientific norm by any concrete scientist. It suffices to be able to use and hear an oral form of a language to form an in-group of hearing experts, who prefer similar attributes to those they recognise in themselves. In other words, scientific norms – such as the oral norm - express what is commonly approved in a certain (in-)group. On this reading, a deviant modality becomes denied[107], marginalised[108], or appraised[109]. Statements that are based on in-group favouritism also explain why hearing experts ascribed complexity, flexibility, precision, independence, perfection, efficiency, development, and so on to oral, but not manual modalities of 'language'. By virtue of the defining power of the majority of the group, social order is maintained. This is the case because it is the deviant part of society that has to change its behaviour. While in many cases the existence of norms makes societies or even scientific practice function smoothly, in this particular case the oral norm had a negative impact on the education and lives of deaf people. As will be shown in the next chapter, the construction of a broader concept of 'language' alleviated the situation.

3. Widening the lens: When communication becomes 'language'.

Lou Fant, a pioneering expert on ASL, once wrote: "I had signed ASL since infancy, but I had never thought of it as a language, it was just a way to communicate."[110] Taking this statement seriously, one may wonder what distinguishes 'a way to communicate' from 'a language', and, as a consequence thereof, from 'language' in general. That much is clear that the changing status of a sign language emerged from detailed structural research into ASL, initiated by William Stokoe. This change was confirmed and extended by various influential linguistic studies[111]. It becomes apparent, however, that some of the earlier, more structural analyses of sign languages defined 'language' merely as a unimodal, linearly structured grammatical code. Gradually Stokoe's successors softened this view of

[104] Haun, Over: 'Like Me', 2013.
[105] Haun, Rekers, Tomasello: 'Majority-Biased Transmission in Chimpanzees and Human Children, but Not Orangutans', 2012; Fu et al: 'Evolution of in-Group Favoritism', 2012; Campbell, de Waal: 'Chimpanzees Empathize with Group Mates and Humans, but Not with Baboons or Unfamiliar Chimpanzees.', 2014.
[106] Lane: When the Mind Hears, 1984, p. xiii.
[107] e.g. 'if auditory receptive language cannot be established [...], then acquisition of all of the succeeding language functions will be impeded'; Myklebust: The Psychology of Deafness., 1964, p. 233.
[108] e.g. 'the deaf child is reduced to gestures to indicate his wants'; Whetnall, Fry: The Deaf Child, 1964, p. 1.
[109] e.g. 'Speech is the highest and most developed of all forms of communication'; Kainz: Die 'Sprache' Der Tiere, 1961, p. 3.
[110] Fant: 'Two Memorable Meals with Ursula and Ed', 2000, p. 4.
[111] Klima, Bellugi: The Signs of Language, 1979; Sandler, Lillo-Martin: Sign Language and Linguistic Universals, 2006; Pfau, Steinbach, Woll: Sign Language. An International Handbook, 2012.

'language'. Instead of transferring the 'linear' oral linguistic theory to a sign language, major empirical studies found 'simultaneity of structure' at every level of phonology, morphology, prosody, and discourse[112]. The oral modality of languages was identified as an empirical filter that had allowed only a subset of languages to become subsumed to the concept of 'language'[113]. The generative grammar theory has been criticised as too narrow[114], the separation of speech and gesture's grammar as too artificial[115]. Unimodal views that regarded speech as the most developed manifestation of 'language', with gesture and other non-oral forms of communication serving a mere scaffold function[116] became replaced in favour of more multimodal (i.e. involving more than one communicative modality at a time) or modality-independent accounts of 'language'[117]. Among others, Nick Enfield noted that meaning in 'language' does not originate in a certain linguistic structure, but is rather a result of intention, composition, and context[118]. While the structural properties of 'established' sign languages such as German, French, or American Sign Language are influenced by the social circumstances of their communities, researchers began to investigate 'village signs' to control for those influences[119]. Researchers realised that putative linguistic universals such as 'duality of patterning'[120] cannot be found everywhere[121]. Furthermore, none of the village sign languages showed inflectional morphology, some do not use syntactical subordination, and still others show an unexpected high degree of lexical variation[122]. Consequently, current research interprets more and more linguistic forms and structures as 'cultural add-ons', rather than biologically grounded[123]. In the face of a shifting oral norm, research into 'language' has become more independent of form and the focus has moved towards pragmatics, i.e. function and context[124]. Altogether there is a growing awareness of and allowance for more diversity in terms of what constitutes 'language'[125]. The impression is created that the oral norm could be forgone, because it no longer provided a suitable explanation for human uniqueness. What remains is a long-standing discourse that defines human exceptionalism by contrasting humans with non-human animals. The shift helped to loosen the long-held fixation on modality or structure in defining 'language'. To make sense of this point, one must

[112] Aronoff, Meir, Sandler: 'The Paradox of Sign Language Morphology', 2005, p. 314.
[113] Meier: 'Language and Modality', 2012, p. 575.
[114] Liddell: Grammar, Gesture, and Meaning in American Sign Language, 2003, p. 71 ff.
[115] Fricke: 'Towards a Unified Grammar of Gesture and Speech', 2013.
[116] as can be found in Myklebust: The Psychology of Deafness., 1964, p. 232.
[117] Slocombe, Waller, Liebal: 'The Language Void', 2011; Perniss, Özyürek, Morgan: 'The Influence of the Visual Modality on Language Structure and Conventionalization' ,2015; McNeill: Gesture and Thought, 2005; Levinson, Holler: 'The Origin of Human Multi-Modal Communication', 2014; Vigliocco, Perniss, Vinson: 'Language as a Multimodal Phenomenon', 2014.
[118] Enfield: The Anatomy of Meaning, 2009.
[119] Meir et al: 'The Influence of Community on Language Structure', 2012.
[120] Hockett: 'The Origin of Speech', 1960.
[121] Sandler et al: 'The Gradual Emergence of Phonological Form in a New Language', 2011; Levinson, Dediu: 'The Interplay of Genetic and Cultural Factors in Ongoing Language Evolution', 2013, p. 222.
[122] Meir et al: 'The Influence of Community on Language Structure', 2012.
[123] Wray, Grace: 'The Consequences of Talking to Strangers, 2007.
[124] e.g. Scott-Phillips: 'Nonhuman Primate Communication, Pragmatics, and the Origins of Language', 2015.
[125] Evans, Levinson: 'The Myth of Language Universals', 2009.

look at how far the scientific discourse proceeds with non-human species when the term ‘language’ is in use in current research.

4. Familiar terms, familiar discourse: ‘Language’ in non-human animals.

The closing chapter aims to show the following: the use of the term ‘language’ in the scientific discourse on non-human animals (henceforth ‘animals’) mirrors the developments that have been described so far regarding the maintenance of the oral norm towards deaf people. Note, that this is not an attempt to compare human ‘language’ to animal communication systems. It is in fact scientists themselves who draw a link between animal communication and human ‘language’.

The legacy of the oral norm is still observable in animal communication discourse today. For the sake of brevity the main focus will be on primate communication. In the early 20th century researchers were fascinated by the idea of teaching human ‘language’ to great apes, to show ‘language’ equivalents in these species closely related to humans. The initial motivation to show this was to challenge the idea that there is no evolutionary continuity between animal communication and human ‘language’[126]. Two strategies of teaching were implemented: an oral in the early 20th century and a non-oral strategy, in the 1960s and 70s. After the first methodologically unsound attempt of Richard Garner to teach oral ‘language’ to an ape[127] at least five methodologically refined studies tried to do the same, all with little success[128]. Additionally, in 1925 Robert Yerkes suggested the teaching of sign or symbolic ‘language’ to apes[129] – a suggestion independently supported by Lev Vygotsky[130]. However, it was only after William Stokoe published his structural analysis of ASL that the first researchers implemented these ideas. Referring to Stokoe’s work, Gardner and Gardner claimed to have taught conventional signs of ASL to a chimpanzee[131]. This, as well as another non-oral study conducted by Premack[132] gave rise to an influential, yet already familiar idea: the gestural origin of ‘language’. Gordon Hewes who is cited frequently as the initiator of this idea refers in his paper[133] to the long historical tradition of the gestural origin, leading back to the 18th and 19th century. However, it was only after ASL was acknowledged as a natural ‘language’, and only after the failure of oral and alleged success of manual experiments, that the gestural origin was debated again. Since then, two contrary positions dominated the ‘language origin’ discourse: the oral and the manual[134] position. While the oral position could find complexity and flexibility in oral utterances only[135] referring to

[126] Radick, The Simian Tongue, 2007, p.16 & 107
[127] Radick, The Simian Tongue, 2007, p.233-234.
[128] Kellogg, ‘Communication and Language in the Home-Raised Chimpanzee’, 1968.
[129] Yerkes, Almost Human, 1925, p.180
[130] Vygotsky, 1986 [1934], Thought and Language, p.75
[131] Gardner, Gardner, ‘Teaching Sign Language to a Chimpanzee’, 1969.
[132] Premack, ‘Language in Chimpanzee?’, 1971
[133] Hewes, ‘Primate Communication and the Gestural Origin of Language’, 1973.
[134] Note; in order to keep cogency of writing style (previous chapters) and to highlight parallels in the line of argument I prefer the technical terms ‘oral’ vs ‘manual’ to ‘vocal’ vs ‘gestural’.
[135] “[…] it is clear that any wild chimpanzee who spent a lot of time doing deaf sign […] would soon be a very dead chimpanzee.” Hill, ‘On the Evolutionary Foundations of Language’, 1972, p.311.

proper vocal learners such as songbirds[136], the manual position downgraded primate vocalisations as rather emotional and restricted, speaking lightly of the parrot's and pigeon's general learning abilities, denoting their vocalisations as a "clearly negative" example of 'language'[137].

However, the optimistic results of the Gardners did not survive for long. In 1979 Herbert Terrace expressed fundamental criticisms to Premack, Gardner and Rumbaugh, about the applied methods and interpretation of the results[138]. Following the publication of this paper, there was a decline in the number of experimental procedures that attempted to train animals for human-like 'language'. During the second half of the 20th century observations of animals communicating in their natural environment became more influential. Again, initial studies focused on the oral modality of communication, such as the alarm call system of vervet monkeys[139]. This study gave rise to a novel field of research that examined this phenomenon in various animal species as well as initiating a debate about the role of functionally referential calls[140]. At the same time, only few studies concerned with non-oral channels of communication were published. A metastudy analysed 553 primate communication studies from 1960 to 2008 and found that the majority (64%) focussed on oral communication[141]. Only 22% looked at facial, 9% at gestural and 5% at multimodal communication. 89% of all studies conducted in the primate's natural habitats focussed on the oral domain. Furthermore, the majority of oral studies used experimental methods (62%) while observational methods were favoured in the other modalities.

Together, this suggests that when considering research aiming at identifying the origin of human 'language' by studying primates, researchers traditionally focus on the oral modality of communication. The simple reason for this is the assumption that 'language' must be something oral. Only if no similarities between humans and primates are found in the oral domain, researchers will start to explore other communicative modalities. It is part of an ongoing scientific debate to question the role of modalities involved in 'language' evolution. To date researchers interested in this evolution no longer blindly follow a pre-empirical oral norm. However, there are some tendencies in current research that can be explained only by reference to the history of the discourse, and seem to emerge as aftermaths of the oral norm. First, as already suggested, there is an ongoing general research bias towards oral communication[142]. This focus might also explain the vast interest in songbird communication. Researchers involved in this field continue to stress the similarities between birdsong and 'language'[143]. This observation by no means implies that birdsong studies are of no use

[136] Nottebohm, 'The Origins of Vocal Learning', 1972, p.133.
[137] Premack, 'Language in Chimpanzee?', 1971, p.808; Premack, 'The Codes of Man and Beasts', 1983, p. 133.
[138] Terrace et al, 'Can an Ape Create a Sentence?', 1979; for a more recent review see: Rivas, 'Recent Use of Signs by Chimpanzees (Pan Troglodytes) in Interactions with Humans', 2005.
[139] Seyfarth, Cheney, Marler, 'Monkey Responses to Three Different Alarm Calls', 1980.
[140] Price et al, 'Vervets Revisited', 2015, p. 2 and the references therein.
[141] Slocombe, Waller, Liebal, 'The Language Void', 2011.
[142] Slocombe, Waller, Liebal, 'The Language Void', 2011.
[143] Bolhuis, Okanoya, Scharff, 'Twitter Evolution', 2010.

when attempting to unfold presumed roots of human 'language'. However, it demonstrates why researchers often see similarities between courtship song and oral 'language' but rarely draw similar conclusions from courtship dance or any combination of modalities. Second, most studies about animal communication do not acknowledge that all animals communicate with their entire body. The present bias for unimodal research can also be explained by reference to the historical oral norm, which implies that something called 'language' might be fully realised in a single modality. The multimodal approach of 'language' is a relatively contemporary concept (see previous chapter, fn 117). Until 2008 only 28 out of 533 primate communication studies focussed on this specifically[144]. Third, research advocating an oral origin of 'language' to date, sometimes uses similar explanatory models as can be found in the sign languages discourse in the mid-20th century. Klaus Zuberbühler writes:

> *"Language is mainly a vocal behaviour. Of course, it is true that rudimentary language-like gestural systems have emerged in deaf populations, but this is not the default pattern in normally developing humans."*[145]

Given the analysis of the previous chapters this statement might remind the reader of earlier citations. Despite this formulation, Zuberbühler does not set the oral trait as constitutive for 'language'. But he, and some of his colleagues[146], seem convinced that oral modality plays a 'default' role in human 'language' and likewise animal communication. What makes this statement interesting is that it reveals its normative foundations. To date there is no empirical evidence for such a 'default pattern' of 'language'. The statement is not as descriptive as it might seem. A certain species may communicate via oral, seismic, or visual signals. Whether it uses any combination of the three, is not a result of any human or animal proficiency, but rather, among other things, it is a function of the environment[147]. As there is no 'default environment' for any organism and no 'default development' in humans[148], there cannot be a 'default modality' of 'language'. For the current case, oral communication is more prominent, relative to its marginalised alternatives, and might be identified as a legacy of the oral norm.

However, an empirical aim leads the current discourse of 'language' origin to describe similarities and dissimilarities between human and animal communicational systems. Scientific norms, such as the oral norm, no longer hinder a broad empirical sampling, nor do they produce 'wrong' results. Empirical facts can never be 'wrong', but their interpretation can lead to a misguided conception of

[144] Slocombe, Waller, Liebal, 'The Language Void', 2011.
[145] Zuberbühler: 'Linguistic Capacity of Non-Human Animals', 2015, p. 315. Emphasis added.
[146] e.g. (i) "speech is the biological default modality": Brainard, Fitch: 'Editorial Overview: Animal Communication and Human Language', 2014, p. v; (ii) in a foreword to a special issue titled 'Language and Birdsong' one can read: "[…] one essential characteristic of human language is that it is based on a capacity for vocal learning": Brenowitz, Perkel, Osterhout: 'Language and Birdsong', 2010, p.2; (iii) a review of vocal communication as relevant to earliest stages of 'language' evolution: Cheney, Seyfarth: 'Constraints and Preadaptations in the Earliest Stages of Language Evolution', 2005.
[147] Candolin: 'The Use of Multiple Cues in Mate Choice', 2003.
[148] see Lane: 'Construction of Deafness', 1997.

'language'. In fact it is scientific narrative that is influenced by scientific norms. Conclusions such as: "Human language is a vocal behaviour, so a natural focus has been the study of non-human primate vocal behaviour."[149] construct a primarily oral narrative of 'language'. Thereby it perpetuates the narrative of human uniqueness, since "humans are enormously vocal primates, especially when compared with their nearest primate relatives"[150]. This is the point where shared values constitute an identity, by setting humans apart from even closely related animals based on features that are chosen and constructed by the researchers themselves. Research narratives can be found in every academic publication. They have a significant effect on the interpretation and exploration of empirical facts. Hence, the aim of the current paper is to unveil a specific scientific norm as an unuttered background assumption allowing the reader of scientific research to form an opinion about empirical facts. As the term 'language' is used from several scientific perspectives, it is worth continuing this line of inquiry by unveiling more normative background assumptions in this domain. Bringing to light the scaffold upholding a concept called 'language' might assist future empirical research in clarifying its implications.

[149] Zuberbühler: 'Linguistic Capacity of Non-Human Animals', 2015, p. 318.
[150] Zuberbühler: 'Linguistic Capacity of Non-Human Animals', 2015, p. 315.

Literature

Allchin, Douglas, 'Values in Science: An Educational Perspective', *Science Education*, 8 (1999), 1–12

Amman, Conrad J., *A Dissertation on Speech* (London: S. Low, Marston, Low and Searle, 1873)

Armstrong, David / Karchmer, Michael, 'William C. Stokoe and the Study of Signed Languages', *Sign Language Studies*, 9 (2009), 389–397

Aronoff, Mark / Meir, Irit / Sandler, Wendy, 'The Paradox of Sign Language Morphology', *Language*, 81 (2005), 301–344

Axelrod, Robert, 'An Evolutionary Approach to Norms', *The American Political Science Review*, 80 (1986), 1095–1111

Battison, Robbin M., 'American Sign Language Linguistics 19701980: Memoir of a Renaissance', in *The Signs of Language Revisited*, ed. by Emmorey, Karen and Lane, Harlan (London: Lawrence Erlbaum Associates, Inc., 2000), pp. 5 – 16

Bauman, H-Dirksen, 'Audism: Exploring the Metaphysics of Oppression.', *Journal of deaf studies and deaf education*, 9 (2004), 239–46

Bauman, Richard / Briggs, Charles L., *Voices of Modernity. Language Ideologies and the Politics of Inequality.* (Cambridge, New York: Cambridge University Press, 2003)

Baynton, Douglas, *Forbidden Signs: American Culture and the Campaign against Sign Language* (Chicago: University of Chicago Press, 1996)

Berwick, Robert / Friederici, Angela / Chomsky, Noam / Bolhuis, Johan, 'Evolution, Brain, and the Nature of Language', *Trends in Cognitive Sciences*, 17 (2013), 98

Bicchieri, Cristina, *The Grammar of Society.* (New York: Cambridge University Press, 2006)

Bicchieri, Cristina / Ryan Muldoon, 'Social Norms', *The Stanford Encyclopedia of Philosophy* (Edward N. Zalta (ed.), 2014)

Bickerton, Derek, *Adam's Tongue.* (New York: Hill and Wang, 2009)

Bloomfield, Leonard, *Language [1933]* (Chicago, London: The University of Chicago Press, 1984)

Bolhuis, Johan / Okanoya, Kazuo / Scharff, Constance, 'Twitter Evolution: Converging Mechanisms in Birdsong and Human Speech.', *Nature reviews. Neuroscience*, 11 (2010), 747–59

Bornstein, Harry / Roy, Howard, 'Comment on "Linguistic Deficiency and Thinking: Research with Deaf Subjects 1964-1969"', *Psychological Bulletin*, 79 (1973), 211–214

Brainard, Michael S. / Fitch, Tecumseh W., 'Editorial Overview: Animal Communication and Human Language', *Current Opinion in Neurobiology*, 28 (2014), v–viii

Brenowitz, Eliot / Perkel, David J. / Osterhout, Lee, 'Language and Birdsong: Introduction to the Special Issue', *Brain and Language*, 115 (2010), 1–2

Brueggemann, Brenda, *Lend Me Your Ear: Rhetorical Constructions of Deafness* (Washington DC: Gallaudet University Press, 1999)

Campbell, Matthew / de Waal, Frans, 'Chimpanzees Empathize with Group Mates and Humans, but Not with Baboons or Unfamiliar Chimpanzees.', *Proceedings of the Royal Society B: Biological Sciences*, 281 (2014), 20140013

Candolin, Ulri, 'The Use of Multiple Cues in Mate Choice', *Biological Reviews*, 78 (2003), 575–595

Cheney, Dorothy / Seyfarth, Robert, 'Constraints and Preadaptations in the Earliest Stages of Language Evolution', *The Linguistic Review*, 22 (2005), 135–159

Chomsky, Noam, 'Language and Other Cognitive Systems. What Is Special About Language?', *Language Learning and Development*, 7 (2011), 263–278

Chough, Steven / Dobyns, Kristina, 'How Is Asian Deaf Culture Different from American Deaf Culture', in *The Deaf Way II Reader*, ed. by Goodstein, Harvey (Washington D.C.: Gallaudet Univ. Press, 2006), pp. 227–232

Christiansen, Morten / Chater, Nick, 'The Language Faculty That Wasn't: A Usage-Based Account of Natural Language Recursion', *Frontiers in Psychology*, 6 (2015), 1–18

van Cleve, John, *A Place of Their Own: Creating the Deaf Community in America* (Gallaudet: Gallaudet University Press, 1989)

Crystal, David, 'Language', *Encyclopædia Britannica Online* (Encyclopædia Britannica Inc., 2015)

Davis, Hallowell / Silverman, Richard, *Hearing and Deafness* (New York: Holt, Rinehart & Winston, 1970)

DeMatteo, Asa, 'Visual Imagery and Visual Analogues in American Sign Language', in *On the other hand : new perspectives on American sign language*, ed. by Friedman, Lynn (New York: Academic Press, 1977), pp. 109–136

Douglas, Heather E, *Science, Policy, and the Value-Free Ideal* (Pittsburgh: University of Pittsburgh Press, 2009)

Elster, Jon, *The Cement of Society: A Study of Social Order* (New York: Cambridge Univ Press, 1989)

Enfield, Nick J, *The Anatomy of Meaning* (Cambridge: Cambridge University Press, 2009)

Evans, Nicholas, 'Language Diversity as a Resource for Understanding Cultural Evolution', in *Cultural Evolution: Society, Technology, Language, and Religion*, ed. by Richerson, Peter J. and Christiansen, Morten (Cambridge: MIT Press, 2013), pp. 233 – 268

Evans, Nicholas / Levinson, Stephen, 'The Myth of Language Universals: Language Diversity and Its Importance for Cognitive Science.', *The Behavioral and brain sciences*, 32 (2009), 429–448

Fant, Lou, 'Two Memorable Meals with Ursula and Ed', in *The Signs of Language Revisited*, ed. by Emmorey, Karen and Lane, Harlan (Mahwah, New Jersey, London: Lawrence Erlbaum Associates, Inc., 2000), pp. 3–4

Fisch, R., 'Psychology of Science', in *Science, Technology and Society* (London, Beverly Hills: SAGE Publications, 1977), pp. 277 – 299

Fisher, Simon E / Marcus, Gary F, 'The Eloquent Ape: Genes, Brains and the Evolution of Language.', *Nature reviews. Genetics*, 7 (2006), 9–20

Fitch, Tecumseh W., *The Evolution of Language* (Cambridge, Massachusetts: Cambridge University Press, 2010)

Fricke, Ellen, 'Towards a Unified Grammar of Gesture and Speech: A Multimodal Approach', in *Body – Language – Communication.*, ed. by Müller, Cornelia et al (Berlin, Boston: De Gruyter Mouton, 2013), pp. 733–754

Fu, Feng et al, 'Evolution of in-Group Favoritism', *Scientific Reports*, 2 (2012), 1–6

Furth, Hans G, 'Linguistic Deficiency and Thinking: Research with Deaf Subjects 1964-1969', *Psychological Bulletin*, 76 (1971), 58–72

---, 'The Influence of Language on the Development of Concept Formation in Deaf Children', *Journal of Abnormal and Social Psychology*, 63 (1961), 386–389

Gallaudet, Edward M., *The Deaf and Their Possibilities* (Chicago: University of Chicago Press, 1898)

Gannon, Jack R., *Deaf Heritage: A Narrative History of Deaf America*, ed. by Butler, Jane and Gilbert, Laura-Jean, 2nd print (Washington D.C.: Gallaudet Univ. Press, 2012)

Gärdenfors, Peter, 'The Role of Cooperation in the Evolution of Protolanguage and Language', in *Evolution of Mind, Brain, and Culture*, ed. by Hatfield, Gary and Pittman, Holly (Philadelphia: University of Pennsylvania Press, 2013), pp. 193 – 216

Gardner, Richard / Gardner, Beatrix, 'Teaching Sign Language to a Chimpanzee.', *Science*, 165 (1969), 664–72

Garner, Richard L., *The Speech of Monkeys* (London: William Heineman, 1892)

Gessinger, Joachim, 'Der Ursprung Der Sprache Aus Der Taubstummheit', in *Theorien vom Ursprung der Sprache. Band II*, ed. by Gessinger, Joachim and von Rahden, Wolfert (Berlin, New York: de Gruyter, 1989), pp. 345 – 387

Gómez, Rebecca L. / Gerken, Lou A., 'Infant Artificial Language Learning and Language Acquisition.', *Trends in cognitive sciences*, 4 (2000), 178–186

Goodhill, V, 'The Educational Treatment of the Pre-School Deaf Child.', *The Laryngoscope*, 57 (1947), 555–63

Harlan, Lane, 'Construction of Deafness', in *The Disability Studies Reader*, ed. by Davis, Lennard J. (New York, London: Routledge, 1997), pp. 153 – 171

Harris, Roy, *The Language Myth* (London: Gerald Duckworth & Company Ltd., 1981)

Haun, Daniel / Over, Harriet, 'Like Me: A Homophily-Based Account of Human Culture', in *Cultural Evolution: Society, Technology, Language, and Religion*, ed. by Richerson, Peter J. and Christiansen, Morten H. (Cambridge: MIT Press, 2013), pp. 75 – 85

Haun, Daniel / Rekers, Yvonne / Tomasello, Michael, 'Majority-Biased Transmission in Chimpanzees and Human Children, but Not Orangutans', *Current Biology*, 22 (2012), 727–731

Hewes, Gordon, 'Primate Communication and the Gestural Origin of Language', Current Anthropology, 14 (1973), 5 – 24

Hill, Jane, ‘On the Evolutionary Foundations of Language’, American Anthropologist, 74 (1972), 308–317

Hintikka, Jaakko, ‘Paradigms for Language Theory’, in *Paradigms for Language Theory and other Essays* (Dordrecht: Kluwer Academic Publishers, 1998), 146 – 174

Hockett, Charles, ‘The Origin of Speech’, *Scientific American*, 203 (1960), 88–111

Huxley, Thomas H., *Man’s Place in Nature* (New York: D. Appleton and Company, 1863)

Jackendoff, Ray, ‘What Is the Human Language Faculty?: Two Views’, *Language*, 87 (2011), 586–624

---, ‘Your Theory of Language Evolution Depends on Your Theory of Language’, in *The Evolution of Human Languages: Biolinguistic Perspectives*, ed. by Larson, Richard K. / Dèprez, Viviane and Yamakido, Hiroko (Cambridge: Cambridge University Press, 2010), pp. 63 – 72

Jespersen, Otto, ‘Language’, *Encyclopaedia Britannica* (William Benton, 1951), pp. 696 – 703

Joseph, John E. / Taylor, Talbot J., *Ideologies of Language*, ed. by Joseph, John E. and Taylor, Talbot J. (New York: Routledge, 1990)

Kainz, Friedrich, *Die ‘Sprache’ Der Tiere* (Stuttgart: Ferdinand Enke Verlag, 1961)

Kellogg, W. N., ‘Communication and Language in the Home-Raised Chimpanzee’, *Science*, 162 (1968), 423–427

Klima, Edward / Bellugi, Ursula, *The Signs of Language*, *New York* (Cambridge, Massachusetts: Harvard University Press, 1979)

Kroeber, AL, ‘Sign Language Inquiry’, *Intern. Journal of American Linguistics*, 24 (1958), 1–19

Lane, Harlan, ‘The Medicalization of Cultural Deafness in Historical Perspective’, in *Looking back*, ed. by Fischer, Renate and Lane, Harlan (Hamburg: Signum Verlag, 1993), pp. 479 – 493

---, *When the Mind Hears: A History of the Deaf* (New York: Random House, 1984)

Levine, John M, ‘Solomon Asch’s Legacy for Group Research.’, *Personality and Social Psychology Review*, 3 (1999), 358–364

Levinson, Stephen C / Dediu, Dan, ‘The Interplay of Genetic and Cultural Factors in Ongoing Language Evolution’, in *Cultural Evolution: Society, Technology, Language, and Religion*, ed. by Richerson, Peter J. and Christiansen, Morten H. (Cambridge: MIT Press, 2013), pp. 219 – 231

Levinson, Stephen C. / Holler, Judith, ‘The Origin of Human Multi-Modal Communication’, *Philosophical Transactions of the Royal Society B*, 369 (2014), 1–9

Liddell, Scott K., *Grammar, Gesture, and Meaning in American Sign Language* (Cambridge, New York: Cambridge University Press, 2003)

Lieberman, Philip, ‘Language Did Not Spring Forth 100,000 Years Ago.’, *PLoS biology*, 13 (2015), e1002064

Longino, Helen E, *Science as Social Knowledge: Value and Objectivity in Scientific Inquiry* (Princeton, New Jersey: Princeton University Press, 1990)

Lotz, John, ‘Speech and Language’, *Journal of Acoustical Society of America*, 22 (1950), 712 – 717

Love, Nigel, ‘Cognition and the Language Myth’, *Language Sciences*, 26 (2004), 525–544

Maher, Jane, *Seeing Language in Sign: The Work of William C. Stokoe* (Washington D.C.: Gallaudet Univ. Press, 1996)

Mallery, Garrick, *Sign Language among North American Indians : Compared with That among Other Peoples and Deaf-Mutes* (The Hague: Mounton, 1881)

Mally, Gertrud, ‘The Long Road to Self-Confidence of the Deaf in Germany’, in *Looking back* , ed. by Fischer, Renate and Lane, Harlan (Hamburg: Signum Verlag, 1993), pp. 177 – 198

Maynard-Smith, John / Harper, David, ‘Animal Signals’ (Oxford: Oxford University Press, 2003)

McDonnell, Patrick / Saunders, Helena, ‘Sit on Your Hands’, in *Looking back*, ed. by Fischer, Renate and Lane, Harlan (Hamburg: Signum Verlag, 1993), pp. 255 – 260

McNeill, David, *Gesture and Thought* (Chicago and London: University of Chicago Press, 2005)

Meier, Richard P., ‘Language and Modality’, in *Sign Language. An International Handbook*, ed. by Pfau, Roland / Steinbach, Markus and Woll, Bencie (Berlin, Boston: De Gruyter Mouton, 2012), pp. 574–601

Meir, Irit et al, ‘The Influence of Community on Language Structure’, *Linguistic Variation*, 12 (2012), 247–291

Mulkay, Michael J., 'Norms and Ideology in Science', *Social Science Information*, 15 (1976), 637–656
Myklebust, Helmer, *The Psychology of Deafness. Sensory Deprivation, Learning, and Adjustment.*, 2nd edn (New York, London: Grune & Stratton, 1964)
Myklebust, Helmer / Brutten, Milton, 'A Study of the Visual Perception of Deaf Children', *Acta oto-laryngologica. Supplementum*, 105 (1953), 1 – 126
Nomeland, Melvia M. / Nomeland, Ronald E., *The Deaf Community in America: History in the Making* (North Carolina: McFarland & Company, Inc., Publishers, 2012)
Northcutt, R G, 'Changing Views of Brain Evolution.', *Brain research bulletin*, 55 (2001), 663–674
Nottebohm, Fernando, 'The Origins of Vocal Learning', American Naturalist, 106 (1972), 116–140
Padden, Carol / Humphries, Tom, 'Inside Deaf Culture', *Sign Language Studies* (Harvard: Harvard University Press, 2006), 217
Payne, Arnold Hill, 'Deaf and Dumb', *Encyclopaedia Britannica* (Encyclopaedia Britannica Co., 1911), 880 – 894
Perniss, Pamela / Özyürek, Asli / Morgan, Gary, 'The Influence of the Visual Modality on Language Structure and Conventionalization: Insights From Sign Language and Gesture', *Topics in Cognitive Science*, 7 (2015), 2–11
Pfau, Roland / Steinbach, Markus / Woll, Bencie eds., *Sign Language. An International Handbook* (Berlin, Boston: De Gruyter Mouton, 2012)
Premack, David, 'The Codes of Man and Beasts', Behavioral & Brain Sciences, 6 (1983), 125–167
---, 'Language in Chimpanzee?', Science, 172 (1971), 808–822
Radick, Gregory, *The Simian Tongue: The Long Debate about Animal Language* (Chicago, London: The University of Chicago Press, 2007)
Reiss, Julian / Sprenger, Jan, 'Scientific Objectivity', *The Stanford Encyclopedia of Philosophy* (Edward N. Zalta (ed.), 2014)
Rivas, Esteban, 'Recent Use of Signs by Chimpanzees (Pan Troglodytes) in Interactions with Humans', Journal of comparative psychology, 119 (2005), 404–417
Robins, R. Henry, 'Language', *Encyclopaedia Britannica Micropaedia* (William Benton, 1977), 32
---, 'Language', *Encyclopaedia Britannica Micropaedia* (William Benton, 2002), 147
Sacks, Oliver, *Seeing Voices. A Journey into the World of the Deaf*, (NY: Harper Perennial, 1990)
Sandler, Wendy et al, 'The Gradual Emergence of Phonological Form in a New Language', *Natural Language and Linguistic Theory*, 29 (2011), 503–543
Sandler, Wendy / Lillo-Martin, Diane, *Sign Language and Linguistic Universals* (Cambridge: Cambridge University Press, 2006)
Schlesinger, Hilde S. / Meadow, Kathryn P., *Sound and Sign* (Berkeley: University of California Press, 1982)
Scott-Phillips, Thomas C., 'Nonhuman Primate Communication, Pragmatics, and the Origins of Language', *Current Anthropology*, 56 (2015), 56–80
---, *Speaking Our Minds* (London, New York: Palgrave Macmillan, 2015)
Scott-Phillips, Thomas C. / Kirby, Simon, 'Information, Influence and Inference in Language Evolution', in *Animal Communication Theory: Information and Influence*, ed. by Stegmann, Ulrich E. (Cambridge: Cambridge University Press, 2013), pp. 421–442
Seyfarth, Robert / Cheney, Dorothy, 'Signalers and Receivers in Animal Communication', *Annual Review of Psychology*, 54 (2003), 145–173
Seyfarth, Robert M. / Cheney, Dorothy / Marler, Peter, 'Monkey Responses to Three Different Alarm Calls: Evidence of Predator Classification and Semantic Communication.', *Science*, 210 (1980), 801–803
Simms, Laurene, 'Deaf Education: Whose "Way" Is It?', in *The Deaf Way II Reader.*, ed. by Goodstein, Harvey (Washington D.C.: Gallaudet Univ. Press, 2006), pp. 81 – 86
Simpson, Jane, 'Language', *The Encyclopedia of Language and Linguistics* (Pergamon Press, 1994), p. 1894
Slocombe, Katie E. / Waller, Bridget M. / Liebal, Katja, 'The Language Void: The Need for Multimodality in Primate Communication Research', *Animal Behaviour*, 81 (2011), 919–924
Smith, Eric A., 'Communication and Collective Action: Language and the Evolution of Human Cooperation', *Evolution and Human Behavior*, 31 (2010), 231–245

Southwood, Nicholas / Eriksson, Lina, 'Norms and Conventions', *Philosophical Explorations*, 14 (2011), 195–217

Sperber, Dan / Origgi, Gloria, 'A Pragmatic Perspective on the Evolution of Language', in *The Evolution of Human Language*, ed. by Larson, Richard K / Deprez, Viviane, and Yamakido, Hiroko (Cambridge: Cambridge University Press, 2010), pp. 124 – 131

Stokoe, William C, 'Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf [1960]', *Journal of deaf studies and deaf education*, 10 (2002), 3–37

Tajfel, Henri, 'Social Categorization, Social Identity and Social Comparison', in *Differentiation between Social Groups: studies in the social psychology of intergroup relations*, ed. by Tajfel, Henri (London: Academic Press, 1978), pp. 61 – 76

Tajfel, Henri et al, 'Social Categorization and Intergroup Behaviour', *European Journal of Social Psychology*, 1 (1971), 149–178

Tattersall, Ian, 'A Putative Role for Language in the Origin of Human Consciousness', in *The Evolution of Human Language*, ed. by Larson, Richard K / Deprez, Viviane, and Yamakido, Hiroko (Cambridge: Cambridge University Press, 2010), pp. 193 – 198

Taylor, Talbot J., 'Folk-Linguistic Fictions and the Explananda of the Language Sciences', *New Ideas in Psychology*, in press (2015), 1–7

---, 'The Origin of Language: Why It Never Happened.', *Language sciences*, 19 (1997), 67–77

Terrace, Herbert / L a Petitto / R J Sanders / T G Bever, 'Can an Ape Create a Sentence?', *Science*, 206 (1979), 891–902

Tervoort, Bernhard, *Developmental Features of Visual Communication* (Amsterdam, Oxford: North-Holland Publishing Company, 1975)

Tomasello, Michael, *Origins of Human Communication* (Cambridge, Massachusetts, London: The MIT Press, 2008)

Trager, George L., 'Language', *Encyclopaedia Britannica* (William Benton, 1962), p. 696

Vaesen, Krist, 'Chimpocentrism and Reconstructions of Human Evolution (a Timely Reminder)', *Studies in History and Philosophy of Biological and Biomedical Sciences*, 45 (2014), 12–21

Veditz, George W., 'Notices of Publications: The Deaf-Mute and His Language.', *American Annals of the Deaf*, 35 (1890), 271 – 275

Vigliocco, Gabriella / Perniss, Pamela / Vinson, David, 'Language as a Multimodal Phenomenon', *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences*, 369 (2014), 20130292

Vygotsky, Lev, *Thought and Language [1934]*, ed. by Alex Kozulin (Cambridge: MIT Press, 1986)

Vinck, Dominique, *The Sociology of Scientific Work* (Northhampton, Massachussetts: Edward Elgar Publishing Inc., 2010)

Whetnall, Edith / Fry, Dennis, *The Deaf Child* (London, Tonbridge: William Heinemann Medical Books Limited, 1964)

Woolard, Kathryn A / Schiefflein, Bambi B, 'Language Ideology', *Annual Review of Anthropology*, 23 (1994), 55–82

Wray, Alison / Grace, George, 'The Consequences of Talking to Strangers: Evolutionary Corollaries of Socio-Cultural Influences on Linguistic Form', *Lingua*, 117 (2007), 543–578

Wundt, Wilhelm, *Völkerpsychologie. Erster Band. Die Sprache.*, (Leipzig: Verlag von Wilhelm Engelmann, 1904)

Yerkes, Robert, Almost Human (London: J. Cape, 1925)

Zuberbühler, Klaus, 'Linguistic Capacity of Non-Human Animals', *Wiley Interdisciplinary Reviews: Cognitive Science*, 6 (2015), 313 – 321

---, 'Primate Communication', in *New Perspectives on the Origins of Language*, ed. by Lefebvre, Claire / Comrie, Bernard, and Cohen, Henri (Amsterdam: John Benjamins Publishing Co., 2013), pp. 187 – 210

---, 'The Phylogenetic Roots of Language: Evidence from Primate Communication and Cognition', *Current Directions in Psychological Science*, 2005, 126–130

Publication 2: Scala Naturae

Ullrich, R., Mittelbach, M., & Liebal, K. (2018). Scala Naturae: The Impact of Historical Values on Current "Evolution of Language" Discourse. *Journal of Language Evolution*, *3*(1), 1–12. https://doi.org/10.1093/jole/lzx017

Journal of Language Evolution, 2017, https://doi.org/10.1093/jole/lzx017

Scala Naturae: The impact of historical values on current 'evolution of language' discourse

Robert Ullrich[1*], Moritz Mittelbach[2], Katja Liebal[1]

[1] Department of Education and Psychology, Comparative Developmental Psychology, Freie Universität Berlin, Habelschwerdter Allee 45

[2] Department of Botany, Ecology of Plants, Freie Universität Berlin, Altensteinstr 6, Berlin 14195, Germany

[*]Corresponding author: robert.ullrich@fu-berlin.de; +49 30 838 57 690

Abstract

Various complaints about the consistent use of a non-epistemological 'norm of progress' (also known as 'Scala Naturae') can be found frequently in recent evolution of language and communication literature. Affiliated to earlier studies that addressed quantification of some overt indicators such as word combinations of 'high + species', the current account aims to go beyond the obvious in describing the presumed phenomena. Using a mixed-methodology approach, we quantify the general use of vocabulary, range of study species, amount of 'progressionist attributes' and subsequently qualify the context of some key words. Investigating 915 peer-reviewed articles from a species-comparative evolution of language and communication discourse, we found that articles focussing on species groups historically regarded as 'high' make more use of attributes implying directed progress than otherwise. We subdivided all articles in two distinct corpora. Articles using the term 'language' or 'speech' in title, abstract or keywords were labelled 'language'. Those using other terms than language were labelled 'communication'. We could identify a more diverse focus on studied species groups and a more behaviouristic vocabulary in corpus 'communication' as compared to the corpus 'language'. Additionally, articles from the latter corpus tend to stress a narrative of human uniqueness. Our results, taken together, do not provide clear evidence for a structural and active promotion of a 'norm of progress', but hint towards historical aftermaths exercising indirect influence and worthy of further study.

Background

Over the last decade a growing number of articles discuss the so called 'replication crisis' in psychology and other scientific disciplines (Ioannidis, 2012). A subsequent boost in meta-research found that many empirical results are not as robust as they originally seemed. Publication biases (Fanelli, 2010), insufficient replication (Makel, Plucker, & Hegarty, 2012), lack of data sharing (Wicherts, Borsboom, Kats, & Molenaar, 2006), questionable research practices (John, Loewenstein, & Prelec, 2012), or low statistical power (Bezeau & Graves, 2001) are just some of the factors identified as possible explanations for the crisis. However, the efforts so far seem to underestimate subjective influences such as personal expectations, use of terminology, as well as physical and psychological constraints in designing hypotheses and conducting experiments. In order to investigate these factors, meta-research must reach beyond quantitative methods and supplement them with a mixed-method approach. This means that the current investigation takes as a basis quantitative data (e.g. from text mining) and mixes it with qualitative material (e.g. analysis of context). Such an approach may be especially helpful in identifying and reflecting upon the structure and function of non-epistemological values or norms, such as ethical, social, or political considerations (Douglas, 2016). Given those non-epistemological norms on the one side, epistemological norms, such as 'reproducibility', 'scope', and 'transparency', on the other side, are an accepted integral part of scientific reasoning (Douglas, 2009, p. 17). According to this, it is only the first group of norms that is sometimes perceived as threat to scientific objectivity (see Hudson, 2016). From the perspective of a value-free ideal, non-epistemological norms (i.e. moral considerations) should not affect scientific practice, because their subjective element is a "remit of art, not of science" (Mogie, 2000, p. 869). However, several classic studies (Feyerabend, 1975; Latour & Woolgar, 1979) as well as some recent publications (Elliott & McKaughan, 2009, Davis, 2013; Douglas, 2016; Mascolo, 2016) question the value-free ideal. In general, so the criticism goes, scientists are part of society and therefore inextricably linked to its values (Douglas, 2016). As a consequence, any description of human or non-human behaviour that goes beyond mere observations draws inevitably on the bias of preconceptions: "The privileging of measurement over meaning puts the empirical cart before the conceptual horse." (Mascolo, 2016, p. 5). Following this line of reasoning, not only is data input influenced by subjective values such as 'preference for similar others' ("homophily"; Haun & Over, 2013) or a priori rejection of 'human-animal similarity' ("anthropodenial"; de Waal, 1999), but scientists' data output also has consequences in the social and ethical domain (Douglas, 2009, p. 115). Illustrative examples can be found with reference to sign languages. Until the mid- 20th century, a dominant preconception in science understood the oral modality as a necessary prerequisite for having 'language', which itself was supposed to be responsible for rationality and flexible communication (Ullrich, 2016, p. 185). As a consequence, deaf humans were forced to learn oral forms of communication instead of better-suited sign languages, with various negative consequences for decades (Ullrich, 2016, p. 189). One

cannot simply blame scientific concepts and the use of terminology for those developments, but they did play a major role. Values in science become more important where "social categories and the images they embed are inescapably value-laden" (Davis, 2013, p. 554). We believe this also to be the case with 'evolution of language and communication' discourse, where scientists try to create a valid human self-conception with reference to a supposedly human unique characteristic, namely language (e.g. Berwick, Friederici, Chomsky, & Bolhuis, 2013; Hauser et al., 2014; Scott-Phillips, 2015). Given the entanglements between science and non-epistemological values or norms, it appears to us more productive to monitor norms instead of combating them. In that respect we want to qualify and quantify one potentially lasting norm in order to enable future investigations towards experimental design, the formulation of questions and subsequent interpretation of data. As such, the study contributes to the process of scientific self-correction.

The potential norm at issue is the 'norm of progress' (also known as 'Scala Naturae' or 'Great Chain of Being'), which assumes that evolution proceeds in a linear 'upward' way from a simple/primitive condition towards an 'improved' state. Although modern evolutionary theory rejects this prediction (Johnson, Lahti, & Blumstein, 2012), a number of scientists complain about the persistence of the norm (Chittka, Rossiter, Skorupski, & Fernando, 2012, p. 2678; Cimatti & Vallortigara, 2015, p. 6; de Waal, 1999, p. 257; Emery & Clayton, 2004, p. 37; Fitch, Huber, & Bugnyar, 2010, p. 796; Nee, 2005).

A number of qualitative studies focus on the history and current influence of the 'norm of progress' (Ghiselin, 2005; Hodos & Campbell, 1969; Lovejoy, 1936; Ruse, 1996). By design they do not quantify the phenomenon in recent discourse. Thus, despite the frequent complaints regarding the persistence of the 'norm of progress', to date there are only two attempts to study the existence of this norm in more quantitative ways. In 2000 Mogie searched scientific papers published between 1995 and 1999 using the attributes 'higher' or 'lower' in descriptions of species within the title. A low-tech query returned over 700 positive hits, mostly in studies of plants (n=665) (Mogie, 2000). Following that study, in 2013, Rigato and Minelli performed a scientometric analysis of over 67,413 biological articles published between 2005 and 2010 in 16 different scientific journals. Their queries on journal websites identified 1,287 out of 67,413 articles (1.91%) using "Scala Naturae language" (Rigato & Minelli, 2013). Another query in the course of the same study on PubMed confirmed that more than 55% of all positive hits derive from Botany (Rigato & Minelli, 2013). Yet despite providing first evidence for possible implications of a non-epistemic norm within a discourse, neither study continues beyond overtly quantifiable issues, and both fail to identify any of the phenomena other academic peers have attributed to the realm of the norm. For instance, the historical exclusion of birdsong as a model of language (Sereno, 2014, p. 5), addressed by qualitative research, escaped these quantitative accounts.

Given the prevailing value-free ideal (Reiss & Sprenger, 2014, Chapter 3.1), it is assumed that non-epistemological norms are mostly deployed unintentionally and not overtly, and are therefore difficult to identify. For these reasons, the current study aims to go beyond easily accessible 'higher/lower classifications'. Instead, it quantifies implicit indicators of a 'norm of progress' in peer-reviewed publications on language/communication across a variety of species groups.

In order to do that we intent to divide articles from the evolution of language discourse into two distinct corpora: 'language' and 'communication'. The only reason we count an article to the corpus 'language' lies in the presence of the predefined terms 'language' or 'speech' in abstract, title, or keywords. On the other side, articles of corpus 'communication' use terms like 'signal', 'song', 'vocalisation', 'gesture', or 'communication'. The categorization makes no assertion about the actual focus of a publication.

We are aware that sometimes the terms 'language' and 'communication' are used in similar ways within the current study. It is not our attempt to equate these terms. However, definitions of 'language' are notoriously diverse (Botha, 2000). That includes perspectives which see 'communication' as mere side-effect of 'language' (e.g. Chomsky, 2011, pp. 264–65), as well as the opposite claim that regards communication as main driver for 'language evolution' (e.g. Okanoya, 2017; Zuberbühler, 2013, p. 188). Still, others interpret 'language' as part of a broader "communicative toolkit" which also includes music and animal song (Rohrmeier, Zuidema, Wiggins, & Scharff, 2015). In general we use a broad definition of 'language' that includes various cognitive (e.g. learning and memory) and physiological mechanisms (e.g. perception and motor control) (Fitch, 2017, p. 5).

The reason for dividing all articles in two corpora is the following: We hypothesise that authors using the word 'language' at prominent sections of an article, implicitly tie their research to a more human-centred perspective of research than researchers avoiding the term. If a 'norm of progress' exists, we would expect an increase of 'progressionist vocabulary' in the corpus 'language'. 'Progressionist vocabulary', like 'higher' or 'sophisticated', implies the existence of an improved, more sophisticated or more complex 'end state' (mostly realised in humans). Since evolutionary theory is not based on a teleological framework, an 'end state' cannot exist and the ranking of structures or abilities along a scale of improvement appears mostly human-centred and/or arbitrary. Therefore use of 'progressionist language' is not only ineloquent, but value-laden.

Hence, we assume that if the 'norm of progress' exists, we should find biased sampling of study species in the corpus 'language' compared to the corpus 'communication'. Within the total of 915 journal articles, we expect to identify value-laden 'progressionist' vocabulary, dependent on species, article format or corpus group.

Material & Methods

In order to gather corpus material we performed search queries on the citation database 'Scopus'. Our aim was to identify a specific fraction of articles concerned with evolution of language from a species-comparative point of view. We chose to select those specific articles for two reasons: first, in both past and current debates it is notoriously difficult to identify a generally accepted definition of 'language' (Botha, 2000), which makes the whole field of research an ideal candidate for speculation and value-laden narratives. Second, one controversial point in research on the 'language origin' concerns the question as to whether 'language' evolved either continuously across species (Wilcox, 1999; Hurford, 2014) or abruptly in human beings (Berwick et al., 2013). An answer might have wide-ranging implications for the human self-concept and, thus giving reason to expect value-loading on that issue in particular.

For the years 2005-2015 we selected from 16 Journals that have a high impact in the particular field of research (see **Tab. 1**).

Table 1 Composition of corpus 'language' and 'communication' by journal. Since two journals were founded in 2010 and 2011, respectively, they were not available for analysis before that year. Furthermore, publications from Behav. Brain. Sci. were not available as full text HTML before 2006 and thereby excluded for 2005.

Journal name	no. of papers in corpus 'language'	no. of papers in corpus 'communication'
Anim. Behav.	36	205
Anim. Cogn.	22	17
Behav. Brain. Sci.	106	7
Curr. Anthropol.	37	4
Curr. Biol.	44	43
Evol. Hum. Behav.	9	2
Evol. Psychol.	4	1
Front. Psychol.	22	2
J. Comp. Psychol.	8	1
Nat. Commun. (*2010)	9	7
Phil. Trans. R. Soc. B	33	11
PLoS Biol.	9	2
PLOS ONE	42	70
PNAS	33	18
Proc. R. Soc. B	19	74
Sci. Rep. (*2011)	6	12

Articles using 'language' or 'speech' in their abstract, title, or as keywords are collected in a corpus termed 'language' (n=890). To contrast the results we also wanted to identify publications focussing

on communication, signal, song, gesture or vocalisation. Articles using one of those terms in their abstract, title, or as keywords are collected in a corpus termed 'communication' (n=1107)[151]. All articles examined were manually checked for relevance by reading abstracts and key words. Articles were included in the corpus of investigation when they fulfilled the following requirements: they (i) use a comparative, cross-species approach; (ii) focus on language/communication (not cognition in general); (iii) focus on biological evolution (i.e. exclude machines); (iv) consider multicellular organisms (but not plants, fungi), and (v) focus on inter-individual communication.

Relevant articles ('language' n=439; 'communication' n=476) were supplemented with meta-information such as (a) species focus, (b) modality, and (c) full-text download link. With regard to (a) nine groups of species were identified (1. human primate, 2. non-human primate, 3. non-primate mammals, 4. marine mammal, 5. bird, 6. other vertebrates, 7. invertebrate, 8. fish, 9. unspecified). With reference to (b) seven modalities were identified (1. acoustic, 2. visual, 3. chemical, 4. tactile, 5. thermal, 6. cross-modal, 7. multimodal). Most articles were automatically retrieved[152] based on their link, converted from source HTML into a raw text format, and broken down to the level of individual words. Specific word classes were attributed automatically via TreeTagger using default settings (Schmid, 1995). In addition to obvious lemmas like 'high' and 'low' used by Rigato & Minelli 2013, we consider a greater number of terms as contributing to a 'norm of progress'. We created two groups of 56 handpicked lemmas (see Supplementary Material) to investigate the use of 'progressionist vocabulary', i.e. words that in a broader sense allow a linear differentiation between 'high' and 'low'. Those potentially value-laden lemmas were identified by earlier research as relating to the 'norm of progress' (Güntürkün & Bugnyar, 2016; Jarvis et al., 2005; Karten, 2015; McShea, 2011; Ruse, 1996; Ullrich, 2016) or were mentioned within an open survey by members of the Comparative Developmental Psychology group in Berlin (see Supplementary Material). For brevity we named those word groups 'high' and 'low' respectively and used them in order to compare the appearance of lemmas between corpora and various meta-data.

All quantitative analyses were performed using R 3.2 (R Development Core Team, 2016). A list of additional R-packages in use can be found at the Supplementary Methods section. To capture even subtle indicators of the 'norm of progress', the study combines quantitative text analysis and a qualitative audit of context (a mixed-methods approach). For qualitative analysis of context, we extracted respective text snippets into Excel Sheets and rated for context manually (i.e. 'opposite meaning', 'species related', 'neutral'). All R-Scripts used and consulted material are open and can be downloaded (doi 10.17605/OSF.IO/EGFHV).

[151] One example (for more see DOI 10.17605/OSF.IO/EGFHV) of a 'Scopus' query for the Journal *Animal Cognition* contributing to the corpus 'communication': TITLE-ABS-KEY (communication OR song OR signal* OR vocali?ation OR gesture AND evol* AND NOT language AND NOT speech) AND ISSN (1435-9456) OR ISSN (1435-9448) AND PUBYEAR AFT 2004 AND PUBYEAR BEF 2016

[152] Due to technical oddities this procedure had to be done by hand for two Journals: "Journal of Comparative Psychology" & "Current Anthropology".

Results & Discussion

Authors mostly avoid direct linkage of 'high/low' to various species groups

Rigato & Minelli (Rigato & Minelli, 2013) concluded that "the great chain of being is still there". When we reproduced their methodology for 915 publications from our corpus, we could identify 8 cases of direct linkage between 'high' and several species, but could not find any incidence with 'low'. Hits from 'higher' linked to either 'vertebrates' (Earley, 2010, p. 2676; Hauser et al., 2014, p. 1; Iriki & Taoka, 2012, p. 18) or 'primates' (Cunningham & Ramos, 2014, p. 806; Glickstein, 2007, p. 824; Jablonka, Ginsburg, & Dor, 2012, p. 2155; Sadagopan, Temiz-Karayol, & Voss, 2015, p. 10) with one exception of 'plants' (Caulier, Flammang, Gerbaux, & Eeckhaut, 2013, p. 1). By definition, publications from the field of botany were excluded, whereby 0.87% positive hits from 915 articles nearly resembles those botany-free results presented by Rigato & Minelli (Rigato & Minelli, 2013). Contrary to their interpretation, we do not conclude that results can lead us to state that researchers adhere to a 'norm of progress'. In all affected articles we could identify only one or two singular events linking 'high + species'. When checking those papers manually, we could not identify a systematic use of 'Scala Naturae language'. Instead we consider those findings as singular cases of "historical baggage" (Mogie, 2000, p. 868) where expressions and metaphors echo a long tradition of teleological thinking.

However, as previously mentioned, we did not assume that the linking of overt 'high' and 'low' classifications with various species would occur at a high frequency, since we expected non-epistemological norms to be mostly used unintentionally and therefore not overtly expressed in the text. This is why we started exploratory investigations for more implicit indicators that might impact the discourse.

Primates dominate corpus 'language'

In 2014, Sereno claims that "birdsong has often been dismissed as a model of human language for the reason that monkeys seem much smarter than some birds" (Sereno, 2014, p. 5). We wanted to quantify his complaint regarding 'Scala Naturae thinking' by checking its substance in current literature. As described in our methods section, we divided all articles into two corpora labelled 'language' and 'communication' respectively. Subsequently we decided to compare the range of studied species groups between the corpus 'language' and that of 'communication'.

We found a substantially wider range of studied species groups in the corpus 'communication' as compared to the corpus 'language' (**Fig. 1**). About 70% of all 439 articles using the terms 'language' or 'speech' in title, abstract or keywords focussed on primates. Broken down to specific groups we observed for the corpus 'language' that the majority of articles focussed specifically on human primates (38%), followed by non-human primates (32%), birds (11%), and finally publications without definite species focus (7.2%). In contrast, articles within the corpus 'communication' mostly

focussed on invertebrates (26.89%) and birds (26.68%), followed by other vertebrates and non-human mammals (both 10.29%). From an overall 476 articles investigating communication and its evolution, only 11 focussed on human (2.3%) and 40 on non-human primates (8.4%).

The results pertaining to humans in the corpus 'language' may not surprise, since many researchers regard language to be unique to them (e.g. Berwick et al., 2013; Hauser et al., 2014; Scott-Phillips, 2015). Nonetheless, almost 62% of all articles using 'language' in their title, abstract or keywords do focus on non-human animals, most of them on non-human primates. For our study it is of no importance to distinguish if those articles investigate the origin of 'language' or 'communication'. We also cannot distinguish between the opposite use of the term or its context. Apart from these issues, it strikes us that articles focussing on invertebrates, fish or other vertebrates avoided almost completely the term 'language' in their opening sections. Even when most articles on non-human primates in the corpus 'language' investigated the 'origin of communication' instead of 'language', it surprises us that only 8.4% of articles from the corpus 'communication' are dedicated to the non-human primate group. That leads us to the conclusion that researchers studying primates are more likely to use the term 'language' when investigating communicative behaviours than researchers concerned with other species groups. Given Sereno's statement (Sereno, 2014), we indeed conclude that articles from the corpus 'language' tended in relative numbers to 'dismiss' birds as a model for investigating the evolution of language.

What might explain this phenomenon? Researchers tend to see abilities that they value, which is more easily done in species that closely resemble humans, e.g. primates. For instance, the oral/acoustic modality of human communication is the subject of 58.5% of the studies within the corpus 'language'. Modalities presumably less relevant to average humans or multimodal accounts that received attention only recently are covered comparably less (crossmodal: 22.3%; visual: 11.2%; multimodal: 7.1%; chemical: <1%; see **Fig. S1**). Earlier studies have traced some historical sources of the phenomenon's origin such as an 'oral norm' (Ullrich, 2016), 'a priori biases' (Slocombe, Waller, & Liebal, 2011), 'Primatocentrism' (Cimatti & Vallortigara, 2015; Emery & Clayton, 2004) or 'Chimpocentrism' (Vaesen, 2014). Frequent focus on primates' unimodal behaviour in the early days of comparative communication studies might have caused an underestimation of the communicative abilities of non-primates, which in turn makes non-primate research look less interesting. The circle creates its own evidence and fuels a view of ascending 'complexity' over the course of evolutionary development.

In accordance with this interpretation, we examined if both corpora would differ in their use of directional language. Since in the corpus 'language', there are more species investigated historically considered 'high' than in the corpus 'communication', we expect to find more adjectives representing 'high' in the corpus 'language' than the other way around. We thus decided to identify the 80 most common adjectives used in both corpora.

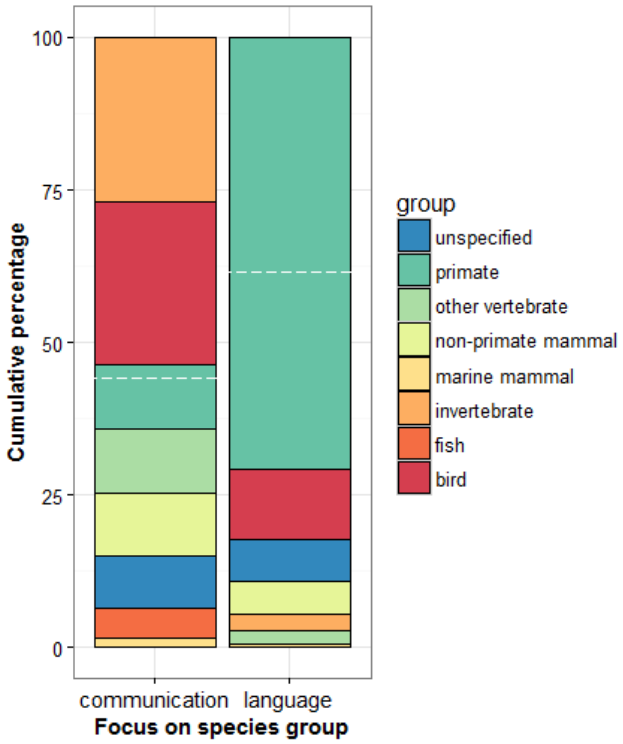

Figure 1 Comparison of the range of studied species groups in corpus 'language' and 'communication'. White horizontal line in primate box subdivides the group into human (above broken line) and non-human primates (below).

A selection of the 80 most common adjectives hints towards substantially different narratives across the two corpora

Adjectives give more precise information about a particular object of interest. Therefore, in case the 'norm of progress' influences scientific publications, we would expect more adjectives implying 'high' in the corpus 'language' as compared to the corpus 'communication'. This expectation is based on the following reflection: if articles from the corpus 'language' focus mostly on species groups that were considered 'high' under the terms of a 'norm of progress', and articles from the corpus 'communication' deal with 'lower' ones, adjectives implying 'high' should appear more frequently in corpus 'language'.

However, the analysis for the 80 most frequent adjectives did not meet our initial expectations. Indeed, the adjective 'complex' occurs more often in 'language' as compared to 'communication', while adjective 'low' followed the opposite pattern (**Fig. 2**). However, we were not able to detect any structural regularity that would systematically ascribe 'high' or 'low' value-laden adjectives to any of the corpora. Instead we became interested in those adjectives without respective counterparts within the list.

With regard to the corpus 'language', examples of some frequently employed adjectives are: cognitive, linguistic, communicative, neural, functional, cultural, syntactic, gestural, and semantic. With regard to the corpus 'communication', some examples are: sexual, reproductive, sensory,

aggressive, conspecific, facial, territorial, and dominant. It appears to us that those words tell very different stories about similar observations. One (corpus 'communication') investigates the communicative behaviour of a species for the sake of the species itself, while the other corpus ('language') aims to compare communicative behaviour between non-human and human animals.

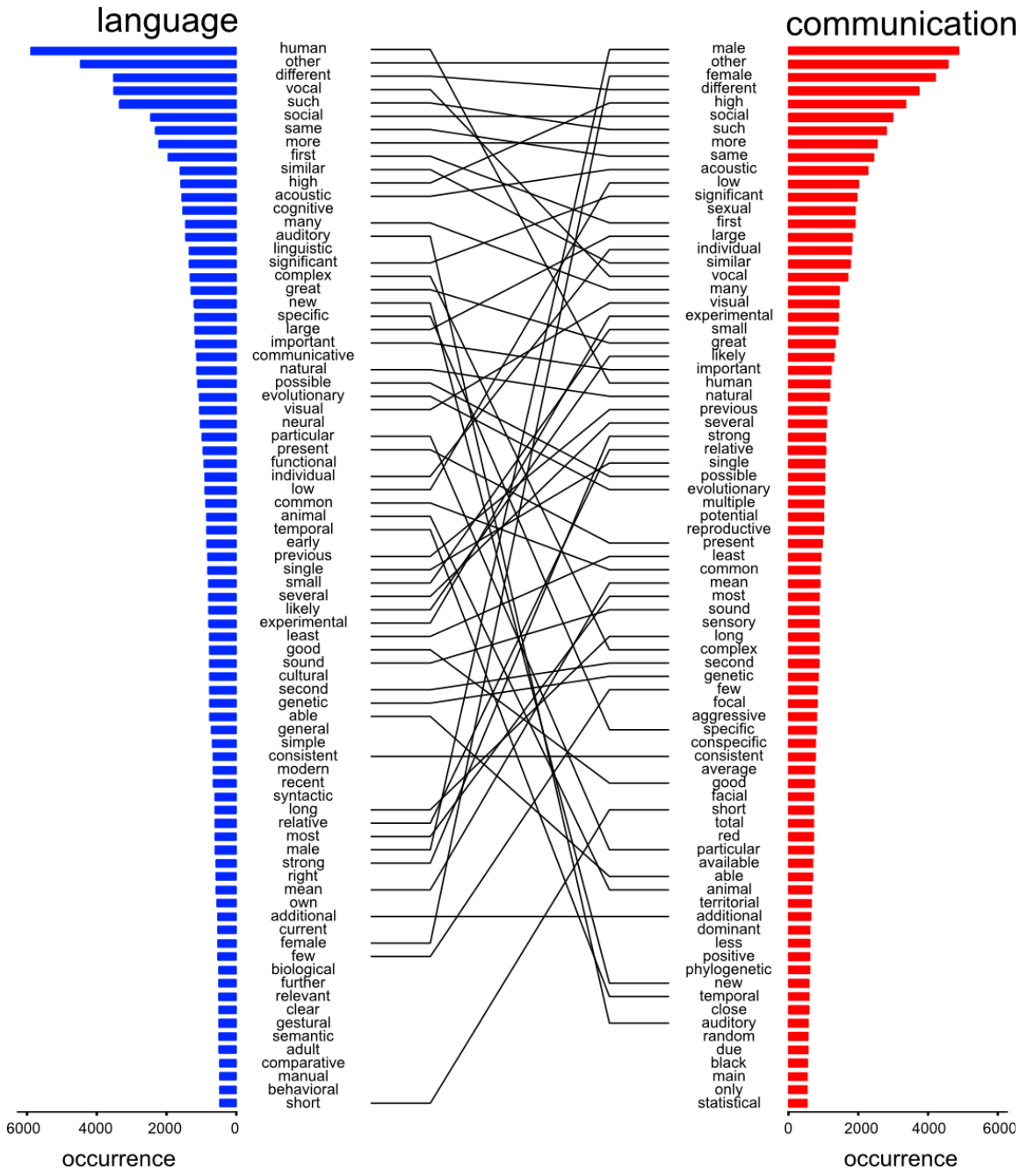

Figure 2 List of the 80 most common adjectives of the respective corpus, ordered by their occurrence. Adjectives that appear on either side are linked by lines. Adjectives without line do not have a respective counterpart among the most frequent 80.

Articles using the term ‘language’ in the abstract, title or key words tend to link and compare their findings to cognition and linguistic concepts, aspects that were investigated in former times under the umbrella term ‘animal psychology’. Articles avoiding ‘language’ concentrate on ecology and ethology, aspects that are investigated under the umbrella term ‘behaviourism’.

However, since extracting the 80 most common adjectives did not answer the question as to whether one of the corpora would feature the more frequent deployment of ‘high’ or ‘low’ classifications, we then decided to directly create a list of target words with the objective of comparing them accordingly.

No difference in directional vocabulary between corpus, but between species group and articles type

Due to the different emphasis on species groups between the corpora and the identification of two diverging uses of vocabulary when writing up results, we were interested in whether a selected list of words could also reveal a difference in the use of lemmas classified as ‘high’ or ‘low’. We predicted that under the terms of a persistent ‘norm of progress’, articles in the corpus ‘language’ would use more lemmas implying values of ‘high’ while avoiding those implying ‘low’, as compared to the corpus ‘communication’. To quantify frequencies of word appearances, we created a list of 58 words which either imply evolutionary ‘improvement’ or ‘simplicity’. The choice of words was based on earlier research (Güntürkün & Bugnyar, 2016; Jarvis et al., 2005; Karten, 2015; McShea, 2011; Ruse, 1996; Ullrich, 2016) and an open survey among researchers in comparative psychology (see Supplementary Material). To account for different text length, we corrected all hits by the total number of words per article. Altogether we found that words of the category ‘high’ are used 40% more often in the corpus ‘language’ and 32% more often in the corpus ‘communication’ as compared to words of category ‘low’. However, the difference of direct hits between the corpora was not as clear as expected. Indeed, publications of corpus ‘language’ did use words classified as ‘high’ 10.51% more often and words classified as ‘low’ 1.55% less often as compared to the corpus ‘communication’. When related to other factors such as ‘species group’ and ‘article type’, these results shift in weight and appear rather comparable.

Indeed, relative frequency differed substantially between various groups of species (**Fig. 3**). Publications in the corpus ‘language’ focussing on non-human primates use vocabulary from the word group ‘high’ more often than, for instance, articles focussing on birds (+27% in corpus ‘language’; +23% in corpus ‘communication’). Similarly, articles in the corpus ‘language’ without any focus on a species group used words classified as ‘high’ with increasing frequency as compared to articles with a focus on birds (+35% in corpus ‘language’; +18% in corpus ‘communication’). In general, we observed a tendency by which articles focussing on species groups ranking ‘high’ according to a ‘norm of progress’ increase their use of words valuing ‘high’. The small sample for articles focussing on humans (n=10) in the corpus ‘communication’ constitute an exception to this observation. Since

many articles with 'unspecified' species groups appear to consist of comments, review or theory pieces, we also wanted to quantify differences for that factor. We found that not only does the species group influence linguistic usage, but also the article format (**Fig. 4**). In general we observed a tendency whereby articles with less experimental or empirical focus increase their use of words defined as 'high/low'. For instance, in articles classified as theoretical publications we identified an increase of words categorised as 'high' by 23.6% (corpus 'language') and 26.5% (corpus 'communication') respectively. Following various conference discussions, most scientists eagerly deny reference to any such 'norm of progress'[153]. As said earlier mostly there is no active promotion for such a norm as it appears 'unscientific' (Mogie, 2000, p. 869).

Of course, a quantitative text analysis cannot determine to what degree vocabulary is used deliberately, or in which context. For that reason we checked context for one specific word that scientists usually value: 'unique'. '

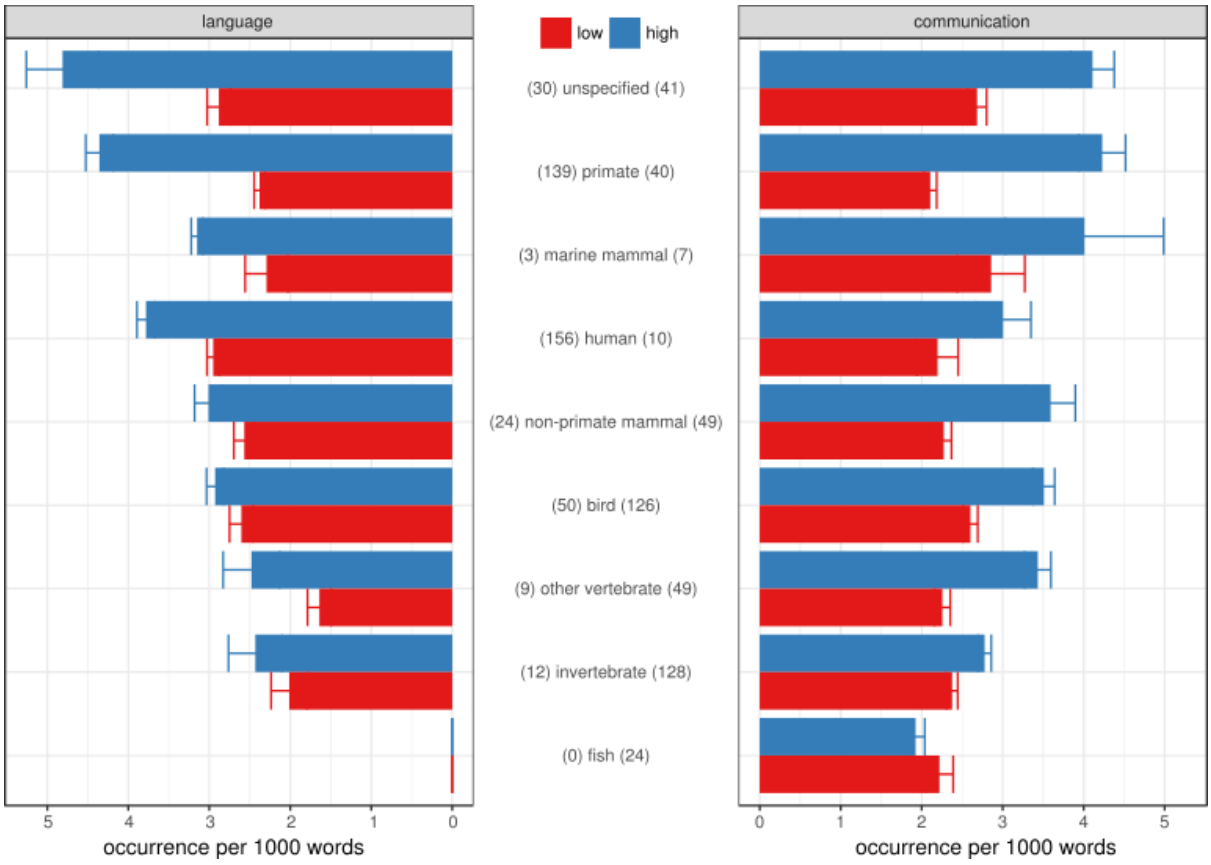

Figure 3 Occurrence of words classified as 'high' (blue) and 'low' (red) computed per 1000 words of original article. Results are broken down for species group and ordered by their summarized mean for 'high'. Horizontal bars indicate mean for all articles of a document type. Numbers in brackets represent articles under investigation. Error bars depict the standard error.

[153] e.g. personal communication to Andrew Whiten, T. Scott-Phillips.

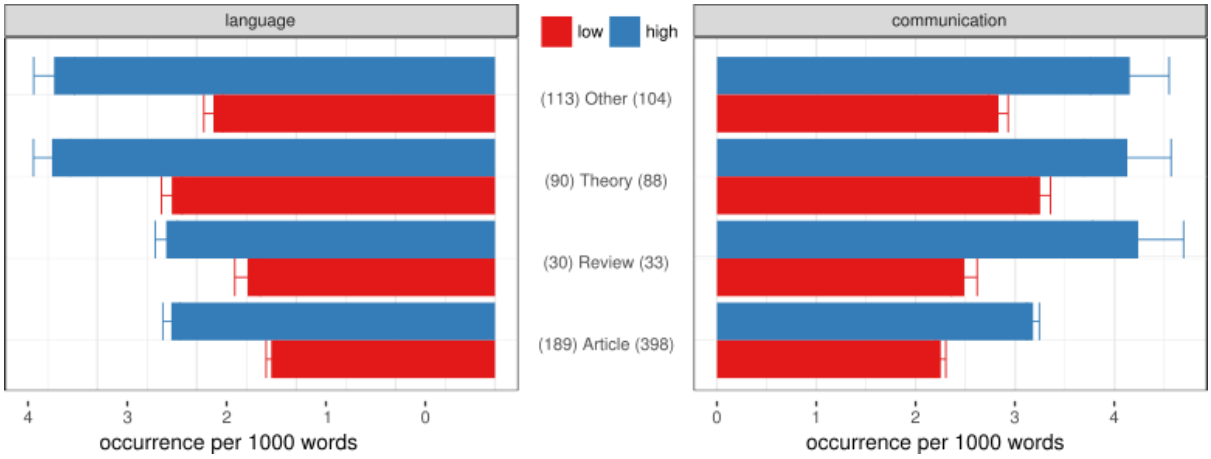

Figure 4 Occurrence of words classified as 'high' (blue) and 'low' (red) computed per 1000 words of original article. Results are broken down for article type and ordered by their summarized mean for 'high'. Horizontal bars indicate mean for all articles of a species group. Numbers in brackets represent articles under investigation. Error bars depict the standard error.

Language' more unique than 'communication'

In order to evaluate our previous results, we wanted to approach the problem of context blindness for one case example. We chose the lemma 'unique', because usually its usage does not imply directional connotations. Furthermore, from a biological point of view there is nothing special about being 'unique', since every species is defined by its autapomorphy – a derived trait that defines the status as a species.

However, based on previous qualitative research we hypothesised that when publications repeatedly state something as uniquely human, but do not mention anything else as unique in non-humans, than this one-sided view might hint towards values in use. In order to test the hypothesis we first quantified the phenomenon and subsequently qualified the results.

We found that 52% of all articles in the corpus 'language' make use of the lemma 'unique', but only 37% of articles in the corpus 'communication'. If the lemma is used, articles in the corpus 'language' use it on average 3.2 times, while articles in the corpus 'communication' employ it 1.9 times. Altogether we identified 57% more instances of the lemma 'unique' in the corpus 'language' as compared to 'communication'.

In order to investigate the qualitative context of the lemma, we extracted all occurrences, including the context, and validated its usage. When 'unique' referred to any species group, we labelled it accordingly. When used to the contrary (e.g. 'not unique'), we labelled it 'opposite'. When used without reference to any species, we labelled it 'neutral'. When used in context of an unanswered question or within quotations, we labelled it 'undecided'. We found that in nearly half the cases from the corpus 'language' the lemma 'unique' referred to humans, while in 78% of all incidences in the corpus 'communication' it was used in neutral manner (**Fig. 5**).

Certainly, it might come of no surprise that the term ‘unique’ appears frequently compared to ‘human’, since language is regarded as one of the important autapomorphies of the species. However, all species-specific forms of communication are unique by definition. Either someone takes the view that human language is unique and thus not comparable to any non-human form of communication, or one conducts species comparative research and therefore allows a comparison of language and animal communication. When following the second strategy, the consequence is that not only language is unique to humans, but also ultrasonic social communication to bats, electric communication signals to electric fish, and multimodal chemo-acoustic signals to lemurs. Still, in only 44 cases, ‘unique’ relates to the behaviour of a species in the corpus ‘communication’, while we could find 388 such cases (mostly in reference to humans) in the corpus ‘language’. That might constitute a scientific narrative that justifies human speciality as an evolutionary ‘achievement’. As such it hints towards a somewhat chauvinistic function where non-human species are not actively discriminated, but implicitly eclipsed. While scientists highlight human uniquely features, they also feel the urge to find biological ‘roots’ of behaviours and thus start testing and observing ‘downwards’ along the ‘evolutionary tree’. In this respect such a research agenda could be classified as motivated by the vestiges of a historical ‘norm of progress’.

After all, non-epistemological norms do indeed play a role within scientific reasoning (see Douglas, 2000). However, the task of monitoring them is always valuable and never completed, enabling readers to develop a critical view of hypotheses, questions, and results.

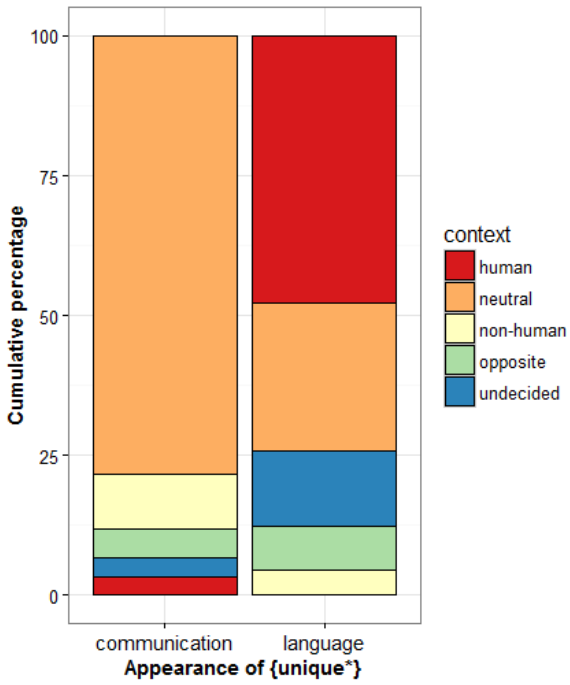

Figure 5 Comparison of context from the lemma ‘unique*’ {including: ‘uniquely’ and ‘uniqueness’} between corpora. In the corpus ‘communication’, the lemma ‘unique*’ is mostly used in neutral context, whereas its use in the corpus ‘language’ refers in almost half of the incidences to humans. See text for details and definition of individual labels.

Conclusion

In order to quantify a possible non-epistemological 'norm of progress' within a current scientific discourse of language evolution, we applied a quantitative text and qualitative context analysis to a corpus consisting of 915 articles. Historically one can find clear evidence for the existence of a 'norm of progress' in scientific publications. A reproduction of a study by Rigato and Minelli (Rigato & Minelli, 2013), however, could show only minor evidence for an open and active promotion of that norm.

Hence the focus of subsequent tests was put on implicit factors such as species range, use of vocabulary and values in language. Although papers from corpus 'language' and 'communication' focus on a similar phenomenon, their narratives appear strikingly different as indicated by the frequency of 80 of the most commonly employed adjectives. In addition, both corpora differ widely in their range of studied species groups and the usage of the lemma 'unique'. In all cases, the corpus 'language' establishes a narrative of human speciality, compared to other species, as could be additionally shown by qualifying all uses of the lemma 'unique' within the corpus. However, both corpora use more frequently words in the category 'high' with reference to primates as compared to birds or insects. Taken together, there is no evidence for a structural and overt promotion of a non-epistemological 'norm of progress' within the discourse. Still, several implicit factors hint at the lingering historical aftermaths of norm-related ideas and an associated subconscious function as leading forces in identifying and formulating current and future research questions.

Competing Interest
We have no competing interests.
Authors' Contributions
RU conceived and designed the study, collected corpus material, did qualitative analysis, drafted the manuscript, analysed and interpreted quantitative data and wrote parts of R code.
MM substantially participated in quantitative data analysis & wrote the majority of R code.
KL did critical revision of article drafts and provided important intellectual content.
All authors gave final approval for publication.
Acknowledgements
We thank Prof. Markus Wild and participants from the Workshop "Minds of Animals" in Bern 2016 for helpful comments.
Research Ethics
The study did not require ethical approval from a local ethics committee.
Data availability
Data and research materials supporting the results in the article are open and available at the Open Science Framework: DOI 10.17605/OSF.IO/EGFHV.
Source of funding
RU & MM received no specific funding for this work. KL receives a DFG Excellence Initiative Grant.

Literature

Berwick, R. C., Friederici, A. D., Chomsky, N., & Bolhuis, J. J. (2013). Evolution, brain, and the nature of language. *Trends in Cognitive Sciences*, *17*(2), 98. http://doi.org/10.1016/j.tics.2012.12.002

Bezeau, S., & Graves, R. (2001). Statistical power and effect sizes of clinical neuropsychology research. *Journal of Clinical and Experimental Neuropsychology*, *23*(3), 399–406. http://doi.org/10.1076/jcen.23.3.399.1181

Bolhuis, J. J., Tattersall, I., Chomsky, N., & Berwick, R. C. (2014). How Could Language Have Evolved? *PLoS Biology*, *12*(8), e1001934. http://doi.org/10.1371/journal.pbio.1001934

Botha, R. (2000). Discussing the evolution of the assorted beasts called language. *Language & Communication*, *20*, 149–160. http://doi.org/10.1016/S0271-5309(99)00022-1

Caulier, G., Flammang, P., Gerbaux, P., & Eeckhaut, I. (2013). When a repellent becomes an attractant: harmful saponins are kairomones attracting the symbiotic Harlequin crab. *Scientific Reports*, *3*, 2639. http://doi.org/10.1038/srep02639

Chittka, L., Rossiter, S. J., Skorupski, P., & Fernando, C. (2012). What is comparable in comparative cognition? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*(1603), 2677–85. http://doi.org/10.1098/rstb.2012.0215

Chomsky, N. (2011). Language and Other Cognitive Systems. What Is Special About Language? *Language Learning and Development*, *7*(4), 263–278. http://doi.org/10.1080/15475441.2011.584041

Cimatti, F., & Vallortigara, G. (2015). So little brain, so much mind. Intelligence and behaviour in non human animals. *Reti, Saperi, Linguaggi*, *4*(7), 5–20. http://doi.org/10.12832/81287

Cunningham, C. L., & Ramos, M. F. (2014). Effect of training and familiarity on responsiveness to human cues in domestic dogs (Canis familiaris). *Animal Cognition*, *17*, 805–814. http://doi.org/10.1007/s10071-013-0714-z

Davis, J. E. (2013). Social Science, Objectivity, and Moral Life. *Society*, *50*(6), 554–559. http://doi.org/10.1007/s12115-013-9710-9

de Waal, F. B. M. (1999). Anthropomorphism and Anthropodenial: Consistency in Our Thinking about Humans and Other Animals. *Philosophical Topics*, *27*(1), 255–280.

Douglas, H. (2000). Inductive Risk and Values in Science. *Philosophy of Science*, *67*(4), 559. http://doi.org/10.1086/392855

Douglas, H. (2009). *Science, policy, and the value-free ideal*. Pittsburgh: University of Pittsburgh Press.

Douglas, H. (2016). Values in Science. In P. Humphreys (Ed.), *The Oxford Handbook of Philosophy of Science* (pp. 609–632). Oxford, New York: Oxford University Press. http://doi.org/10.1093/oxfordhb/9780199368815.013.28

Earley, R. L. (2010). Social eavesdropping and the evolution of conditional cooperation and cheating strategies. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *365*(1553), 2675–2686. http://doi.org/10.1098/rstb.2010.0147

Elliott, K. C., & McKaughan, D. J. (2009). How Values in Scientific Discovery and Pursuit Alter Theory Appraisal. *Philosophy of Science*, *76*(5), 598–611. http://doi.org/10.1086/605807

Emery, N., & Clayton, N. S. (2004). Comparing the Complex Cognition of Birds and Primates. In L. J. Rogers & G. Kaplan (Eds.), *Comparative Vertebrate Cognition* (pp. 3–55). New York: Springer Science+Business Media.

Fanelli, D. (2010). Do pressures to publish increase scientists' bias? An empirical support from US states data. *PLoS ONE*, *5*(4). http://doi.org/10.1371/journal.pone.0010271

Feyerabend, P. (1975). *Against Method* (3rd 1993). London: Verso.

Fitch, W. T. (2017). Empirical approaches to the study of language evolution. *Psychonomic Bulletin & Review*. http://doi.org/10.3758/s13423-017-1236-5

Fitch, W. T., Huber, L., & Bugnyar, T. (2010). Social Cognition and the Evolution of Language: Constructing Cognitive Phylogenies. *Neuron*, *65*(6), 795–814. http://doi.org/10.1016/j.neuron.2010.03.011

Ghiselin, M. T. (2005). The Darwinian revolution as viewed by a philosophical biologist. *Journal of the History of Biology*, *38*(1), 123–136. http://doi.org/10.1007/s10739-004-6513-2

Glickstein, M. (2007). What does the cerebellum really do? *Current Biology*, *17*(19), 824–827. http://doi.org/10.1016/j.cub.2007.08.009

Güntürkün, O., & Bugnyar, T. (2016). Cognition without Cortex. *Trends in Cognitive Sciences*, *xx*, 1–13. http://doi.org/10.1016/j.tics.2016.02.001

Haun, D., & Over, H. (2013). Like me: a homophily-based account of human culture. In P. J. Richerson & M. H. Christiansen (Eds.), *Cultural Evolution: Society, Technology, Language, and Religion* (pp. 75–85). Cambridge: MIT Press.

Hauser, M. D., Yang, C., Berwick, R. C., Tattersall, I., Ryan, M. J., Watumull, J., … Lewontin, R. C. (2014). The mystery of language evolution. *Frontiers in Psychology*, *5*(MAY), 1–12. http://doi.org/10.3389/fpsyg.2014.00401

Hodos, W., & Campbell, C. G. B. (1969). Scala Naturae: Why There Is No Theory In Comparative Psychology. *Psychological Review*, *76*(4), 337–350.

Hudson, R. (2016). Why We Should Not Reject the Value-Free Ideal of Science. *Perspectives on Science*, *24*(2), 167–191. http://doi.org/10.1162/POSC_a_00199

Hurford, J. R. (2014). *Origins of Language: A Slim Guide*. Oxford: Oxford University Press.

Ioannidis, J. P. A. (2012). Why Science Is Not Necessarily Self-Correcting. *Perspectives on Psychological Science*, *7*(6), 645–654. http://doi.org/10.1177/1745691612464056

Iriki, A., & Taoka, M. (2012). Triadic (ecological, neural, cognitive) niche construction: a scenario of human brain evolution extrapolating tool use and language from the control of reaching actions. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*(1585), 10–23. http://doi.org/10.1098/rstb.2011.0190

Jablonka, E., Ginsburg, S., & Dor, D. (2012). The co-evolution of language and emotions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1599), 2152–2159. http://doi.org/10.1098/rstb.2012.0117

Jarvis, E. D., Güntürkün, O., Bruce, L., Csillag, A., Karten, H., Kuenzel, W., … Butler, A. B. (2005). Avian brains and a new understanding of vertebrate brain evolution. *Nature Reviews. Neuroscience*, *6*(February), 151–159. http://doi.org/10.1038/nrn1606

John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the Prevalence of Questionable Research Practices With Incentives for Truth Telling. *Psychological Science*, *23*(5), 524–532. http://doi.org/10.1177/0956797611430953

Johnson, N. a., Lahti, D. C., & Blumstein, D. T. (2012). Combating the Assumption of Evolutionary Progress: Lessons from the Decay and Loss of Traits. *Evolution: Education and Outreach*, *5*(1), 128–138. http://doi.org/10.1007/s12052-011-0381-y

Karten, H. J. (2015). Vertebrate brains and evolutionary connectomics: on the origins of the mammalian "neocortex." *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1684), 20150060–20150060. http://doi.org/10.1098/rstb.2015.0060

Latour, B., & Woolgar, S. (1979). *Laboratory Life. The Construction of Scientific Facts.* Princeton, New Jersey: Princeton University Press.

Lovejoy, A. O. (1936). *The great chain of being. A Study of the History of an Idea* (26th repri). Cambridge, Massachusetts, London: Harvard University Press.

Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in Psychology Research: How Often Do They Really Occur? *Perspectives on Psychological Science*, *7*(6), 537–542. http://doi.org/10.1177/1745691612460688

Mascolo, M. F. (2016). How objectivity undermines the study of personhood: Toward an intersubjective epistemology for psychological science. *New Ideas in Psychology*, 1–8. http://doi.org/10.1016/j.newideapsych.2016.11.005

McShea, D. W. (2011). Evolutionary Progress. In M. Ruse & J. Travis (Eds.), *Evolution: The First Four Billion Years* (pp. 550–557). Cambridge, Massachusetts: Harvard University Press.

Mogie, M. (2000). Historical baggage in biology: The case of "higher" and "lower" species. *BioEssays*, *22*(9), 868–869. http://doi.org/10.1002/1521-1878(200009)22:9<868::AID-BIES13>3.0.CO;2-A

Naguib, M. (2006). Animal Communication: An Overview. In *The encyclopedia of language & linguistics - Second edition* (Keith Brow, Vol. 1, pp. 276–280). Elsevier.

Nee, S. (2005). The great chain of being. *Nature*, *435*(May), 429. http://doi.org/10.1177/00221678930333006

Okanoya, K. (2017). Sexual communication and domestication may give rise to the signal complexity necessary for the emergence of language: An indication from songbird studies. *Psychonomic Bulletin & Review*, 1–5. http://doi.org/10.3758/s13423-016-1165-8

R Development Core Team, R. (2016). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.r-project.org/

Reiss, J., & Sprenger, J. (2014). Scientific Objectivity. In *The Stanford Encyclopedia of Philosophy* (Fall 2014). Edward N. Zalta (ed.). Retrieved from http://plato.stanford.edu/entries/scientific-objectivity/

Rigato, E., & Minelli, A. (2013). The great chain of being is still here. *Evolution: Education and Outreach*, *6*(18), 1–6. http://doi.org/10.1186/1936-6434-6-18

Rohrmeier, M., Zuidema, W., Wiggins, G. A., & Scharff, C. (2015). Principles of structure building in music, language and animal song. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 370(1664), 20140097. http://doi.org/10.1098/rstb.2014.0097

Ruse, M. (1996). *Monad to Man. The Concept of Progress in Evolutionary Biology*. Cambridge, Massachusetts, London, England: Harvard University Press.

Sadagopan, S., Temiz-Karayol, N. Z., & Voss, H. U. (2015). High-field functional magnetic resonance imaging of vocalization processing in marmosets. *Scientific Reports*, *5*, 10950. http://doi.org/10.1038/srep10950

Schmid, H. (1995). Improvements In Part-of-Speech Tagging With an Application To German. In *Proceedings of the ACL SIGDAT-Workshop* (pp. 1–9). Dublin, Ireland. Retrieved from http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/

Scott-Phillips, T. C. (2015). Nonhuman Primate Communication, Pragmatics, and the Origins of Language. *Current Anthropology*, *56*(1), 56–80. http://doi.org/10.1086/679674

Sereno, M. I. (2014). Origin of symbol-using systems : speech, but not sign, without the semantic urge. *Philosophical Transactions of the Royal Society B*, *369*(August), 20130303. http://dx.doi.org/10.1098/rstb.2013.0303

Slocombe, K. E., Waller, B. M., & Liebal, K. (2011). The language void: the need for multimodality in primate communication research. *Animal Behaviour*, *81*(5), 919–924. http://doi.org/10.1016/j.anbehav.2011.02.002

Ullrich, R. (2016). From "speech" to "gesture": The "oral" as norm in "language" research. *Interdisziplinäre Anthropologie Jahrbuch: Wahrnehmung*, *4*, 179–208. http://doi.org/10.1007/978-3-658-14264-3

Vaesen, K. (2014). Chimpocentrism and reconstructions of human evolution (a timely reminder). *Studies in History and Philosophy of Science Part C :Studies in History and Philosophy of Biological and Biomedical Sciences*, *45*(1), 12–21. http://doi.org/10.1016/j.shpsc.2013.12.004

Wicherts, J. M., Borsboom, D., Kats, J., & Molenaar, D. (2006). The poor availability of psychological research data for reanalysis. *The American Psychologist*, *61*(7), 726–728. http://doi.org/10.1037/0003-066X.61.7.726

Wilcox, S. (1999). The Invention and Ritualization of Language. In B. King (Ed.), *The Origins of Language: What Nonhuman Primates Can Tell Us* (pp. 351–384). Oxford: James Currey Ltd.

Zuberbühler, K. (2013). Primate Communication. In C. Lefebvre, B. Comrie, & H. Cohen (Eds.), *New Perspectives on the Origins of Language* (pp. 187–210). Amsterdam: John Benjamins

Supplementary Methods

Words used for quantifying the frequency of attributions of 'high' and 'low' in a broader sense:

category high: **advanced**[1,5] {advance, advancement}; **better**[1]; **complex**[1,6] {complexity, complexly}; **deliberate**[5] {deliberately}; **developed**[1,2]; **elaborate** {elaborately, elaboration}; **flexible**[2,5] {flexibility, flexibly}; **high**[1,2] {highly, higher, highest, high-level, higher-level, high-quality, higher-quality, higher-order}; **improved**[4]; **infinite**[5]; **intelligent**[1,3]; **intentionality**[3] {intention, intentional, intentionally}; **judicious**[3]; **metapsychology**[3] {metapsychological}; **perfect**[4] {perfection, perfectly}; **powerful**[5] {powerfully}; **precise**[4]; **progressive**[3] {progressively}; **ratchet**[7]; **reasonable**[3] {reasonably}; **remarkable**[5] {remarkably}; **rich**[5] {richly}; **sophisticated**[1,5] {sophisticate, sophisticatedly, sophistication}; **surpass**[5]; **volitional**[5] {volitionally}

category low: **ancient**[4]; **automatic**[1] {automatically}; **basic**[1]; **constrained**[4] {constraint}; **deficient**[4] {deficit; deficiency}; **finite**[4]; **fixed**[5] {fixedness}; **heritable**[5]; **immature**[4]; **imperfect**[2] {imperfection, imperfectly}; **imprecise**[2] {imprecisely, imprecision}; **inefficient**[1,2] {inefficiency}; **inflexible**; **instinct**[1] {instinctive, instinctively}; **involuntary** {involuntarily}; **less**[1] {less-productive}; **limited**[1,2]; **low**[1]; **minor**[1]; **narrow**[1]; **precursor**[1]; **premature**[4] {prematurely; prematurity}; **primitive**[1,4]; **protolanguage**[4]; **reflexive**[1,5] {reflexively}; **restrict**[2,5] {restricted, restriction, restrictive}; **rigid**[5] {rigidity, rigidly}; **rudimentary**[2,4] {rudiment}; **simple**[1]; **stereotypic**[5] {stereotype, stereotyped, stereotypical, stereotypy}; **underdeveloped**[4,5]

[1] Attributes mentioned within an open survey conducted by members of the Comparative Developmental Psychology group (FU Berlin).

[2] Attributes identified in earlier research related to sign languages used by the deaf. Historically they were used to discredit human manual, or to value human oral, forms of communication. For literature examples see (Ullrich, 2016).

[3] According to Ruse, these attributes are valued by humans because human achievements supposedly depend on them: "One thinks here particularly of intellectual capacity, broadly construed. […] even though it might be the case that these qualities are themselves contingently connected with (human-valued) organic capacities or features." (Ruse, 1996, p. 39).

[4] According to Ruse, these attributes imply direction towards an improved state – an end result: "Progress that people desire, especially when (by and large) everyone has an interest in the end result, centrs on value." (Ruse, 1996, pp. 19–20)

[5] Attributes identified by earlier research related to the 'classic view' on songbird brains. Based on a neuroanatomy that widely assumed an 'underdeveloped' bird brain which was illustrated by a large phylogenetic distance to humans (i.e. birds were viewed as closer to reptiles than to mammals),

authors frequently compared ‘heritable’, ‘instinctive’ behaviour of birds with ‘flexible’, ‘voluntarily’ traits in mammals. (Güntürkün & Bugnyar, 2016, p. 2; Jarvis et al., 2005, p. 161; Karten, 2015, p. 5)

[6] According to McShea, ‘complex’ serves as ‘stealth candidate’ for progress: “because a direct connection with progress is rarely acknowledged. But the way complexity is used in evolutionary studies suggests that it functions as a kind of code word for progress, superficially value free and therefore scientific sounding but still subtly connoting advancement. The absence of a widely known technical definition makes this usage problematic.” (McShea, 2011, p. 555)

[7] According to Ruse the ‘ratchet theory’ is opposed to Darwinian thinking and links to the notion of progress in the writings of Stephen Jay Gould, which again connect to German ‘Naturphilosophie’: “a vision of humankind that puts us above the vulgarly biological” (Ruse, 1996, p. 506).

The following R-packages were in use for data analysis in R 3.2 (R Development Core Team, 2016):
"dplyr" (Wickham & Francois, 2016),
"gdata" (Warnes et al., 2015),
"ggmap" (Kahle & Wickham, 2013),
"ggplot2" (Wickham, 2009),
"gridExtra" (Auguie, 2016),
"koRpus" (Michalke, 2016).
"lazyeval" (Wickham, 2016a),
"qdap" (Rinker, 2013),
"RColorBrewer" (Neuwirth, 2014),
"RCurl" (R Development Core Team, 2016),
"rvest" (Wickham, 2016b),
"stringi" (Gagolewski & Tartanus, 2016),
"svglite" (Wickham, Luciani, Decorde, & Lise, 2016),
"tidyr" (Wickham, 2016),
"tm" (Feinerer & Hornik, 2015),
"XML" (R Development Core Team, 2016)

Supplementary Results

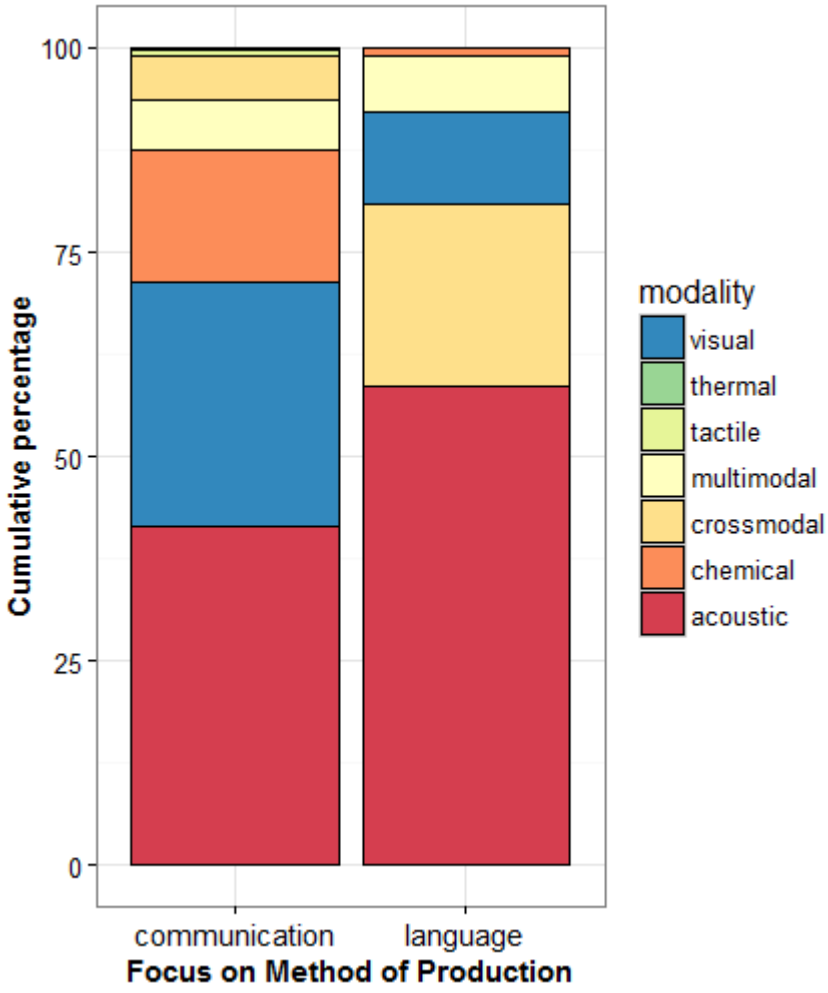

Figure S1 Comparison of the focus on modality between the corpus 'language' and 'communication'. After reading title and abstract, a decision was made if studies primarily focus on acoustic, visual, thermal, tactile or chemical communication. If there was no deliberate focus on any modality, we labelled 'crossmodal'. If there was deliberate focus on combining several modalities, we labelled 'multimodal'. In both corpora, studies mostly focused on acoustic modality.

Supplementary References

Auguie, B. (2016). GridExtra: Miscellaneous Functions for "Grid" Graphics. Retrieved from https://cran.r-project.org/package=gridExtra

Feinerer, I., & Hornik, K. (2015). tm: Text Mining Package. Retrieved from https://cran.r-project.org/package=tm

Gagolewski, M., & Tartanus, B. (2016). R package stringi: Character string processing facilities. http://doi.org/10.5281/zenodo.32557

Güntürkün, O., & Bugnyar, T. (2016). Cognition without Cortex. *Trends in Cognitive Sciences*, *xx*, 1–13. http://doi.org/10.1016/j.tics.2016.02.001

Jarvis, E. D., Güntürkün, O., Bruce, L., Csillag, A., Karten, H., Kuenzel, W., … Butler, A. B. (2005). Avian brains and a new understanding of vertebrate brain evolution. *Nature Reviews. Neuroscience*, *6*(February), 151–159. http://doi.org/10.1038/nrn1606

Kahle, D., & Wickham, H. (2013). ggmap: Spatial Visualization with ggplot2. *The R Journal*, *5*(1), 144–161. Retrieved from http://journal.r-project.org/archive/2013-1/kahle-wickham.pdf

Karten, H. J. (2015). Vertebrate brains and evolutionary connectomics: on the origins of the mammalian "neocortex." *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1684), 20150060–20150060. http://doi.org/10.1098/rstb.2015.0060

Lang, D. T., & R Development Core Team, R. (2016a). RCurl: General Network (HTTP/FTP/...) Client Interface for R. Retrieved from https://cran.r-project.org/package=RCurl

Lang, D. T., & R Development Core Team, R. (2016b). XML: Tools for Parsing and Generating XML Within R and S-Plus. Retrieved from https://cran.r-project.org/package=XML

McShea, D. W. (2011). Evolutionary Progress. In M. Ruse & J. Travis (Eds.), *Evolution: The First Four Billion Years* (pp. 550–557). Cambridge, Massachusetts: Harvard University Press.

Michalke, M. (2016). koRpus: An R Package for Text Analysis. Retrieved from http://reaktanz.de/?c=hacking&s=koRpus

Neuwirth, E. (2014). RColorBrewer: ColorBrewer Palettes. Retrieved from https://cran.r-project.org/package=RColorBrewer

R Development Core Team, R. (2016). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.r-project.org/

Rinker, T. W. (2013). qdap: Quantitative Discourse Analysis Package. Retrieved from http://github.com/trinker/qdap

Ruse, M. (1996). *Monad to Man. The Concept of Progress in Evolutionary Biology*. Cambridge, Massachusetts, London, England: Harvard University Press.

Ullrich, R. (2016). From "speech" to "gesture": The "oral" as norm in "language" research. In G. Hartung & M. Herrgen (Eds.), *Interdisziplinäre Anthropologie Jahrbuch 4/2016: Wahrnehmung* (pp. 179–208). Wiesbaden: Springer VS. http://doi.org/10.1007/978-3-658-14264-3

Warnes, G. R., Bolker, B., Gorjanc, G., Grothendieck, G., Korosec, A., Lumley, T., … Rogers, J. (2015). gdata: Various R Programming Tools for Data Manipulation. Retrieved from https://cran.r-project.org/package=gdata

Wickham, H. (2009). ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag. Retrieved from http://had.co.nz/ggplot2/book

Wickham, H. (2016). lazyeval: Lazy (Non-Standard) Evaluation. Retrieved from https://cran.r-project.org/package=lazyeval

Wickham, H. (2016). rvest: Easily Harvest (Scrape) Web Pages. Retrieved from https://cran.r-project.org/package=rvest

Wickham, H. (2016). tidyr: Easily Tidy Data with `spread()` and `gather()` Functions. Retrieved from https://cran.r-project.org/package=tidyr

Wickham, H., & Francois, R. (2016). dplyr: A Grammar of Data Manipulation. Retrieved from https://cran.r-project.org/package=dplyr

Wickham, H., Luciani, T. J., Decorde, M., & Lise, V. (2016). svglite: An "SVG" Graphics Device. Retrieved from https://cran.r-project.org/package=svglite

Publication 3: Times change, values change

Ullrich, R., & Liebal, K. (in print). Times change, values change: Criteria for attributing *language* in species comparative research [expected summer 2018]. In E. Luef & M. M. Marin (Eds.), *The talking species: Perspectives on the evolutionary, neuronal, and cultural foundations of language.* Graz: Unipress Graz-Verlag.

Times change, values change: Criteria for attributing *language* in species comparative research

Robert Ullrich, Katja Liebal
Freie Universität Berlin, Department of Psychology

Abstract

Charles Darwin's idea of a common origin of species stimulated cross-species comparative research on all kinds of phenomena, among them *language*. Research on language, however, is faced with the problem of defining the term at issue. Across times and disciplines, researchers ascribed a notoriously diverse set of properties to the faculty of language. The consequent ambiguity surrounding the term still exists, which is – as we hypothesize – the result of divergent scientific norms and historical influences. The current chapter aims to reconstruct three selected properties of language that historically had an important impact on species comparative language research, but which emerged in fact from social norms and subjective values, namely: (i) the *norm of directed progress*; (ii) the *oral norm*; and (iii) the *behavioristic norm*. The idea of *primitive* compared to *more complex* species (i), for instance, marginalized the complexity of birdsong. A narrow focus on the oral modality (ii) precluded the serious investigation of gestures in humans and non-humans. Also, excluding inner mental processes from the area of scientific knowledge (iii) disqualified non-humans from cognitive comparison. In the history of the species comparative language discourse, those value-based norms often created a narrative of human specialty by constraining the applicability of the defining properties to a narrow subset of skills. The current chapter aims to reconstruct the change of values over time, in order to point to recurring thoughts and methodological pitfalls such as sampling biases, a priori assumptions, and anthropomorphism. By consulting the history of the language discourse, it is possible to explain and reveal the aftermaths of the norms, which strongly influence current research using cross-species comparisons and consequently enter current debates about language definitions.

1. Introduction

Some scientists consider language to be the "most salient behavioural difference" and "separator" between humans and non-human animals (Wallman, 1992, p. 5). From cross-species comparative perspectives, prominent questions are: *Is human language unique?* (e.g., Smit, 2016) and *Do other species possess analogue or homologue language abilities?* (e.g., Fitch, 2017). Yet, underneath those and similar questions rests an ongoing discussion about the definition of language. The closer life sciences investigate the topic, the more a theory of *language* disintegrates. Instead of a commonly shared definition, there remain constructions with sometimes incompatible conclusions (see Botha, 2016; Wacewicz & Zywiczynski, 2015). For instance, for some scientists the communicative function of language is a side effect complementing the more important *system of thought* (e.g., Chomsky, 2011), while for others it is the other way around where the function of communication caused "language to be a vehicle of thought" (e.g., Okanoya, 2017, p. 1).

The apparent lack of unity between scientists regarding the defining properties of language is as old as the debate about the origins of language. According to Rudolf Botha, this might have its roots in (a) conflicting judgments about the theory at issue, (b) different ideas about the methods of finding evidence, and (c) divergent interpretations of the strength of evidence (Botha, 2016). Consequently, in all aspects, it is the scientist's *judgment* that is at the core of diverging assessments. Judgments, however, are not made by unbiased brains, but are underpinned by a set of implicit historical influences and social values. Those values – sometimes denoted as *social norms* – not only influence judgments on (a) to (c), but also leave their marks on the theory of biological evolution and linguistic ontology that scientists employ. For that reason, it is the aim of the current chapter to focus on social norms influencing the assessment of the design properties of language. Norms are especially potent where "social categories and the images they embed are inescapably value-laden" (Davis, 2013, p. 554). That is assumed to be the case with the species comparative language discourse, where scientists try to create a valid human self-conception with reference to a supposedly unique human characteristic, namely language. Given the entanglements between science and values or norms, it appears to be more productive to monitor them instead of combating them in their entirety. The current chapter aims to present three examples of how historical norms influenced and sometimes still influence the discourse. Unveiling those historical norms might help future empirical research in clarifying the experimental design, the formulation of questions and subsequent interpretation of data. As such, the chapter contributes to the process of scientific self-correction.

2. From scala to continuum: Teleology in brain architecture

In 1999, the primatologist Frans de Waal wrote: "the comparative aspect of comparative psychology is essentially anthropocentric: extrapolations are generally from animals to humans along a *linear progression from lower to higher* forms" (de Waal, 1999, p. 257, italics added). The direct consequences of what de Waal criticizes are spelled out for the language discourse by Martin Sereno:

> There is a powerful perennial tendency outside fields explicitly focused on evolutionary processes to think of evolution in terms of a 'Great Chain of Being' and to ignore the mosaic nature of evolution. Thus, birdsong has often been dismissed as a model of human language for the reason that monkeys seem much smarter than some birds (Sereno, 2014, p. 5).

The observation from Sereno and the phenomenon addressed by de Waal belong to a social norm whose validity in fact was supposed to belong to a bygone era. There is talk of the norm of *directed progress*, also known as *Scala Naturae* or *Great Chain of Being* (Ghiselin, 2005; Hodos & Campbell, 1969; Ruse, 1996). Behind the idea of a *Scala Naturae* there is the assumption that evolution proceeds in a linear *upward* way from a simple or *primitive* condition towards an *improved* (mostly human) state.

Although some researchers stick to the metaphor even today (e.g., Lourenço & Bacci, 2017, p. 1 "[…] with Homo sapiens putatively at the top of the scale"), this is not how evolution proceeds. Evolutionary processes are initiated by random variation with natural selection, but they are not directed towards a *most sophisticated* end state (Johnson, Lahti, & Blumstein, 2012). Indeed, as mentioned by Sereno, until the beginning of the 20th century, species were divided into *lower* and *higher* ones. Birds, at this point, are just one example for an allegedly *primitive* species among others (see Hodos & Campbell, 1969). The empirical foundation of the idea was laid in earlier works by neurologists Hughlings Jackson (1835–1911) and Ludwig Edinger (1855–1918), to name a few. Jackson was highly influenced by Herbert Spencer (1820–1903), also known as "Britain's chief prophet of Progress" (Ruse, 1996, p. 30). Following Spencer's teleological assumptions about evolution in general[154], Jackson proclaimed a "climax of nervous evolution" (Jackson, 1884, p. 591) where *higher* neurological areas – such as the cortex – suppress the function of the *lower* parts (York & Steinberg, 2011). Edinger came to similar conclusions, when he compared brains of fish, birds, amphibians and mammals. He was the first to assign the names *palæëncephalon* to *lower* brain areas and *neëncephalon* to the *higher* ones. The names were later modified into *paleocortex* and *neocortex* by Dutch neurologist Ariëns Kappers (Kappers, 1929). The prefixes *paleo* and *neo* were supposed to represent the alleged age of their evolutionary origin. Edinger identified the palæëncephalon in all species, but he found the neëncephalon only "above fish [where] it increases to that enormous organ, the cerebrum, which in man fills almost the entire skull" (Edinger, 1908, p. 438).

The typical narrative of that time leads from *primitive* reptiles to humans *at the top*. As a consequence of that narrative, scientists considered birds as incapable of any "intelligent action" because of their "purely instinctive behavior" which is "governed by emotion" (Emlen, 1948, p. 37). Similar statements can be found frequently at the beginning of the 20th century, for instance in a book from Judson Herrick, who was a comparative neurologist and publisher of the *Journal of Comparative Neurology*. He wrote: "It is everywhere recognized that birds possess highly complex instinctive

[154] "Evolution is a passage from the most simple to the most complex" (Jackson, 1884, p. 591)

endowments and that their intelligence is very limited" (Herrick, 1924, p. 213). Those claims found their way into classic textbooks such as *Principles of Animal Psychology* from Maier and Schneirla: "Birds possess an extensive repertoire of highly stereotyped activities" and "seem to behave rather stupidly" (Maier & Schneirla, 1935, pp. 235 & 478). Bird's *primitive* origin and their *simple* brains were perceived as exhibiting reflexes and instincts only[155], while "in the mammals we meet a brain which has so large a neëncephalon that we may well expect a subordination of reflexes and instincts to associative and intelligent actions" (Edinger, 1908, p. 453).

Because of this line of argumentation, birds were excluded from the mainstream of comparative language discourse especially from the 1880s to the 1950s (see Baker, 2001 for more details). Birdsong was perceived as something purely instinctive, in stark contrast to intentionally produced human language. While language "must be learned laboriously by the human child", song was considered to be a "purely innate code" hard-wired and genetically determined (Lorenz, 1949/2002, p. 74). In 1951, Otto Koehler noticed that common field guides for songbirds did not mention *subsong* (German "Jugendgesang"), because most scientists did not expect learning to be involved in song acquisition (Koehler, 1951). The birds' small and light brains were not considered to be an ecological adaption for flight, but interpreted as indicator for their *primitive* status (Emery, 2006). Some were convinced that birds and their *reptilian brains* would "lack the neural machinery for verbal communication" (MacLean, 1977, p. 159). A sampling bias towards chicken, quail or pigeons further reinforced prejudices as exemplified by a passage from psychologist David Premack: "Although the rat and pigeon may have property-identifying tags, I rather doubt that I could exploit these tags, associating different pieces of plastic with each of them, thereby teaching these nonprimates something functionally equivalent to the words *color of*, *shape of* and the like. Yet this is exactly what we have been able to do with the chimpanzee" (Premack, 1983, p. 133). Note that parrots and corvids are able to perform those tasks very well (Güntürkün & Bugnyar, 2016; Güntürkün, Ströckens, Scarf, & Colombo, 2017), but unlike primates, bird behavior was usually interpreted from a low-level perspective. While birdsong was often interpreted as mere expression of the callers' emotional state (Thorpe, 1958), primate signals were considered "referential", as means to "convey information about salient objects and events in the environment", and "allow individuals to make adaptive responses" (Hauser, 2000, p. 463; but see Liebal, Waller, Burrows, & Slocombe, 2014, p. 171 for the classical dichotomy between intentional and emotional signals in primate communication).

Back in the 60s of the 20th century, the neurologist Harvey Karten wondered how "the richness of avian behaviour" could exist without "the presence of an intact neocortex" (Karten, 2015, p. 4). But it took until the 21st century to change the perspective on the significance of birdsong drastically (Emery, 2006; Shimizu, 2009). Only very recently, a consortium recognized the *old* bird brain

[155] Edinger admittedly adds: "instincts whose perfection is so great that it has not always been possible to distinguish them from activities which are dependent upon the cortex" (Edinger, 1908, p. 451)

nomenclature as wrong, misleading, and motivated by a *norm of progress*. The Avian Brain Nomenclature Consortium decided to rename and reassess large areas of the bird brain and "eliminated all phylogeny-based prefixes (palaeo-, archi- and neo-) that erroneously implied the relative age of each subdivision" (Jarvis et al., 2005, p. 155). The new paradigm encouraged scientists to question the cognitive performance ranks across species and to proclaim gradual parallels between birdsong and human language in terms of neural circuits (Nottebohm, Stokes, & Leonard, 1976), vocal learning (Thorpe, 1958), imitation (Marler, 1970), and dialects (Baker & Michael, 1985). The failing of the *norm of progress* did now encourage an unbiased debate about what scientists believed to be the shared biological substrates of *language* (see Doupe & Kuhl, 1999). Although implicit indicators sometimes still point to the persistent presence of the norm of progress in current literature (Ullrich, Mittelbach, & Liebal, 2018), the debate no longer explicitly excludes certain species based on their phylogenetic distance to humans (Güntürkün et al., 2017). Widening the scope to non-mammalian species revealed traits in songbirds that were supposed to be uniquely human, such as critical learning periods (e.g., Bolhuis, Okanoya, & Scharff, 2010), social shaping of babbling (e.g., Goldstein, King, & West, 2003), phonology (e.g., Yip, 2013), syntactical structure (e.g., Berwick, Okanoya, Beckers, & Bolhuis, 2011), specialized brain circuits (e.g., Jarvis, 2013), and genes related to vocalizations, such as the transcription factor FoxP2 (e.g., Scharff & Petri, 2011).

In sum, there are various reasons why birdsong became a model for comparisons with human language. However, an additional important reason for the interest in birdsong might also be its vocal modality, shared with human language. The oral-vocal modality did and still does motivate researchers to assume a close relationship between human and bird utterances, which implies another social norm that is described as *oral norm* hereafter.

3. From speech to gesture: Overcoming a too narrow construction

Until the middle of the 20th century, scientists had no doubt that one decisive characteristic of *language* was its verbal nature, and more specifically, the oral modality, which has a number of important implications. The American psychologist James Coleman wrote the popular sentence: "The fish will be the last to discover water" (Coleman, 1960, p. 59). Thus, the fact that the oral-vocal modality was set as a defining feature of language was rarely noticed. Authors from the early modern period interpreted *speech* and *language* as being inextricably linked with each other (Serjeantson, 2001). The linkage entails that "language is a necessary condition of thought […] and was translated by the body machine into action [= *speech*, added by author]" (Kiriazis & Slobodchikoff, 1997, p. 365). Linguist Edward Sapir implicitly continued to share those ideas in 1921, when he published an influential textbook titled *Language: An Introduction to the Study of Speech* (Armstrong & Karchmer, 2009). In his famous article *The Origin of Speech*, Charles Hockett wrote that the "vocal-auditory channel" is the most obvious design feature, which "appear[s] so trivial that no one looking just at language would bother to note" (Hockett, 1960, p. 6). Consequently, scientists using cross-species

comparisons continued to promote an oral picture of language: "a language, if it is to achieve its full potentialities, must be a language of sounds" (Thorpe, 1958, p. 537). The implicit preference for oral-vocal utterances similarly affected research on primates:

> If one were looking for parallels with the process of human vocal learning, the most obvious place to look would be in our closest surviving relatives, the apes and monkeys. Surprisingly, no one has yet discovered a non-human primate with any facility for vocal imitation (Marler, 1970, p. 669).

Peter Marler wrote this sentence after at least six failed attempts to teach verbal utterances to various ape species (see Kellogg, 1968; Miles, 1997; Radick, 2007). To what extent those teaching attempts are informative is the matter of an ongoing debate on a methodological and semantic level (e.g., Kellogg, 1968; Lameira, 2017). However, in the 1970s those results led researchers to conclude that apes were of little use in research about vocal learning:

> In broaching the comparative investigation of vocal learning it might seem logical to study the abilities of nonhuman primates in this regard. This approach has yielded results which though interesting in themselves, are in some respects disappointing. […] Thus, apes demonstrate no great facility for vocal imitation (Marler, 1970, p. 1).
>
> Evidence on hand suggests that the socioecology of present-day non-human primates is an unsatisfactory springboard for vocal learning of any consequence (Nottebohm, 1972, p. 133).

The reasons why these two researchers preferred the communication of songbirds over primates when comparing human language to non-human forms were grounded in their reservations towards other, non-verbal means of communication. Until the 1960s those reservations also inhibited the realization of suggestions from the psychologists Robert Yerkes and Lev Vygotsky, who promoted the teaching of visual-gestural signs or symbols instead of vocalizations (Vygotsky, 1934/1986; Yerkes, 1925). To understand the delay in implementing those suggestions, reference to the separate deaf-discourse provides valuable insights. The institutional education of deaf students at the beginning of the 20th century was dominated by the so-called *oral method*. The use of *manual gestures* was mostly forbidden at schools for deaf students in Europe and North America (McDonnell & Saunders, 1993). Instead, deaf students were forced to learn oral utterances. As a result of inadequate *oral* teaching methods, students suffered dramatic deterioration of education and frequently became functional illiterates (Sacks, 1990). One reason to suppress manual and to force oral communication was a long held conviction that manual signs cannot be a *natural language* and must be *deficient* compared to oral sounds:

> Beside speech there is no other generally used universal sign system. […]. Other sign systems, like the deaf-mute language […] are either transposed, restricted or parasitic (Lotz, 1950, p. 712).
>
> Ideographic language systems, in comparison with verbal symbol systems, lack precision, subtlety, and flexibility […]. Comparatively, a verbal language is more abstract (Myklebust, 1964, p. 241).

It took decades of research by early pioneers like William Stokoe, Robert Johnson, Adam Kendon, and Scott Liddell (linguists); Edward Klima and Ursula Bellugi (psychologists); and Harlan Lane and

John van Cleve (historians) to uncover, challenge, and overcome the pre-empirical claims as expressed by Lotz and Myklebust (see Ullrich, 2016). A new generation of researchers broadened the – in their view – too narrow perspective on *language*. For them, language entails more than the oral-vocal modality. Rather, it includes gesture and body posture (Goldin-Meadow & Brentari, 2017; Kendon, 2008; McNeil, 1992). By and large, the defining properties of language became independent of modality associated with an increasing promotion for multimodal or cross-modal accounts of human language (Vigliocco, Perniss, & Vinson, 2014).

The cross-species comparative language discourse absorbed many of the ideas sketched above. However, it was only after American Sign Language was acknowledged as a natural *language*, and only after the failure of oral-vocal in contrast to alleged success of visual-gestural experiments with apes (Gardner & Gardner, 1969; Premack, 1971) that the gestural origin was reconsidered (Hewes et al., 1973). While the results of both the oral and gestural *ape language* experiments received fundamental criticism after their publication (Leavens, Bard, & Hopkins, 2010; Rivas, 2005; Terrace, 1979), the idea that the defining properties of language are independent of their modality became prevalent. Nowadays, research focuses both on oral accounts (Lameira, 2017) and on gestural accounts (Liebal, 2017), but most importantly, there is an increasing attention on an integration of multimodal or cross-modal research (Partan & Marler, 1999; Slocombe, Waller, & Liebal, 2011). In fact, the last two decades of multi- and cross-modal research have shown that the use of a certain modality is not necessarily a marker for sophisticated communication but may rather be explained by environmental conditions (Partan, Fulmer, Gounard, & Redmond, 2010; Waller, Liebal, Burrows, & Slocombe, 2013). The more the defining properties of language became independent of normative attributions about modality, the more types of animal communication could fit the concept of *language* in principle. However, some scholars strictly reject that perspective, claiming that the external features of language are ontologically different from the inner mental functions (Bolhuis, Tattersall, Chomsky, & Berwick, 2014). This demands a closer look on the presumed dichotomy of *internal* and *external*, to understand the historical roots of the *cognitive turn*, which was meant to overcome the *behavioristic norm*.

4. From external behavior to internal processing: The example of intentionality

The term *intentionality* has made it to the top of the most discussed defining properties of *language* (see, e.g., Liebal et al., 2014; Townsend et al., 2016). Its usage started with Franz Brentano (1838–1917) (Brentano, 1874/2009) who reintroduced the term from scholastic philosophy. Paul Grice (1957, 1969) brought it forward to the language discourse, followed by Daniel Dennett (Dennett, 1971, 1983) who transferred the concept into the species comparative discourse.

The term *intentionality* was introduced in the comparative sciences during the early 20th century, when Behaviorism dominated the discourse on language in the United States. The *behavioristic norm*

(see Graham, 2015 – here denoted as *doctrine*) arose as a countermovement to 19th-century *animal psychology*, where scientists ascribed *purposes*, *feelings* and *abstract thought* to their study species (e.g., Romanes, 1879; Radick, 2007). For critics of that *anthropomorphic mentalism*, the anecdotal approach of animal psychology was nothing other than unfounded metaphysics (Jamieson & Bekoff, 1992). In order to abandon the *metaphysical* mind-matter-dualism and in hopes of unifying science towards a positivist orientation, behaviorists called for rigorously controlled experiments, where the "facts to be observed are external phenomena, physical occurrences in the objective world" (Morgan, 1903, p. 48). For researchers driven by the *behavioristic norm*, any science that does not focus on quantifiable entities (e.g., *introspective psychology*) is not a "proper science" (Radick, 2016, p. 73). In such a climate, a term like *intentionality*, defined by Brentano as *hallmark of the mental* (Jacob, 2014), was difficult to use. Still, it was used by European ethologists because for them descriptions of behavior freed from *mentalistic* vocabulary were hard to sustain. When explaining and predicting behavior, ethologists like Oskar Heinroth (1871–1945), Konrad Lorenz (1903–1989), and Nikolaas Tinbergen (1907–1988) introduced the term "intention movement" (German "Intentionsbewegung") (Heinroth, 1910/1990, p. 680; Lorenz, 1937, p. 292; Tinbergen, 1939, p. 223). According to the definition of Heinroth, the term *intention movement* serves as a methodological stance, where "the trained behaviour student can derive from their study a knowledge of what the animal is intending to do in the next few moments" (Daanje, 1950, p. 48). However, the understanding of *intention* at this stage of the debate was very different from what Brentano (1874/2009) had in mind when he revived it. As Lorenz once clarified, the fulfillment of an intention movement follows a "biological purpose" (Lorenz, 1937, p. 292). The beating heart has the *biological purpose* to keep the body alive. The heart, however, does not have a *psychological intention* to do so. The ethologists at the beginning of the 20th century did not use the term *intention* in the sense of Brentano's coinage of an inner mental state. For them, it was a method of explaining behavior as *biologically purposeful* (see Millikan, 1997).

The ethologist Peter Marler was clearly a direct descendant from ethologists with behavioristic influence. In his article *The logical analysis of animal communication*, he warns "about the dangers of the introspective method in animal studies" that comes with "anthropocentric preconceptions" (Marler, 1961, p. 297). In his view, animal communication has to be described as behavior in "objective terms" where "semantics are of doubtful value" (Marler, 1961, p. 299). In an earlier paper, Marler defines language as "a means of communication between individuals, by means of sound signals" (Marler, 1956, p. 245). As for many ethologists, Marler believes that language is first of all a communicative system, whereas research must transfer attention "from pragmatics to syntactics to consider the physical nature of some of the signals used" (Marler, 1961, p. 309). Those statements might be reminiscent of the American structuralist Leonard Bloomfield. His linguistic theory was also motivated by the *behavioristic norm* and led to the conviction that objective research on language

must exclusively concentrate on form, including phonology, syntax, or morphology (Bloomfield, 1943; Levelt, 2013).

During the 1970s, Peter Marler started a collaborative project with Dorothy Cheney and Robert Seyfarth to investigate the alarm calls of vervet monkeys. Given his earlier remarks on the value of comparative research, the first sentence of their paper in *Science*, outlining the results from the collaboration, may come as a surprise: "A central but neglected issue in the study of animal communication is that of semantics" (Seyfarth, Cheney, & Marler, 1980, p. 801). Semantics, which he earlier denoted of "doubtful value" (Marler, 1961, p. 299), is now at the centre of the narrative. However, Marler did not turn into an anti-behaviorist. In principle, he opens up the possibility for unobservable mental representations, but he still uses behavioristic vocabulary and sticks to the study of *objective physical phenomena*. It is Daniel Dennett, philosopher and proponent of a concept of *intentionality* in the tradition of Brentano, who celebrates Seyfarth, Cheney and Marler as "new ethologists, having cast of the straightjacket of behaviourism" (Dennett, 1983, p. 343). It is he who exploits the vast potential of the study for his own *intentional system theory*. Dennett borrows his theoretical foundation from Brentano and Grice. He uses *intentionality* in the philosophical rich sense as a form of "directedness of the mind towards a content or object" (Glock, 2001). As Millikan (1997, p. 194) once explained:

> The difference between merely biological purposes and intentional purposes is that in the latter case the animal's biological purposes are implemented via the manufacture and use of inner representations.

Consequently, Dennett experiences research with vervet monkeys as an opportunity to discuss *representations*, *beliefs*, *desires* and the like in animals. In his opinion, reference to inner mental realities is needed to sufficiently explain and predict the vervets' behavior and to answer questions about language and communication in general. Dennett was not the first to introduce *intentionality* in the philosophically rich sense into species comparative research (see Révész, 1944; von Glasersfeld, 1974, 1976), but his writings fell on fertile ground at a time when the *cognitive turn* in linguistics and psychology was well underway. The introduction of *intentionality* achieved further support from psychologists like David Premack (Premack & Woodruff, 1978) and Michael Tomasello (1985), as well as from the philosopher John Searle (Searle, 1984). Robbins Burling summarized the consequence of the *cognitive shift* for the language discourse: "Given that language is inseparably bound up with human cognition, the most promising place to look for the antecedents of language is in primate cognitive abilities" (Burling, 1993, p. 25). Hence, questions about mental phenomena like *reference* (Sievers & Gruber, 2016), *recursion* (Martins, 2012), and *deception* (Oesch, 2016) became the subject of cross-species approaches in language research, with research about *intentionality* representing just one of many candidates serving as a potential defining property of language.

Peter Marler, like other ethologists influenced by behaviorism, never stepped back from his so-called *objective research approach*. Thus, he continues to use terms like *phonological syntax* instead of *syntax*, *functional reference* instead of *reference*, and explains behavior from the perspective of *biological purpose* instead of *psychological intention*. Marler justifies his position by explaining that "the role of the many dimensions of mindfulness still remains unclear" because of the impossibility of "introspection" and a lack of "appropriate experiments" (Marler, 2000, p. 32). Ethologists in behavioristic tradition nowadays use the term *intentionality* as means of a *biologically purposive behavior* (e.g., Vail, Manica, & Bshary, 2013). In contrast, cognitive scientists often use the term *intentionality* in the philosophically rich sense as *psychologically purposive behavior*. The *behavioristic norm* that banished the mind from *objective science* deceased, but the divergent use of the term *intentionality* still draws conclusions about its historical background. Once used to describe external behavior, it is deployed frequently nowadays to explain internal processing and thereby considered an integral defining property of language.

5. Conclusion

We have highlighted the influence of value-driven norms for defining the term *language* in the discourse involving species comparative approaches to language evolution. While the examples presented show how a theory of language was restricted or modified by various values throughout the 20th century, the use of values in science among all disciplines and all times is pervasive, as a body of classic studies (Feyerabend, 1975; Kuhn, 1977; Latour & Woolgar, 1979) and some recent publications (Davis, 2013; Douglas, 2016; Elliott & McKaughan, 2009) testify. However, while it is easy to identify values in research papers retrospectively, it becomes more complicated for contemporary publications because of their implicit character. Certainly, the history of the discourse can teach us that norms still govern recent language definitions. Those norms frame the narrative of publications and constitute the theoretical basis for defining properties as associated with *language*, as suggested by Ray Jackendoff (2010): "Your theory of language evolution depends on your theory of language". This chapter adds to this notion that the theories themselves are influenced by subjective norms. Yet, norms and values must not be immediate indicators for *bad science*. Instead, they can motivate scientists to choose a certain study species, to design their experimental procedures, to use a specific vocabulary, and to weigh the evidence found (Douglas, 2016). As shown in the current chapter, values might change over time, but they cannot be excluded from scientific practice. Also, they do not have to be excluded, as values might expose alternative answers to questions where empirical evidence is scarce. *Good science* is not to deny subjective influence on scientific practice, but to bring that influence to light. That requires making transparent the individual scientific background of researchers and their personal motivation for the topic. As long as such details are not considered, meta-research about those issues is needed to uncover current norms influencing the discourse and to understand the latest attempts of answering the question: *What is language?*

References

Armstrong, D., & Karchmer, M. (2009). William C. Stokoe and the study of signed languages. Sign Language Studies, 9, 389–397.

Baker, M. C. (2001). Bird song research: The past 100 years. Bird Behavior, 14, 3–50.

Baker, M. C., & Michael, A. (1985). The biology of bird-song dialects. Behavioral & Brain Sciences, 8, 85–133.

Berwick, R. C., Okanoya, K., Beckers, G. J. L., & Bolhuis, J. J. (2011). Songs to syntax: The linguistics of birdsong. Trends in Cognitive Sciences, 15, 113–21.

Bloomfield, L. (1943). Meaning. Monatshefte für deutschen Unterricht, 35, 101–106.

Bolhuis, J. J., Okanoya, K., & Scharff, C. (2010). Twitter evolution: Converging mechanisms in birdsong and human speech. Nature Reviews Neuroscience, 11, 747–59.

Bolhuis, J. J., Tattersall, I., Chomsky, N., & Berwick, R. C. (2014). How could language have evolved? PLoS Biology, 12, e1001934.

Botha, R. (2016). Language evolution. Cambridge: Cambridge University Press.

Brentano, F. (1874/2009). Psychology from an empirical standpoint. New York: Routledge.

Burling, R. (1993). Primate calls, human language, and nonverbal communication. Current Anthropology, 34, 25–37.

Chomsky, N. (2011). Language and other cognitive systems: What is special about language? Language Learning and Development, 7, 263–278.

Coleman, J. C. (1960). Personality dynamics and effective behavior. Chicago: Scoth Foresman and Company.

Daanje, A. (1950). On locomotory movements in birds and the intention movements derived from them. Behaviour, 3, 48–98.

Davis, J. E. (2013). Social science, objectivity, and moral life. Society, 50, 554–559.

de Waal, F. B. M. (1999). Anthropomorphism and anthropodenial: Consistency in our thinking about humans and other animals. Philosophical Topics, 27, 255–280.

Dennett, D. C. (1971). Intentional systems. The Journal of Philosophy, 68, 87–106.

Dennett, D. C. (1983). Intentional systems in cognitive ethology: The "Panglossian paradigm" defended. Behavioral and Brain Sciences, 6, 343.

Douglas, H. (2016). Values in science. In P. Humphreys (Ed.), The Oxford handbook of philosophy of science (pp. 609–632). New York: Oxford University Press.

Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: Common themes and mechanisms. Annual Review of Neuroscience, 22, 567–631.

Edinger, L. (1908). The relations of comparative anatomy to comparative psychology. The Journal of Comparative Neurology and Psychology, 18, 437–457.

Elliott, K. C., & McKaughan, D. J. (2009). How values in scientific discovery and pursuit alter theory appraisal. Philosophy of Science, 76, 598–611.

Emery, N. J. (2006). Cognitive ornithology: The evolution of avian intelligence. Philosophical Transactions of the Royal Society of London. Series B, 361, 23–43.

Emlen, J. T. (1948). Importance of the study of bird behavior. The American Biology Teacher, 10, 36–41.

Feyerabend, P. (1975). Against method. London: Verso.

Fitch, T. W. (2017). Empirical approaches to the study of language evolution. Psychonomic Bulletin & Review, 24, 3-33.

Gardner, A. R., & Gardner, B. T. (1969). Teaching sign language to a chimpanzee. Science, 165, 664–672.

Ghiselin, M. T. (2005). The Darwinian revolution as viewed by a philosophical biologist. Journal of the History of Biology, 38, 123–136.

Glock, H. J. (2001). Intentionality and language. Language and Communication, 21, 105–118.

Goldin-Meadow, S., & Brentari, D. (2017). Gesture, sign and language: The coming of age of sign language and gesture studies. Behavioral and Brain Sciences, 40, e46.

Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. Proceedings of the National Academy of Sciences USA, 100, 8030–8035.

Graham, G. (2015). Behaviorism. In E. N. Zalta (Ed.), The Stanford encyclopedia of philosophy. Retrieved from http://plato.stanford.edu/archives/spr2015/entries/behaviorism/

Güntürkün, O., & Bugnyar, T. (2016). Cognition without cortex. Trends in Cognitive Sciences, 20, 291–303.

Güntürkün, O., Ströckens, F., Scarf, D., & Colombo, M. (2017). Apes, feathered apes, and pigeons: Differences and similarities. Current Opinion in Behavioral Sciences, 16, 35–40.

Hauser, M. D. (2000). A primate dictionary? Decoding the function and meaning of another species' vocalizations. Cognitive Science, 24, 445–475.

Heinroth, O. (1990). Beiträge zur Biologie, namentlich Ethologie und Psychologie der Anatiden [1910]. Wien: Verein für Ökologie und Umweltforschung.

Herrick, J. (1924). Neurological foundations of animal behavior. New York: Henry Holt and Company.

Hewes, G. W., Andrew, R. J., Carini, L., Choe, H., Gardner, R. A., Kortlandt, A., … Wescott, R. W. (1973). Primate communication and the gestural origin of language. Current Anthropology, 14, 5–24.

Hockett, C. F. (1960). The origin of speech. Scientific American, 203, 88–111.

Hodos, W., & Campbell, C. G. B. (1969). Scala naturae: Why there is no theory in comparative psychology. Psychological Review, 76, 337–350.

Jackendoff, R. (2010). Your theory of language evolution depends on your theory of language. In R. K. Larson, V. Deprez, & H. Yamakido (Eds.), The evolution of human language (pp. 63–72). New York: Cambridge University Press.

Jackson, J. H. (1884). The Croonian lectures on evolution and dissolution of the nervous system. The British Medical Journal, 1, 591–593.

Jacob, P. (2014). Intentionality. In E. N. Zalta (Ed.), The Stanford encyclopedia of philosophy. Retrieved from https://plato.stanford.edu/entries/intentionality/

Jamieson, D., & Bekoff, M. (1992). On aims and methods of cognitive ethology. PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association, 2, 110–124.

Jarvis, E. D. (2013). Evolution of brain pathways for vocal learning in birds and humans. In J. J. Bolhuis & M. Everaert (Eds.), Birdsong, speech, and language: Exploring the evolution of mind and brain (pp. 63–107). Cambridge: MIT Press.

Jarvis, E. D., Güntürkün, O., Bruce, L., Csillag, A., Karten, H., Kuenzel, W., … Butler, A. B. (2005). Avian brains and a new understanding of vertebrate brain evolution. Nature Reviews Neuroscience, 6, 151–159.

Johnson, N. A., Lahti, D. C., & Blumstein, D. T. (2012). Combating the assumption of evolutionary progress: Lessons from the decay and loss of traits. Evolution: Education and Outreach, 5, 128–138.

Kappers, C. U. A. (1929). The evolution of the nervous system in invertebrates, vertebrates and man. Haarlem: Bohn.

Karten, H. J. (2015). Vertebrate brains and evolutionary connectomics: On the origins of the mammalian "neocortex". Philosophical Transactions of the Royal Society B, 370, pii: 20150060.

Kellogg, W. N. (1968). Communication and language in the home-raised chimpanzee. Science, 162, 423–427.

Kendon, A. (2008). Some reflections on the relationship between "gesture" and "sign". Gesture, 8, 348–366.

Kiriazis, J., & Slobodchikoff, C. N. (1997). Anthropocentrism and the study of animal language. In R. W. Mitchell, N. S. Thompson, & H. L. Miles (Eds.), Anthropomorphism, anecdotes, andaAnimals (pp. 365–369). Albany: SUNY Press.

Koehler, O. (1951). Der Vogelgesang als Vorstufe von Musik und Sprache. Journal für Ornithologie, 93, 3–20.

Kuhn, T. S. (1977). Objectivity, value judgment, and theory choice. In T. S. Kuhn (Ed.), The Essential Tension: Selected Studies in Scientific Tradition and Change (pp. 320–339). Chicago: University of Chicago Press.

Lameira, A. R. (2017). Bidding evidence for primate vocal learning and the cultural substrates for speech evolution. Neuroscience and Biobehavioral Reviews, 83, 429-439.

Latour, B., & Woolgar, S. (1979). Laboratory life. The construction of scientific facts. Princeton: Princeton University Press.
Leavens, D. A., Bard, K. A., & Hopkins, W. D. (2010). BIZARRE chimpanzees do not represent "the chimpanzee". The Behavioral and Brain Sciences, 33, 100–1.
Levelt, W. (2013). A history of psycholinguistics: The pre-Chomskyan era. Oxford: Oxford University Press.
Liebal, K. (2017). Gestural communication. In A. Fuentes (Ed.), The international encyclopedia of primatology (pp. 466-467). New Jersey: Wiley-Blackwell.
Liebal, K., Waller, B. M., Burrows, A. M., & Slocombe, K. E. (2014). Primate communication: A multimodal approach. New York: Cambridge University Press.
Lorenz, K. (1937). Über die Bildung des Instinktbegriffes. Die Naturwissenschaften, 19, 289–300.
Lorenz, K. (2002). King Solomon's ring [1949]. University Paperbacks. London: Routledge.
Lotz, J. (1950). Speech and language. Journal of the Acoustical Society of America, 22, 712–717.
Lourenço, J., & Bacci, A. (2017). Human-specific cortical synaptic connections and their plasticity: Is that what makes us human? PLoS Biology, 15, 1–5.
MacLean, P. (1977). An evolutionary approach to brain research on prosematic (nonverbal) behavior. In J. Rosenblatt & B. R. Komisaruk (Eds.), Reproductive Behavior and Evolution (pp. 137–164). New York: Springer.
Maier, N. R. F., & Schneirla, T. C. (1935). Principles of animal psychology. London: McGraw-Hill Book Company.
Marler, P. (1956). The voice of the chaffinch and its function as language. Ibis, 98, 231–261.
Marler, P. (1961). The logical analysis of animal communication. Journal of Theoretical Biology, 1, 295–317.
Marler, P. (1970). Birdsong and speech development: Could there be parallels? American Scientist, 58, 669–673.
Marler, P. (2000). Origins of music and speech: Insights from animals. In N. L. Wallin, B. Merker, & S. Brown (Eds.), The origins of Music (pp. 31–48). Cambridge: MIT Press.
Martins, D. M. (2012). Distinctive signatures of recursion. Philosophical Transactions of the Royal Society B, 367, 2055–2064.
McDonnell, P., & Saunders, H. (1993). Sit on your hands. In R. Fischer & H. Lane (Eds.), Looking back (pp. 255–260). Hamburg: Signum Verlag.
McNeil, D. (1992). Hand and mind. What gestures reveal about thought. Chicago: The University of Chicago Press.
Miles, H. L. (1997). Anthropomorphism, apes, and language. In R. W. Mitchell, N. S. Thompson, & H. L. Miles (Eds.), Anthropomorphism, anecdotes, and animals (pp. 383–404). Albany: SUNY Press.
Millikan, R. G. (1997). Varieties of purposive behavior. In R. W. Mitchell, N. S. Thompson, & H. L. Miles (Eds.), Anthropomorphism, anecdotes, and animals (pp. 189–197). Albany: SUNY Press.
Morgan, L. C. (1903). An introduction to comparative psychology. London: Walter Scott.
Myklebust. (1964). The psychology of deafness: Sensory deprivation, learning, and adjustment. New York: Grune & Stratton.
Nottebohm, F. (1972). The origins of vocal learning. American Naturalist, 106, 116–140.
Nottebohm, F., Stokes, T. M., & Leonard, C. M. (1976). Central control of song in the canary, Serinus canarius. The Journal of Comparative Neurology, 165, 457–86.
Oesch, N. (2016). Deception as a derived function of language. Frontiers in Psychology, 7, 1–7.
Okanoya, K. (2017). Sexual communication and domestication may give rise to the signal complexity necessary for the emergence of language: An indication from songbird studies. Psychonomic Bulletin & Review, 24, 106–110.
Partan, S. R., Fulmer, A. G., Gounard, M. A. M., & Redmond, J. E. (2010). Multimodal alarm behavior in urban and rural gray squirrels studied by means of observation and a mechanical robot. Current Zoology, 56, 313–326.
Partan, S. R., & Marler, P. (1999). Communication goes multimodal. Science, 283, 1272–1273.
Premack, D. (1971). Language in chimpanzee? Science, 172, 808–822.
Premack, D. (1983). The codes of man and beasts. Behavioral & Brain Sciences, 6, 125–167.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? Behavioral and Brain Sciences, 4, 515–526.

Radick, G. (2007). The simian tongue: The long debate about animal language. Chicago: The University of Chicago Press.

Radick, G. (2016). The unmaking of a modern synthesis: Noam Chomsky, Charles Hockett, and the Politics of Behaviorism, 1955–1965. Isis, 107, 49–73.

Révész, G. (1944). The language of animals. The Journal of General Psychology, 30, 117–147.

Rivas, E. (2005). Recent use of signs by chimpanzees (Pan troglodytes) in interactions with humans. Journal of Comparative Psychology, 119, 404–417.

Romanes, G. J. (1879). Intellect in brutes. Nature, 20, 122–125.

Ruse, M. (1996). Monad to man. Cambridge: Harvard University Press.

Sacks, O. (1990). Seeing voices: A journey into the world of the deaf. New York: Harper Perennial.

Scharff, C., Petri, J. (2011). Evo-devo, deep homology and FoxP2: Implications for the evolution of speech and language. Philosophical Transactions of the Royal Society of London. Series B, 366, 2124–40.

Searle, J. R. (1984). Minds, brains, and science. Cambridge: Harvard University Press.

Sereno, M. I. (2014). Origin of symbol-using systems : Speech, but not sign, without the semantic urge. Philosophical Transactions of the Royal Society B, 369, 20130303.

Serjeantson, R. (2001). The passions and animal language, 1540 - 1700. Journal of the History of Ideas, 62, 425–444.

Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980). Monkey responses to three different alarm calls: Evidence of predator classification and semantic communication. Science, 210, 801–803.

Shimizu, T. (2009). Why can birds be so smart? Background, significance, and implications of the revised view of the avian brain. Comparative Cognition & Behavior Reviews, 4, 103–115.

Sievers, C., & Gruber, T. (2016). Reference in human and non-human primate communication: What does it take to refer? Animal Cognition, 19, 759–768.

Slocombe, K. E., Waller, B. M., & Liebal, K. (2011). The language void: The need for multimodality in primate communication research. Animal Behaviour, 81, 919–924.

Smit, H. (2016). The transition from animal to linguistic communication. Biological Theory, 11, 158–172.

Terrace, H. S. (1979). Is problem-solving language? Journal of Experimental Analysis Of Behavior, 31, 161–175.

Thorpe, W. H. (1958). The learning of song patterns by birds, with especial reference to the song of the chaffinch Fringilla coelebs. Ibis, 100, 535–570.

Tinbergen, N. (1939). On the analysis of social organization among vertebrates, with special reference to birds. The American Midland Naturalist, 21, 210–234.

Tomasello, M., George, B. L., Kruger, A. C., Jeffrey, M., & Evans, A. (1985). The development of gestural communication in young chimpanzees. Journal of Human Evolution, 14, 175–186.

Townsend, S., Koski, S. E., Byrne, R. W., Slocombe, K. E., Bickel, B., Boeckle, M., … Manser, M. B. (2016). Exorcising Grice's ghost: An empirical approach to studying intentional communication in animals. Biological Reviews, 92, 1427-1433.

Ullrich, R. (2016). From "speech" to "gesture": The "oral" as norm in "language" research. In G. Hartung & M. Herrgen (Eds.), Interdisziplinäre Anthropologie (pp. 179–208). Wiesbaden: Springer.

Ullrich, R., Mittelbach, M., & Liebal, K. (2018). Scala naturae: The impact of historical values on current "evolution of language" discourse. Journal of Language Evolution, 2017, 3, 1-12.

Vail, A. L., Manica, A., & Bshary, R. (2013). Referential gestures in fish collaborative hunting. Nature Communications, 4, 1765.

Vigliocco, G., Perniss, P., & Vinson, D. (2014). Language as a multimodal phenomenon: Implications for language learning, processing and evolution. Philosophical Transactions of the Royal Society of London. Series B, 369, 20130292.

von Glasersfeld, E. (1974). Signs, communication, and language. Journal of Human Evolution, 3, 465–474.

von Glasersfeld, E. (1976). The development of language as purposive behavior. Annals of the New York Academy of Sciences, 280, 212–226.

Vygotsky, L. (1934/1986). Thought and language. Cambridge: MIT Press.
Wacewicz, S., Zywiczynski, P. (2015). From the narrow to the broad: Multiple perspectives on language evolution. Theoria et Historia Scientiarum, 11, 5–18.
Waller, B. M., Liebal, K., Burrows, A. M., & Slocombe, K. E. (2013). How can a multimodal approach to primate communication help us understand the evolution of communication? Evolutionary Psychology, 11, 538–49.
Wallman, J. (1992). Aping language. New York: Cambridge Univ Press.
Yerkes, R. (1925). Almost Human. London: J. Cape.
Yip, M. (2013). Structure in human phonology and in birdsong: A phonologist's perspective. In J. J. Bolhuis & M. Everaert (Eds.), Birdsong, speech, and language (pp. 181–208). Cambridge: MIT Press.
York, G. K., & Steinberg, D. A. (2011). Hughlings Jackson's neurological ideas. Brain, 134, 3106–3113.

Publication 4: Introducing intention

Ullrich, R., Mittelbach, M., & Liebal, K. (submitted March 2018). Introducing *intention*: How an idea has spread within species-comparative science. *Animal Cognition*

Please note, the article was originally submitted to Animal Cognition in March 2018. After rejection it was re-submitted to the Journal of Comparative Psychology in October 2018 under the title: "Disseminating intention: How a term has spread within cross-species comparative research". The here printed version is a pre-refeered version from March 2018.

Introducing *intention*: How an idea has spread within cross-species comparative science

Robert Ullrich[1*], Moritz Mittelbach[2], Katja Liebal[1]

[1] Department of Education and Psychology, Comparative Developmental Psychology, Freie Universität Berlin, Habelschwerdter Allee 45

[2] Department of Botany, Ecology of Plants, Freie Universität Berlin, Altensteinstr 6, Berlin 14195, Germany

[*]Corresponding author: robert.ullrich@fu-berlin.de; +49 30 838 57 690

Abstract

Advances in information technology and ever-increasing volumes of digitised bibliographic data allow for new meta-scientific approaches. With them, the current study takes a holistic view of cross-species comparative research and investigates the introduction of the term *intention* as representative of the so called 'cognitive revolution'. All references from 653 articles are used to analyse a citation network, covering the period 1948-2017. The analysis visualises and identifies prominent articles in the scientific debate and locates them structurally on a map. In addition, each article is categorised in terms of the *school of thought*, its position within the discourse (e.g. opposing, supporting), the order of intentionality (e.g. 1st or 2nd-order), and the species under consideration. By employing a mixed-method-approach, which combines qualitative and quantitative methods, we could identify two divergent schools of thought (a cognitive and a behaviouristic). Both schools introduced the term *intention* mostly independently from each other and show little overlap in citation habits. Both notions of *intention* have influenced comparative science until today. However, while the term finds limited application in various schools, only in connection with more cognitive approaches has it enjoyed a successful career, as indicated by the increasing number of articles in which it is employed. Here, most controversy does not surround the concept of *intention* itself, but its order. Furthermore, taking account of which species are investigated could reveal a pronounced primate bias in past discourse. Articles on non-primate species that use the term *intention* in the cognitive sense are markedly outnumbered by those on primates.

Introduction

Animal Cognition is celebrating its 20th anniversary. While this may be a respectable age for a journal, it is a brief timespan in terms of the history of science. More than just the name of a journal, *Animal Cognition* is also a largely empirical scientific endeavour, whose practical foundations rest and rely on theoretical arguments and assumptions – for instance regarding the nature of mind or the essence of communication (Andrews, 2014). Before scientists began to use the term *cognitive system* and to design experiments for the purpose of its investigation, they first and foremost had to accept that non-human animals are minded, that is, cognitively endowed. This consideration required an epistemological shift, a change in the theory of science. The present article is born out of an interest in the nature of this epistemological shift, which led scientists to apply terms like 'cognition', 'mind', and 'intention' in discussions of non-human animals. There exist a number of qualitative studies addressing this shift, which is sometimes referred to as the 'cognitive revolution' (Greenwood, 1999; Miller, 2003; Proctor & Vu, 2006; Sperry, 1993; Watrin & Darwich, 2012). While some prefer to call it a 'counter-revolution' (Miller, 2003, p. 141) or dismiss the term 'revolution' at all (e.g. Hobbs & Burman, 2009; Leahey, 1992; Mandler, 2002), most agree that the shift dates back to the 1950s and 1960s (Cromwell & Panksepp, 2011, p. 2028).

However, the current study adopts a more quantitative perspective on this shift, in order to supplement a rich body of qualitative work. This meta-analysis seeks to investigate the introduction of one term–intention– that is seen as representative of the new *cognitive* vocabulary, creating a citation map to track its spread within a corpus of 653 articles. We further aim to quantify the ongoing controversy surrounding its introduction. In the last decades, much bibliographic data has been digitalised and stored, together with the corresponding metadata (i.e. keywords, affiliations, references). Advances in information technology allow for the collection of large datasets, facilitate new mining procedures and consequently enable, for instance, the construction of citation networks. These developments have fostered the growing field of meta-research (Ioannidis, Fanelli, Dunne, & Goodman, 2015). To date, however, quantitative perspectives in philosophy and history of science are scarce. One of the few studies that has attempted to quantify the change from behaviourism to cognitive science was a bibliometric study carried out by Friman and colleagues (Friman, Allen, Kerwin, & Larzelere, 1993). Counting citations in four leading psychological journals for the years 1979-1988, they found "…an increase in citations to core journals in cognitive psychology…," but no "…corresponding decreases in citations to core journals in behavioural psychology…" (Friman et al., 1993, p. 661). In another meta-analysis, Robins and colleagues (1999) operationally defined three measures of 'prominence' ([i] frequency of publication; [ii] number of dissertations; [iii] frequency of citation) and compared four sub-disciplines of psychology (the 'behavioural-,' 'cognitive-,' 'psychoanalytic-,' and 'neuroscientific-school[s]') for the period 1950-1997 (Robins, Goslin, & Craik, 1999). Based on their analysis, they charted changes in the prominence of the four schools according

to variables [i]-[iii] and concluded that "sometime during the 1970s, the prominence of behavioural psychology gave way to the ascension of cognitive psychology" (Robins et al., 1999, p. 124). In 2004, the same research group furthered this investigation and extended their focus to the year 2002. Again they charted the changes for prominence depending on the variables [i]-[iii] and found "empirical evidence for a cognitive revolution" (Tracy, Robins, & Gosling, 2004, p. 116). Finally, Virues-Ortega and Pear (2014) text-mined all available books in English via Google Ngram for terms like 'behaviour' (as representative of behaviouristic vocabulary) and 'mind' (as representative of cognitive vocabulary) for the years 1900-2008 (Virues-Ortega & Pear, 2014). Their results suggest that "the rise of cognitivism did not start until well past the 1970s" (Virues-Ortega & Pear, 2014, p. 28).

All four studies found evidence for a shift of research traditions in psychology, such that the number of articles or the use of vocabulary referring to cognitive concepts increased. In the current study, we want to focus on how, once underway, this shift proceeded. In concrete terms we want to track the introduction of a new term into the cross-species comparative discourse (i.e. research across different species). We choose to focus on cross-species comparative research because it was the comparisons of different species' behaviour that initiated the empirical endeavour of *cognitive* research in non-humans in the first place (Rosati, Wobber, Hughes, & Santos, 2014). We are also interested in cross-species comparative research because it comprises research from disciplines beyond psychology. This is important, as the cognitive shift also affected other disciplines such as linguistics (Radick, 2016) and biology (Jamieson & Bekoff, 1992). While quantitative studies have provided evidence for a paradigm shift, they could not facilitate an understanding of the processes governing the further evolution of the discourse. To better understand the intellectual structure of science and to visualise the formation of a new field, we used the available meta-information to look back on a specific fraction of the species-comparative discourse. In order to identify leading figures, prominent publications and roots of discussions, we conducted the present case study on *intentionality*.

Why do we focus on *intention*? The *Stanford Encyclopaedia of Philosophy* defines the term *intentionality* as mental capacity that comprises "the power of minds to be about, to represent, or to stand for, things, properties and states of affairs." (Jacob, 2014). We focus on this specific term because we suggest that it serves as an illustrative example of how the idea of cognition in general could have spread within the scientific community (cf. Ullrich & Liebal, in print). Also, searches for the technical term *intention* will yield more relevant hits than searches for more general terms such as 'mind' or 'cognition'.

The concept of intentionality in the philosophical/psychological sense was revived by philosopher Franz Brentano (1838-1917), for whom it comprised the exclusive characteristic of any mental phenomena (Brentano, 2009, p. 68 [1874]). For Brentano who wanted to describe subjective, mental phenomena in *objective* terms; 'mind' was not an *object*, so he called it an 'in-existent state', i.e. an intentional state (Brentano, 2009, p. 68 [1874]). For critics like Lloyd Morgan (1852 - 1936) and John

Watson (1878 - 1958), this was crude metaphysical mind-matter-dualism. They called for rigorously controlled experiments, where the "facts to be observed are external phenomena, physical occurrences in the objective world" (Morgan, 1903, p. 48). For Watson, "introspective psychology" was not a "proper science" (Radick, 2016, p. 73). Morgan and Watson were early proponents of a movement that started at the end of the nineteenth century and fully unfolded in the 50s of the twentieth century: behaviourism.

Behaviourism spread among other disciplines in psychology (Watrin & Darwich, 2012), linguistics (Radick, 2016) and biology (Jamieson & Bekoff, 1992). Behaviourism was most influential in the United States (Mandler, 2002), and as a result the term *intention* enjoyed little currency at the time. In Europe, however, ethologists were less cautious in employing terminology that could convey a mental connotation. At the beginning of the 20th century, some of them introduced the term *intention movement* for methodological reasons, to sustain predictions and explanations when observing non-human animals in the wild (Heinroth, 1990 [1910], p. 680; Lorenz, 1937, p. 292; Tinbergen, 1939, p. 223). The current case study tracks the distribution of the term *intention* from this initial situation. We are aware of the fact that the term *intention movement* was born out of a very different school of thought and perspective from of Brentano's sense of *intention* as philosophical/psychological phenomenon. Ethologists like Nikolaas Tinbergen (1907 - 1980) used the term *intention movement* as a synonym for "displacement activities" (Tinbergen, 1952) or "preparatory movements" (Tinbergen, 1939). Using it in the ethological or biological sense did not initially entail the attribution of mental states to non-human animals.

The epistemological turn from biological intention to the psychological form was initiated by the philosopher Daniel Dennett (1983) and the psychologists David Premack and Guy Woodruff (1978). As is sometimes explained in hindsight, they introduced the psychological notion to US-American behaviourism and European ethology, two biologically-rooted paradigms that were inclined to see in it an inherently anthropomorphic enterprise (Seed & Tomasello, 2010). The first question of the current study is: (1) Do these epistemologically distinct perspectives (biological and psychological) continue to co-exist within the cross-species comparative discourse, or did they exchange views during the last decades? Further questions arise from this initial query:

(2) Which authors helped propagate the term *intention* within the cross-species comparative discourse, and in relation to which epistemological concept? (3) To what extent was the introduction of the term accompanied by empirical evidence? (4) What forces drive the spread of the term in the discourse?

By answering these questions, the aim is to deliver an outline on the development of the term *intention* and its associated concept in historical times. Reflecting on this case might facilitate an

improved understanding of the different epistemological foundations of the term as used by various schools of thought down to the present day.

Material & Methods

We gathered 653 articles from the database 'Web of Science' (https://webofknowledge.com; hereafter: WoS) and used the provided meta-data to create a citation map (i.e. 'who-cites-whom'). By visualising the number of times an article is citied within the corpus, we want to provide a valuable empirical measure of the level of popularity of an article. Not only do we plot these articles and their frequency of citation, but we also categorise every publication according to the author's scope, school of thought, and definition of intentionality applied. The combination of both qualitative and quantitative methods (i.e. mixed-method-approach) is a powerful tool, circumventing the limitations of a purely quantitative programme or purely qualitative approach. While the former is blind towards the contents of publications, the latter is blind to frequencies and large numbers. We see our approach extending, rather than competing with, other meta-research studies that evaluate, reassess, or improve scientific practice (Ioannidis et al., 2015). We use principles from network theory to map an evolving discourse from a new perspective and to gain a more holistic picture of the cross-species comparative research on intention.

Search queries were conducted in WoS in order to identify relevant articles within cross-species comparative sciences containing the term *intention*. To limit the results, the items 'intent*' and 'intend*' were combined with 'animal', 'non-human', 'species'; 'song', 'behaviour', 'signal', 'gesture', etc. (for a full description of search queries, see Supplemental Material). The use of asterisks refers to a wildcard. An initial query, searching the title, abstract and keywords, yielded 5,925 articles. However, most of them, like articles from the WoS categories 'Computer Science', 'Economics' or 'Surgery', originate from fields not associated with cross-species comparative science and were consequently excluded. In addition to that, we checked each of the remaining articles to determine if they met all of the following five criteria for relevance:

a) peer-reviewed journal article (no editorials)
b) written in English or German
c) comparative (research compares at least two different species)
d) no plants, fungi, cells, or computers (to keep results manageable)
e) 'intent*' appears at least twice in full text

After filtering and checking the initial 5,925 hits, 394 of them were considered relevant, based on criteria (a)-(e). In a second step, we sent queries to WoS using terms often used in cross-species comparative literature, such as 'deception' (3,262 hits / 27 relevant); 'referential' (359 hits / 23 relevant); 'joint attention', 'visual attention' or 'attentional state' (211 hits / 17 relevant); 'gaze follow*' (175 hits / 15 relevant); 'ostension' (1437 hits / 9 relevant). Taken together, the queries resulted in 11,369 hits, from which 485 (4.3%) were identified as relevant.

Since there are no searchable abstracts available on WoS for articles published before 1991, older articles were added manually. In a third step, 20 articles from the corpus were selected on the basis of their representing diversity in years, study species, schools of thought and authors. We reviewed all references cited in these 20 articles; if one of the cited references was not yet included in the corpus, we added it. Additionally, 20 randomly chosen articles were reviewed in the same manner. Checking these 40 articles yielded another 103 publications to add to the existing corpus, which now consisted of 587 items.

In a forth step, we analysed which publications are most cited within the total corpus of 587 articles. Sixty-six publications that received at least ten citations and used the term *intention* at least twice in the full text, and which were not already part of the corpus, were finally added. Twelve of them were not available in WoS and had to be added by hand. Apart from those 12 articles, the dataset is limited to publications available on WoS on 29th June 2017. Altogether, 653 articles were used for further analysis.

Construction of the citation map is based on meta-data provided by WoS. The available meta-information was downloaded in a CSV file on 22 August 2017. The column 'CR' contains all 'cited references' of an article. The raw-data file was read into R version 3.4.1 (R Development Core Team, 2016). While parsing the file, a unique identifier was created, both for each article listed in column 'CR' and for the 653 articles from the original corpus. The identifier consists of, in this order, author, year of publication, and abbreviated ISO4 journal title. Since journal abbreviations differed, all of them were standardised to the same abbreviations.

The network was drawn using the 'ggraph-package' (Pedersen, 2017) and its Fruchterman-Reingold layout algorithm. This algorithm is based on the idea that the network can be treated like a physical system in which aesthetic criteria – such as minimising line crossings or distributing nodes evenly – exert an influence on the creation of the graph in terms of minimising the energy in the network (see Fruchterman & Reingold, 1991). The full R script used to produce the figures and analysis, as well as the raw-data files, can be downloaded at the Open Science Framework (doi: 10.17605/OSF.IO/N6PFZ). Quantitative, automatic plotting of a citation map has its own practical limitations, for instance, when it comes to identifying individual nodes and their semantic roles. Since we are not only interested in *how* the term spreads in a discourse, but also *how it is used*, we decided to systematically rate each article, allowing us to ask more qualitative questions (mixed-methods approach). The rating procedure consisted of three steps:

I. Determine the level of intention (i.e. zero-, first-, second-, or higher-order).
II. Determine the author's position (i.e. 'opposing' or 'supportive').
III. Determine the school of thought applied in the article (i.e. biological notion, cognitive notion, mixed).

In order to determine the *order of intention*, we used the definitions proposed by Daniel Dennett (1983):

i. zero-order: "[…] an account that attributes no mentality, no intelligence, no communication, no intentionality at all […]" (p.346)
ii. first-order: "A *first-order* intentional system has beliefs and desires (etc.) but no beliefs and desires *about* beliefs and desires." (p. 345, italics in original)
iii. second-order: "A *second-order* intentional system [...] has beliefs and desires [...] about beliefs and desires [...] - both those of others and its own." (p. 345, italics in original)
iv. higher-order (including all levels from third-order upwards): "A *third-order* intentional system is one that is capable of such states as […] x *wants* y to *believe* that x *believes* he is all alone." (p. 345, italics in original)

In order to determine the *school of thought* employed by the authors, we assigned the article to one of the following categories (**Tab.1**):

Table 1 List of the schools of thought ascribed to each article during the rating procedure. The names listed in the column labelled 'school of thought' function as a handle in the analysis and do not depict actual scientific paradigms. In brief, brown and yellow represent more behaviouristically, green and blue more cognitively oriented research. The colours represent the schools in subsequent figures.

Colour	School of thought	Description
	Game theory	Articles embedded in evolutionary game-theoretic discussions.
	Intentional stance	Articles using *intention* as a means of describing zero-order-intentional behaviour (e.g. *as if* the animal would have beliefs)
	Intention	Articles using the term in Dennett's philosophical sense as a psychological phenomenon.
	Mixed	Articles combining any of the present schools of thought (mostly a combination of behaviouristic and cognitive research).
	None	Articles where no present school of thought can be ascribed, because of an ambiguous or missing definition (not even implicit).

The full rating protocol, as well as a comprehensive decision diagram (**Fig. 5**), can be found in the supplementary materials section.

To identify prominent articles, we counted absolute numbers of citations. Since some authors consider such a measurement too simple and propose more sophisticated strategies (e.g. Ding, 2010), we also implemented a weighted measurement for importance in a directed network graph. This was done by computing Kleinberg's authority centrality scores using the R-package {igraph} (Csardi & Nepusz, 2006). Scores for authority and hub status are supposed to weigh an article's influence. Similar to the PageRank mechanism (Ding, 2010), Kleinberg (1999) defined his scores in a recursive way: Articles that link *many* related authorities are defined as *hubs* (e.g. reviews), while a*uthorities* are articles that are linked by influential publications (i.e. hubs) (Kleinberg, 1999, p. 8). In other words the authority

score weighs citations received from an influential article more than those from an article considered less important.

After arriving at our initial results for the most cited publications and articles of authority, we evaluated whether experiments were done on non-human primates or non-primate animals. This rating was implemented as an exploratory analysis.

In order to test the reliability of the qualitative assessment, a randomly selected subsample of 5% of the corpus (n = 33) was reassessed by a scorer blind to the questions of the study. According to an unweighted test on Cohen's kappa (R's 'kappa2' from package {irr v0.84}; Gamer, Lemon, Fellows, & Singh, 2012), we received a very high inter-rater reliability for all the qualitatively assessed categories (school: $\kappa = 0.935$; opposing: $\kappa = 0.875$; neutral: $\kappa = 0.858$; supportive: $\kappa = 0.9$; supportive-evidence: $\kappa = 0.937$).

Results

One corpus, two clusters: The citation network.

For the analysis, we collected 653 articles from species-comparative research that use the term *intention* in their titles, keywords or abstracts, or, in other cases, at least twice in the main text (see Methods for specifications). One measure of our analysis is the number of citations received per publication within the corpus. The ensuing article citation map depicts all articles that received or referenced at least one citation from another article inside the corpus (n=603) (**Fig.1**). As the figure makes apparent, the structure of the network is comprised of two clusters. The clusters emerge when a community in a network maintains more dense connections with one group than with another (Radicchi, Fortunato, & Vespignani, 2012, p. 242). Few articles cite literature from both clusters, with four nodes standing out. Those nodes represent articles that combine different epistemological schools (Cheney & Seyfarth, 1984; Guilford & Dawkins, 1991; Hauser & Nelson, 1991; Marler, Dufty, & Pickert, 1986). Both clusters are dominated by two distinct *schools of thought*, which are colour-coded in the present study. Here, an element enters the analysis which adds an extra layer to the graph, beyond the purely descriptive distribution of nodes, which helps us answer question (1) from the introduction.

Distinct definitions continue to co-exist: Intentionality according to various schools of thought.

Colours represent the different *schools of thought* that were rated manually for each article (**Tab.1**). In short, while the schools rated as *game* (i.e. game theory) or *stance* (i.e. behaviouristic schools) do not discuss internal mental conditions to non-human animals, schools rated as *intention* (i.e. cognitive science) or *mixed* (i.e. a combination of any of the schools) do. The distribution of colours in **Fig. 1**, a product of force-directed graph drawing (see Methods: Fruchterman-Reingold-analysis), appears to follow a pattern. The more cognitive studies accumulate on one side; the more behaviouristic articles

fall on the other, and few articles combining any of the schools fall between. The pattern that emerges helps us answer question (1): Do different schools of thought continue to co-exist? As the citation map reveals, the term *intention* enters the discourse independently from two sides. First, the term is introduced by European ethologists in the beginning of the 20th century in a purely methodological fashion (i.e. without considering mental states). It is then later introduced in the philosophical/psychological sense by various figures, with the American psychologists Premack and Woodruff leading the way. Yet authors from either tradition rarely cite one other. The analysis of citation links reveals that articles mostly cite literature from the school to which they belong. For instance, articles categorised as belonging to the school of *game theory* tend to cite publications from their own specialisation, at a rate of 70%. Of the remaining references, 23% link to behaviouristically oriented publications. Thus, articles rated as *game theory* rarely cite articles categorised as belonging to a more cognitively oriented school of thought (4.3%). The results for the remaining schools are similar (78% of the articles from the school *stance* cite articles from the school *stance* & *game*; 94% of the articles from the school *intention* cite articles from the school *intention* & *mixed*; 65% of the articles from school *mixed* cite the school *intention*; and 75% of articles tagged as 'none' cite articles from the school *intention*). The visualisation and analysis of citation links shows that two epistemologically distinct uses of *intention* were introduced independently and continue to co-exist to the present.

While different schools co-exist, the use of 'intentionality' has increased in cognitive science:

While two independent origins exist, in only one school–cognitive science–has the term enjoyed a successful career (**Fig. 2**). Under this umbrella term, one can subsume all research focussing on mental content in non-human animals. Hence, the histogram in the panel *intention* implicitly visualises the advent of cognitive science. What it does not show, however, is the overall preponderance of behaviouristic over cognitive research. **Fig. 2** does not lead to statements about the total number of publications within a respective school, but merely visualises the fact that, while cognitive scientists started using the term *intention* more often within the early 1980s, there was no such development in the more behaviouristic schools. One partial answer to question (4), regarding the driving forces behind these developments, is that the immense spread of the term in the cross-species comparative literature can be traced back to researchers interested in the mental abilities of non-human animals. Nothing has been said so far about *how* the term is used in the current corpus. What appears so in **Fig. 1** and **Fig. 2** is actually not as uniform as it seems.

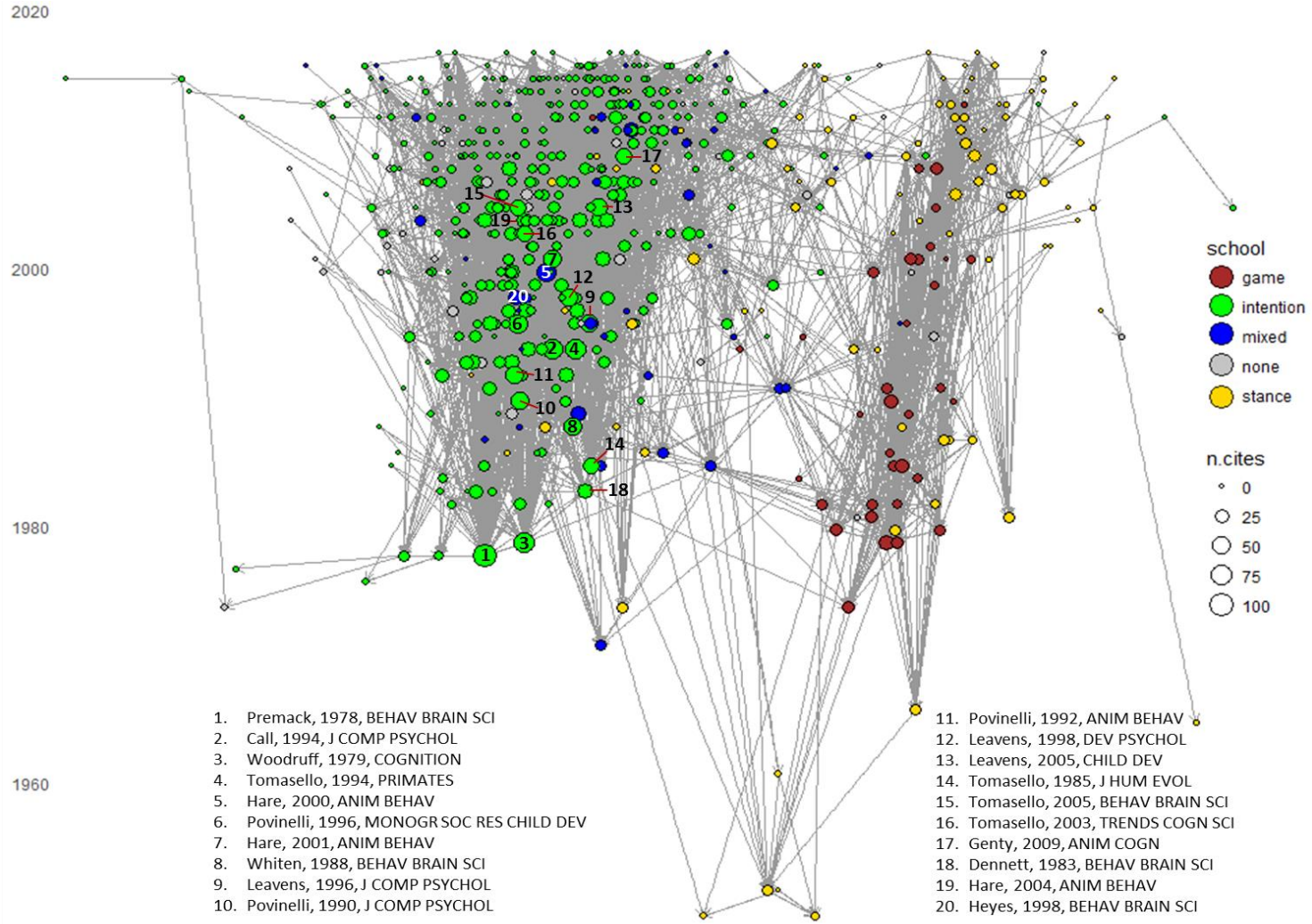

Figure 1 Citation map created from 603 articles (nodes) and 19827 citations (connections). The timeline for 1948-2017 is represented by the y-axis. The position of nodes follows the Fruchterman-Reingold algorithm, which, as a physical simulation, locates nodes according to their minimised energy expenditure. The size of a node represents the number of citations an article *receives*. Articles that do not receive a citation *and* do not cite other articles are absent. The colour of inner circles depicts the *school of thought* of an article. While the brown ('game') and yellow ('stance') nodes represent a more behaviouristic approach that defines 'intention' in a biological sense, without assuming mental content, the green ('intention') and blue ('mixed') nodes represent schools that discuss intention in a cognitive dimension. See the Methods section and supplementary material for more detailed descriptions of the schools. Numbers from 1-20 depict the 20 most cited articles within the corpus in absolute numbers. Articles are sorted in descending order. Title and DOI from the 20 articles can be found in **Tab. 2**.

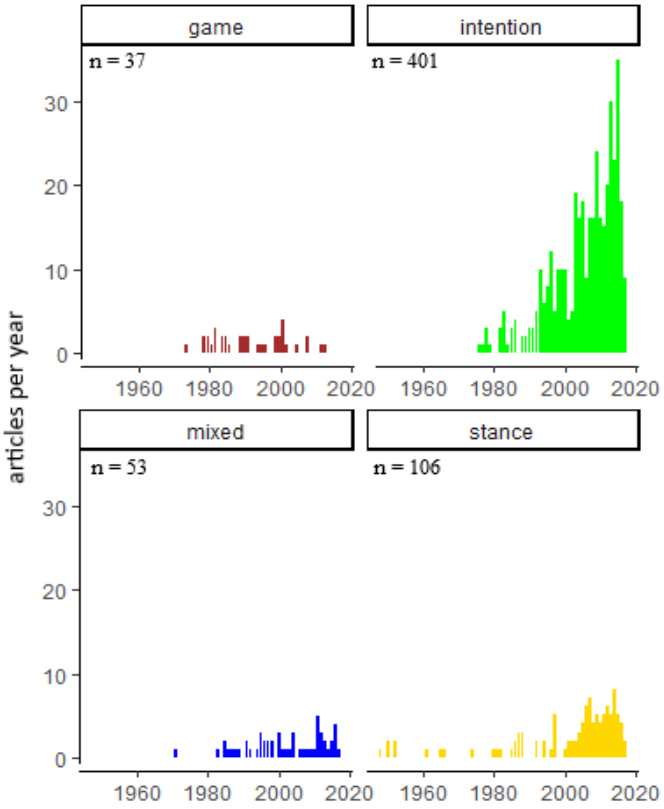

Figure 2 Appearance of the term *intention* according to schools of thought. The histogram depicts the number of publications per year and school. The x-axis represents the timeline for 1948-2017. 'Game' (game theory, brown) and 'Stance' (yellow) represent more behaviouristic schools. 'Intention' (green) and 'Mixed' (blue) represent schools related to cognitive science. Articles that could not be tagged with any of the labels (none) are absent from the present histograms (n=56). See the Methods section for more detailed explanations on the definitions of the respective schools.

Controversy exists not about intentionality, but about orders.

While all articles from the more cognitive schools of thought use the term *intentionality* at least in some cases in the philosophical/psychological sense, the statements therein can vary significantly. It starts with the observation that intentionality in itself is not a single-layered concept but consists of sub-orders as defined by philosopher Daniel Dennett (1983) and explained in the Methods section above. Different researchers ask very different research questions, depending on the orders of intentionality they choose to focus on. While some are interested in questions regarding non-human animals' mental processes per se (i.e. 1st -order intentionality, e.g. Griffin, 1978), others take their existence for granted and investigate the extent to which these mental processes are employed to deceive others (i.e. 2nd-order intentionality, e.g. Woodruff & Premack, 1979). Still others remain either neutral, oppose claims about 2nd-order intentionality (e.g. Heyes, 1993), or reject the very idea of mental content in non-human species (i.e. 1st-order intentionality, e.g. Silverman, 1983). We tagged the articles according to these possibilities and also took into consideration that one and the same article can represent multiple positions within the same text (see Methods). As depicted in **Fig. 3**, the debate about intentionality is more diverse than might appear from **Fig. 1**. The figure casts a light on question (3), regarding the extent to which empirical studies accompanied the term *intention* with evidence. It was only in the mid-1990s that sufficient experimental and observational evidence of 1st-order intentional behaviour was at hand (i.e. subjects possessing any mental content) in non-humans, and it took until 2004 that articles more frequently found evidence of 2nd-order intentionality (i.e. subjects *reflect* on their 1st-order mental content).

Interestingly, we rarely (n=20) found articles that opposed the concept of 1st-order intentionality from the outset. Until 1991, the corpus contains only one article providing positive evidence for 1st-order intentionality (Chevalier-Skolnikoff, Galdikas, & Skolnikoff, 1982), and one article for 2nd-order intentionality (Woodruff & Premack, 1979). At the same time, 26 publications (half of them on primates) are willing to accept the existence of 1st-order intentionality and support the idea of psychological states in non-humans in general. While authors support the idea of 1st-order intentionality in the absence of much evidence, they are much more likely to oppose 2nd-order intentionality. Until 2003, only two articles provide positive evidence for the existence of 2nd-order intentional behaviour (Call & Tomasello, 1994; Woodruff & Premack, 1979); at the same time, 34 articles dismiss it. From 2003 onwards, there are 22 articles tagged as providing evidence for 2nd-order intentionality, but there exist 56 articles on the other side that reject this view. As apparent from **Fig. 3,** most of the scientific controversy is about the order of intentionality and rarely about the concept itself.

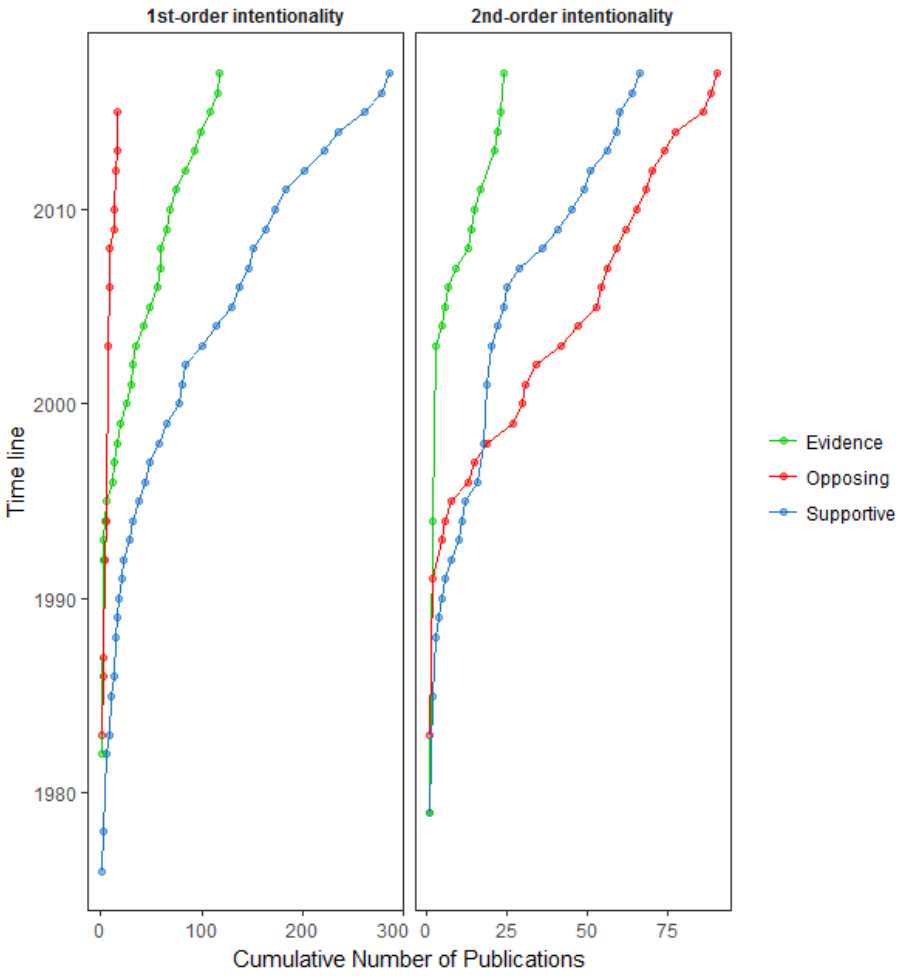

Figure 3 Line graph depicting the cumulative numbers of publications taking positions on orders of intentionality. The two graphs show the cumulative number of publications on 1st- and 2nd-order intentionality (see the Methods section for definitions). Green lines represent the number of publications that provide supportive evidence for the respective order. Blue lines show the number of publications that endorse intentionality in general, but do not provide direct data. Red lines show the number of publications that explicitly or implicitly reject the respective order of intentionality. The timeline for 1948-2017 is represented by the y-axis. Only publications from the school 'intention' (n=401) and 'mixed' (n=53) enter the histogram. Since 35% of all publications represent multiple positions on intentionality (e.g. supporting 1st-order, but rejecting 2nd-order intentionality), the histogram shows more than 454 entries (evidence for 1st-order: n=119; evidence for 2nd-order: n=24; supporting 1st-order: n=294; supporting 2nd-order: n=66; opposing 1st-order: n=20; opposing 2nd-order: n=90)

Articles' popularity: Identifying the most prominent articles.

One of the aims of citation networks is to measure the importance of a node, which in the current case represents an individual article. Due to the ambiguous nature of the term *importance*, node influence metrics often quantify the number of citations in order to identify *popular* nodes (Kas et al., 2012, p. 178). Here, popular nodes are articles that receive many citations. While this is the simplest measure of popularity, it is a powerful tool for revealing the seminal papers in a corpus. The bigger a node in the citation map, the more popular the article. For instance, the 20 most cited articles receive roughly one quarter (23.2%) of all the citations made within the corpus. While those 20 articles are cited at

least minimum 35 times each, another 208 publications do not receive a single mention from one of the 653 corpus articles. Fifty articles neither receive a citation, nor do they mention an article from the corpus. Those 50 articles are excluded from the citation map. In general, most articles are cited between one and four times (n=192) or 5-24 times (n=200). Only 53 articles are cited more than 24 times. We refer to the top 20 as the most prominent articles of the corpus. In order to answer question (2) from the introduction, we labelled them with numbers in descending order in **Fig.1**. A more weighted measure of importance as proposed by Kleinberg (1999) is the computation of authority and hub scores (see Methods). From the 20 most cited articles, 13 are among the top 20 articles with the highest authority-scores (**Tab. 2** in Supplementary Results). However, we found that, for the purposes of answering our proposed research questions, the calculation of authority and hub scores did not contribute much additional information beyond the simple measurement of popularity. Authority and hub scores, citation counts, and titles from the top 27 articles can be found in **Tab. 2** in the supplementary results section. For an extended annotation of the 90 most prominent nodes, see also **Fig.6**.

Interestingly, all of the 20 most-cited articles feature primates, and more than half (n=12) of them focus on chimpanzees. This prompted an exploratory post-hoc analysis in which we categorised each article according to the study species (i.e. whether it focused on non-human primates or non-primate animals). We found that the vast majority of articles from the cognitive science spectrum focus on primates (**Fig. 4**). This concerns nearly three out of four articles ($n_{intention}$ = 401; $n_{primates}$ =297; $n_{non\text{-}primates}$ = 104). On the other hand, as **Fig. 4** shows, primates are not a frequent research subject for behaviouristic schools. Apparently, *intention* in the philosophical/psychological sense is frequently associated with studies on primates.

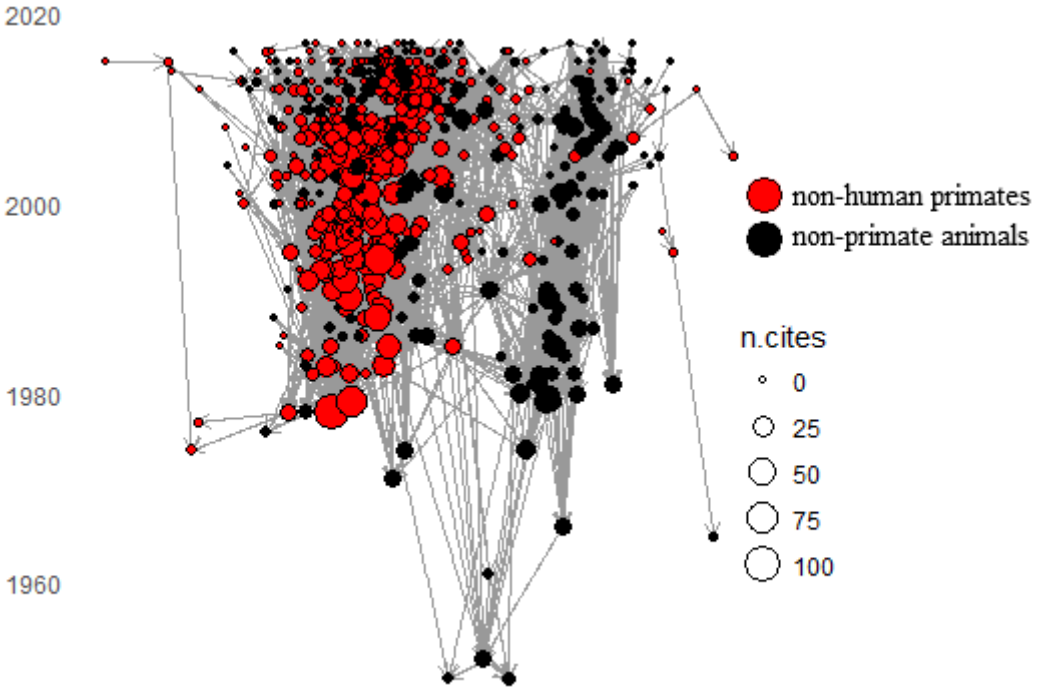

Figure 4 Citation map tagging articles with research focussing on non-human primates (red nodes, n=361) and non-human animals (black nodes, n=242). Red nodes mostly represent research articles with a focus on cognitive science (see **Fig. 1**). The network is created from 603 articles (nodes) and 19827 citations (connections). The timeline for 1948-2017 is represented by the y-axis.

What cannot be seen

Only articles that met criteria (a)-(e) for relevance (see Methods) are reproduced in the citation map. By definition this excluded a number of publications such as books, book chapters, and journal articles from other disciplines or without the key term *intention* in their full text. The ten most cited articles not belonging to the corpus, but frequently cited within it, are listed in **Tab. 3** in the supplementary material section. At position seven, there is a prominent study from three ethologists (Seyfarth, Cheney, & Marler, 1980). Their pioneering study on the predator-specific alarm calls of free-ranging vervet monkeys is frequently cited in the corpus. For instance, Daniel Dennett praised the "new ethologists" as prime examples of researchers' "having cast off the straightjacket of behaviorism" (Dennett, 1983, p. 343). Despite its prominence, the article is not part of the corpus we analyse because it does not employ the word intention. The case of the vervet-monkey study illustrates the extent to which advances in ethology could also act as a driving force in promoting the term *intention* in cognitive research. There might be further influences from ethology, which are not visible in the current map.

Manual search, for instance, could identify 33 articles published before 1978 which are missing in the WoS database. Many of those 33 articles are publications from 'classical' ethologists like Konrad Lorenz, Nikolaas Tinbergen (only represented once in the current corpus[156]), and Robert Hinde, but

[156] see Figure 6; [03] Tinbergen, 1952, Q REV BIOL

also psychologists like Ernst von Glasersfeld. Presumably, these 33 articles are just a fraction of older articles not available in WoS. We can only speculate about the impact of such hidden articles on the current citation map, but would expect that cognitive science until the late 1980s was influenced by ethology to a much greater degree than can be observed in the current citation map.

Discussion

The primary message of the current study should read: no idea is developed in isolation. This is not merely a trivial, common-sense phrase, but it is actually more complex than it might first appear.

Plotting the connections between articles can be understood as visualising the ongoing communication within a scientific community, in this case over a period of 69 years. By using a mixed-methods approach we were able to reveal two independent origins of the term *intention* in cross-species comparative research. These distinct schools of thought coevolved through time to the present day.

The answer to our initial question regarding the forces that contributed to its spread is that, were it not for the *cognitive shift*, *intention* would not have become a prominent term. Its use only increased in more cognitively oriented schools; while behaviouristic schools introduced the term decades earlier and still use it today, there was no comparable increase in use. This might raise the question as to whether the history and use of the term *intention* in different schools is comparable at all. Indeed, although the schools use the expression differently and are structurally separated in terms of citations, the division is not as clear as it might appear from the data. There are various influences flowing from more behaviouristic schools to cognitive science and vice versa. For instance some authors from a more cognitively oriented position use the biological description of *intention movements* in order to explain the origin of intentional gestures (e.g. Tomasello et al., 1997, p. 225; Tomasello, Gust, & Frost, 1989, p. 43). Other more biologically inspired authors have used philosophical/psychological ideas about intention to establish a discipline called *cognitive ethology* (e.g. Beer, 1997; Bekoff & Allen, 1992; Jamieson & Bekoff, 1992). These examples demonstrate that cases do exist in which both schools are mixed and combined. In such cases it seems essential that authors are explicit with their definitions and state clearly the school of thought to which they belong, lest they cause "confusion" by using "quite different concepts" (Hauser & Nelson, 1991, p. 186), an issue also observed by Wharton (2003, p. 448). We hope that the current study encourages a future trend towards increased clarity that promotes more interdisciplinary exchange, but less confusion about the chosen vocabulary.

This aside, there are a number of general challenges that come with citation analysis. For instance, we will have to accept that authors and research groups frequently cite their own work (Aksnes, 2003). We cannot distinguish whether an article is cited for poor methodological quality or because of its ground-breaking results. Another challenge, as pointed out by Simkin and Roychowdhury (2003), is

that authors do not necessarily read their cited studies, but rather copy them from other articles. That is to say, there are many reasons to cite an article, often unrelated to the questions posed here (see Erikson & Erlandson, 2014). The quantitative approach cannot distinguish between them. In order to circumvent these limitations, weighted citations (authority score and hub score) were calculated. However, as listed in **Tab. 2** (Supplementary Results), no large gap exists between the simple measurement of prominence (a count of the total number of citations) and weighted scores. Citation biases might lose weight when citations are analysed for longer periods.

Our second question, as worded in the introduction, refers to the act of identifying the most prominent articles; this is one of the most intriguing aspects of the current study. The mere structure of the citation network allows one to identify articles that, while perhaps not very popular in terms of citations, still connect divergent fields and thus help establish new directions for the debate. At a glance, **Fig. 1** suggests that only a few people dominate the corpus: David Premack, Michael Tomasello, Josep Call (a former graduate student of Tomasello's), Brian Hare (who collaborated with Tomasello and Call for many years), Daniel Povinelli, David Leavens, Andrew Whiten, Daniel Dennett, and Cecilia Heyes (whose postdoctoral thesis was supervised by Dennett). The 20 most cited articles (3%) receive nearly one quarter of all the citations made within the corpus. Our observation that a small group of articles and authors receive most of the attention confirms the hypothesis proposed by other meta-researchers (Ioannidis, Boyack, & Klavans, 2014).

Not to be overlooked is our third question, which seeks evidence regarding the introduction of the term *intention* into the discourse at hand. The categorisation and subsequent analysis of the articles reveals little evidence for the establishment of a cognitive vocabulary before the mid-1990s (**Fig. 3**). Although an increasing number of articles began using cognitive vocabulary in the early 1980s, they do not seem to have received much attention as a result. Even self-described behaviourists like Cecilia Heyes do not in principle object to the notion of "not directly-observable states and processes" (Heyes, 1993, p. 187). In fact, most of the debate concerns the order of intentionality one should ascribe to non-human animals, not about the very idea of mental states. The empirical data provides further evidence to support those who criticise the use of the phrase 'cognitive revolution' (Hobbs & Burman, 2009; Leahey, 1992; Mandler, 2002). If a revolution entails a clash with an established school of thought, we found little evidence for a defence being mounted against the ascendant cognitivists (**Fig. 3)**.

An unexpected answer to our fourth question might be contained in the following observation: Although the corpus is limited to a small fraction of a larger discourse, the striking dominance of primate research within cognitive science was unexpected. Not only did we find quantitative proof for "Chimpocentrism" (Vaesen, 2014), but for Primatocentrism in general. It seems as if researchers working with primates were among the major propagators of the term in its psychological sense. Although most cognitive scientists would likely agree with the proposition that intentional phenomena

– such as goal-directedness, the possession of concepts and problem solving behaviour –also exist in non-primate animals as well, there is a poverty of research done on those species. Only one-quarter of the articles from cognitive science investigates species other than primates. It seems that researchers engaged with primates use cognitive vocabulary more readily than researchers focussing on non-primate species. For instance Xitco and colleagues (Xitco, Gory, & Kuczaj, 2004) grounded their study with dolphins in an experiment developed by Tomasello and Call (1994). Although their "dolphins' performance was similar in many respects to that of the orangutans tested by Call and Tomasello" (Xitco et al., 2004, p. 236), they completely avoided using cognitive vocabulary like 'intention', 'comprehension', 'wishing', 'understanding' and the like, as Call and Tomasello did.[157] Consequently, the dolphin article is missing from the corpus.

There are more examples of the above: When some primatologists observe "gaze alternation", they are using this as a measure for "intentional communication" (Leavens & Hopkins, 1998), while the same phenomenon in spiders is described in purely behaviouristic vocabulary as "directional biases in signal transmission" (Echeverri, Morehouse, & Zurek, 2017). When other primatologists describe behaviour with the term "frustration" (Hare & Tomasello, 2004, p. 578), ethologists working with birds describe a similar behaviour in terms of "aggressive motivation" (Hurd & Enquist, 2001, p. 931). Of course, these differences in vocabulary are grounded differences regarding the layer of analysis (ultimate vs proximate). However, it remains an open question as to why primatologists tend to ask ultimate questions while non-primatologists read the same phenomenon on a proximate level. One possible answer could lie in the historical impact of the old idea of the *Scala Naturae* (Ullrich, Mittelbach, & Liebal, 2018). On the other hand, it must be noted that there does exist a number of cognitive studies on non-primate animals. In the current analysis, however, they are overshadowed by other, more prominent and more authoritative articles, as most important non-primate studies are of a more recent date. These include advances in bird communication research (Bolhuis, Okanoya, & Scharff, 2010; Fitch, 2017) or bird cognition in general (Emery, 2006; Güntürkün & Bugnyar, 2016), as well as cognitive perspectives on fish (Bshary, Wickler, & Fricke, 2002; Vail, Manica, & Bshary, 2013; Brown, 2015) and reptile cognition (Wilkinson & Huber, 2012). Future research will have to devote attention to these new developments, which are beyond the scope of the current article.

This notwithstanding, it is nearly impossible to unravel all of the connections between scientists' published output and the historical roots of ideas. This applies even to the attempt to visualise the introduction of a new term within a limited corpus of 653 articles. Our analysis supports the conclusion that while a citation network cannot replace qualitative historical research, it can help to identify valuable research targets and provide a global historical overview of a diverse and long-evolving discourse.

[157] Although one of the study's keywords is "cognition", its authors avoid referring to unobservable entities.

Conflict of Interest: The authors declare that they have no conflict of interest.

Data availability: The datasets analysed during the current study are available in the OSF repository, DOI 10.17605/OSF.IO/N6PFZ.

Funding: R. Ullrich received funding from an Elsa-Neumann scholarship. K. Liebal was supported by the Excellence Initiative of the German Research Foundation.

Literature

Aksnes, D. W. (2003). A macro study on self-citation. *Scientometrics*, *56*(2), 235–246. http://doi.org/10.1023/A:1021919228368

Andrews, K. (2014). Animal cognition. In *The Stanford Encyclopedia of Philosophy* (Fall 2014). Edward N. Zalta (ed.). http://doi.org/10.1016/j.cub.2004.05.008

Beer, C. (1997). Expressions of Mind in Animal Behavior. In R. W. Mitchell, N. S. Thompson, & H. L. Miles (Eds.), *Anthropomorphism, Anecdotes, and Animals* (pp. 198–212). Albany: SUNY Press.

Bekoff, M., & Allen, C. (1992). Intenional Icons: Towards an Evolutionary Cognitive Ethology. *Ethology*, *91*(1), 89–107. http://doi.org/10.1111/j.1439-0310.1992.tb00846.x

Bolhuis, J. J., Okanoya, K., & Scharff, C. (2010). Twitter evolution: converging mechanisms in birdsong and human speech. *Nature Reviews. Neuroscience*, *11*(11), 747–59. http://doi.org/10.1038/nrn2931

Brentano, F. (2009). *Psychology from an Empirical Standpoint [1874]*. (O. Kraus & L. L. McAlister, Eds.). London, New York: Routledge.

Brown, C. (2015). Fish intelligence, sentience and ethics. Animal Cognition, 18(1), 1–17. http://doi.org/10.1007/s10071-014-0761-0

Bshary, R., Wickler, W., & Fricke, H. (2002). Fish cognition: a primate's eye view. *Animal Cognition*, *5*(1), 1–13. http://doi.org/10.1007/s10071-001-0116-5

Call, J., & Tomasello, M. (1994). Production and comprehension of referential pointing by Orangutans (Pongo pygmaeus). *Journal of Comparative Psychology*, *108*(4), 307–317. http://doi.org/10.1037/0735-7036.108.4.307

Cheney, D. L., & Seyfarth, R. M. (1985). Vervet Monkey Alarm Calls: Manipulation Through Shared Information? *Behaviour*, *94*(1), 150–166. http://doi.org/10.1163/156853985X00316

Chevalier-Skolnikoff, S., Galdikas, B. M. F., & Skolnikoff, A. Z. (1982). The Adaptive Significance of Higher Intelligence in Wild Orang-utans: a Preliminary Report. *Journal of Human Evolution*, *11*(7), 639–652. http://doi.org/10.1016/S0047-2484(82)80010-9

Cromwell, H. C., & Panksepp, J. (2011). Rethinking the cognitive revolution from a neural perspective: How overuse/misuse of the term "cognition" and the neglect of affective controls in behavioral neuroscience could be delaying progress in understanding the BrainMind. *Neuroscience and Biobehavioral Reviews*, *35*(9), 2026–2035. http://doi.org/10.1016/j.neubiorev.2011.02.008

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal*, *Complex Systems*, (5), 1695. Retrieved from http://igraph.org

Dennett, D. C. (1983). Intentional systems in cognitive ethology: The "Panglossian paradigm" defended. *Behavioral and Brain Sciences*, *6*(3), 343. http://doi.org/10.1017/S0140525X00016393

Ding, Y. (2010). Applying Weighted PageRank to Author Citation Networks. *Journal of the American Society for Information Science and Technology*, *62*(3), 236–245. http://doi.org/10.1002/asi.21452

Echeverri, S. A., Morehouse, N. I., & Zurek, D. B. (2017). Control of signaling alignment during the dynamic courtship display of a jumping spider. *Behavioral Ecology*, *0*, 1–9. http://doi.org/10.1093/beheco/arx107

Emery, N. J. (2006). Cognitive ornithology: the evolution of avian intelligence. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *361*(1465), 23–43.

http://doi.org/10.1098/rstb.2005.1736
Erikson, M. G., & Erlandson, P. (2014). A taxonomy of motives to cite. *Social Studies of Science*, *44*(4), 625–637. http://doi.org/10.1177/0306312714522871
Fitch, T. W. (2017). Empirical approaches to the study of language evolution. *Psychonomic Bulletin & Review*. http://doi.org/10.3758/s13423-017-1236-5
Friman, P. C., Allen, K. D., Kerwin, M. L. E., & Larzelere, R. (1993). Changes in Modern Psychology. *American Psychologist*, *48*(6), 658–664. http://doi.org/10.1037//0003-066X.48.6.658
Fruchterman, T. M. J., & Reingold, E. M. (1991). Graph Drawing by Force-Directed Placement. *Software: Practice and Experience*, *21*(11), 1129–1164. http://doi.org/10.1002/spe.4380211102
Gamer, M., Lemon, J., Fellows, I., & Singh, P. (2012). irr: Various Coefficients of Interrater Reliability and Agreement. Retrieved from https://cran.r-project.org/package=irr
Greenwood, J. D. (1999). Understanding the "cognitive revolution" in psychology. *Journal of the History of the Behavioral Sciences*, *35*(1), 1–22. http://doi.org/10.1002/(SICI)1520-6696(199924)35:1<1::AID-JHBS1>3.0.CO;2-4
Griffin, D. R. (1978). Prospects for a cognitive ethology. *Behavioral and Brain Sciences*, *1*(4), 527–538. http://doi.org/10.1017/S0140525X00076524
Guilford, T., & Dawkins, M. S. (1991). Receiver psychology and the evolution of animal signals. *Animal Behaviour*, *42*(1), 1–14. http://doi.org/10.1016/S0003-3472(05)80600-1
Güntürkün, O., & Bugnyar, T. (2016). Cognition without Cortex. Trends in Cognitive Sciences, 20(4), 291–303. http://doi.org/10.1016/j.tics.2016.02.001
Hare, B., & Tomasello, M. (2004). Chimpanzees are more skilful in competitive than in cooperative cognitive tasks. *Animal Behaviour*, *68*(3), 571–581. http://doi.org/10.1016/j.anbehav.2003.11.011
Hauser, M. D., & Nelson, D. A. (1991). "Intentional" Signaling in Animal Communication. *Trends in Ecology & Evolution*, *6*(6), 186–189. http://doi.org/10.1016/0169-5347(91)90211-F
Heinroth, O. (1990). *Beiträge zur Biologie, namentlich Ethologie und Psychologie der Anatiden [1910]*. Wien: Verein für Ökologie und Umweltforschung.
Heyes, C. M. (1993). Anecdotes, training, trapping and triangulating: do animals attribute mental states? *Animal Behaviour*, *461*(1), 177–188. http://doi.org/10.1006/anbe.1993.1173
Hobbs, S., & Burman, J. T. (2009). Is the "cognitive revolution" a myth? *Psychologist*, *22*(9), 812–814.
Hurd, P. L., & Enquist, M. (2001). Threat display in birds. *Canadian Journal of Zoology*, *79*(6), 931–942. http://doi.org/10.1139/cjz-79-6-931
Ioannidis, J. P. A., Boyack, K. W., & Klavans, R. (2014). Estimates of the continuously publishing core in the scientific workforce. *PLoS ONE*, *9*(7). http://doi.org/10.1371/journal.pone.0101698
Ioannidis, J. P. A., Fanelli, D., Dunne, D. D., & Goodman, S. N. (2015). Meta-research: Evaluation and Improvement of Research Methods and Practices. *PLOS Biology*, *13*(10), e1002264. http://doi.org/10.1371/journal.pbio.1002264
Jacob, P. (2014). Intentionality. In *The Stanford Encyclopedia of Philosophy* (Winter2014). Edwad N. Zalta (ed.). Retrieved from https://plato.stanford.edu/entries/intentionality/
Jamieson, D., & Bekoff, M. (1992). On Aims and Methods of Cognitive Ethology. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, *2*, 110–124. http://doi.org/10.1086/psaprocbienmeetp.1992.2.192828
Kas, M., Carley, K. M., & Carley, L. R. (2012). Trends in science networks: understanding structures and statistics of scientific networks. *Social Network Analysis and Mining*, *2*(2), 169–187. http://doi.org/10.1007/s13278-011-0044-6
Kleinberg, J. M. (1999). Authoritative sources in a hyperlinked environment. *Journal of the ACM*, *46*(5), 604–632. http://doi.org/10.1145/324133.324140
Leahey, T. H. (1992). The mythical relolutions of American psychology. *American Psychologist*, *47*(2), 308–318. http://doi.org/10.1037//0003-066X.47.2.308
Leavens, D. A., & Hopkins, W. D. (1998). Intentional communication by Chimpanzees: A Cross-Sectional Study of the Use of Referential Gestures. *Developmental Psychology*, *34*(5), 813–822. http://doi.org/10.1037/0012-1649.34.5.813
Lorenz, K. (1937). Über die Bildung des Instinktbegriffes. *Die Naturwissenschaften*, *19*, 289–300.

http://doi.org/10.1007/BF01491749
Mandler, G. (2002). Origins of the cognitive (r) evolution. *Journal of the History of the Behavioral Sciences*, *38*(4), 339–53. http://doi.org/10.1002/jhbs.10066
Marler, P., Dufty, A., & Pickert, R. (1986). Vocal communication in the domestic chicken: II. Is a sender sensitive to the presence and nature of a receiver? *Animal Behaviour*, *34*(1), 194–198. http://doi.org/10.1016/0003-3472(86)90023-0
Miller, G. A. (2003). The cognitive revolution: A historical perspective. *Trends in Cognitive Sciences*, *7*(3), 141–144. http://doi.org/10.1016/S1364-6613(03)00029-9
Morgan, L. C. (1903). *An introduction to comparative psychology* (2nd ed.). London: Walter Scott.
Pedersen, T. L. (2017). ggraph: An Implementation of Grammar of Graphics for Graphs and Networks. Retrieved from https://cran.r-project.org/package=ggraph
Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *The Behavioral and Brain Sciences*, *4*, 515–526.
Proctor, R. W., & Vu, K. P. L. (2006). The cognitive revolution at age 50: Has the promise of the human information-processing approach been fulfilled? *International Journal of Human-Computer Interaction*, *21*(3), 253–284. http://doi.org/10.1207/s15327590ijhc2103_1
R Development Core Team, R. (2016). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.r-project.org/
Radicchi, F., Fortunato, S., & Vespignani, A. (2012). Citation Networks. In A. Scharnhorst, K. Börner, & P. van den Besselaar (Eds.), *Models of Science Dynamics*. *Understanding Complex Systems.* (pp. 233–257). Berlin, Heidelberg: Springer. http://doi.org/10.1007/978-3-642-23068-4_7
Radick, G. (2016). The unmaking of a modern synthesis: Noam Chomsky, Charles Hockett, and the Politics of Behaviorism, 1955–1965. *Isis*, *107*(1), 49–73. http://doi.org/10.1086/686177
Robins, R. W., Goslin, S. D., & Craik, K. H. (1999). An Empirical Analysis of Trends in Psychology. *American Psychologist*, *54*(2), 117–128. http://doi.org/10.1037//0003-066X.54.2.117
Rosati, A. G., Wobber, V. V, Hughes, K., & Santos, L. R. (2014). Comparative developmental psychology: how is human cognitive development unique? *Evolutionary Psychology*, *12*(2), 448–473. http://doi.org/10.1177/147470491401200211
Seed, A., & Tomasello, M. (2010). Primate cognition. *Topics in Cognitive Science*, *2*(3), 407–419. http://doi.org/10.1111/j.1756-8765.2010.01099.x
Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980). Vervet monkey alarm calls: Semantic communication in a free-ranging primate. *Animal Behaviour*, *28*(4), 1070–1094. http://doi.org/10.1016/S0003-3472(80)80097-2
Silverman, P. S. (1983). Attributing mind to animals: The role of intuition. *Journal of Social and Biological Structures*, *6*(3), 231–247. http://doi.org/10.1016/S0140-1750(83)80045-1
Simkin, M. V, & Roychowdhury, V. P. (2003). Read Before You Cite! *Complex Systems*, *14*(3), 269–274.
Sperry, R. W. (1993). The impact and promise of the cognitive revolution. *American Psychologist*, *48*(8), 878–885. http://doi.org/10.1037/0003-066X.48.8.878
Tinbergen, N. (1939). On the Analysis of Social Organization Among Vertebrates, with Special Reference to Birds. *The American Midland Naturalist*, *21*(1), 210–234.
Tinbergen, N. (1952). "Derived" Activities; Their Causation, Biological Significance, Origin, and Emancipation During Evolution. *The Quarterly Review of Biology*, *27*(1), 1–32. http://doi.org/10.1086/398642
Tomasello, M., Call, J., Warren, J., Frost, G. T., Carpenter, M., & Nagell, K. (1997). The Ontogeny of Chimpazee Gestural Signals: A Comparison Across Groups and Generation. *Evolution of Communication*, *1*(2), 223–259. http://doi.org/10.1075/eoc.l.2.04tom
Tomasello, M., Gust, D., & Frost, G. T. (1989). A longitudinal investigation of gestural communication in young chimpanzees. *Primates*, *30*(1), 35–50. http://doi.org/10.1007/BF02381209
Tracy, J. L., Robins, R. W., & Gosling, S. D. (2004). Tracking trends in psychological science: An empirical analysis of the history of psychology. In T. C. Dalton & R. B. Evans (Eds.), *The Life Cycle of Psychological Ideas: Understanding Prominence and the Dynamics of Intellectual*

Change (pp. 105–130). New York: Kluwer Academic/Plenum. http://doi.org/10.1007/0-306-48010-7_5

Ullrich, R., & Liebal, K. (2018). Times change, values change: Criteria for attributing language in species comparative research [in print: expected summer 2018]. In E. Luef & M. M. Marin (Eds.), *The talking species: Perspectives on the evolutionary, neuronal, and cultural foundations of language*. Graz: Unipress Graz-Verlag.

Ullrich, R., Mittelbach, M., & Liebal, K. (2018). Scala Naturae: The Impact of Historical Values on Current "Evolution of Language" Discourse. *Journal of Language Evolution*, *3*(1), 1–12. http://doi.org/10.1093/jole/lzx017

Vaesen, K. (2014). Chimpocentrism and reconstructions of human evolution (a timely reminder). *Studies in History and Philosophy of Science Part C :Studies in History and Philosophy of Biological and Biomedical Sciences*, *45*(1), 12–21. http://doi.org/10.1016/j.shpsc.2013.12.004

Vail, A. L., Manica, A., & Bshary, R. (2013). Referential gestures in fish collaborative hunting. *Nature Communications*, *4*, 1765. http://doi.org/10.1038/ncomms2781

Virues-Ortega, J., & Pear, J. J. (2014). A History of "Behavior" and "Mind": Use of Behavioral and Cognitive Terms in the 20th Century. *Psychological Record*, *65*(1), 23–30. http://doi.org/10.1007/s40732-014-0079-y

Watrin, J. P., & Darwich, R. (2012). On behaviorism in the cognitive revolution: Myth and reactions. *Review of General Psychology*, *16*(3), 269–282. http://doi.org/10.1037/a0026766

Wharton, T. (2003). Natural Pragmatics and Natural Codes. *Mind and Language*, *18*(5), 447–477. http://doi.org/10.1111/1468-0017.00237

Wilkinson, A., & Huber, L. (2012). Cold-Blooded Cognition: Reptilian Cognitive Abilities. In T. K. Shackelford & J. Vonk (Eds.), *The Oxford Handbook of Comparative Evolutionary Psychology* (pp. 129–143). New Jersey: Oxford University Press. http://doi.org/10.1093/oxfordhb/9780199738182.013.0008

Woodruff, G., & Premack, D. (1979). Intentional communication in the chimpanzee: The development of deception. *Cognition*, *7*(4), 333–362. http://doi.org/10.1016/0010-0277(79)90021-0

Xitco, M. J., Gory, J. D., & Kuczaj, S. A. (2004). Dolphin pointing is linked to the attentional behavior of a receiver. *Animal Cognition*, *7*(4), 231–238. http://doi.org/10.1007/s10071-004-0217-z

Supplementary Methods

R-packages used for analysis

The following R-packages were in use for data analysis in R 3.4.1 (R Development Core Team, 2016):

"ggplot2" (Wickham, 2009)
"plotly" (Sievert, Parmer, Hocking, Chamberlain, & Karthik, 2017)
"RColorBrewer" (Neuwirth, 2014)
"ggraph" (Pedersen, 2017)
"dplyr" (Wickham & Francois, 2016)
"tidyr" (Wickham, 2016)
"irr" (Gamer, Lemon, Fellows, & Singh, 2012)
"igraph" (Csardi & Nepusz, 2006)

'Web of Science' search queries

The full search queries send to the 'Web of Science' database on April 26th 2017 read as follows:

TS=(intent* AND (behavio* OR language OR speech OR communication* OR song OR signal* OR vocalisation* OR vocalization* OR gestur* OR information* OR call) AND (animal* OR nonhuman* OR non-human* OR specie*) NOT (plant* OR robot* OR clinical* OR disease* OR patient* OR syndrome* OR schizophren* OR psychiatry* OR retard* OR milk OR immuno* OR enzym* OR disorder* OR microbiol* OR anatomy* OR alcohol* OR computer* OR membrane*))
OR
TI=(intent* AND (behavio* OR language OR speech OR communication* OR song OR signal* OR vocalisation* OR vocalization* OR gestur* OR information* OR call) NOT (plant* OR robot* OR clinical* OR disease* OR patient* OR syndrome* OR schizophren* OR psychiatry* OR retard* OR immuno* OR enzym* OR disorder* OR microbiol* OR anatomy* OR alcohol* OR milk OR computer* OR membrane*))

The asterisk symbol functions as wildcard. The remaining settings were left on default. Subsequent queries were written by replacing 'intent*' by one of the following terms:

- intend*
- ostensi*
- deception
- joint-attention*
- visual attention
- attentional state
- gaze follow*
- referential*

Qualitative assessment of the articles

After collecting a corpus of 653 articles we did rate each of the articles in a three-step-procedure.

I. Determine the orders of intentionality (values = 0, 1, 2, 3)

II. Determine the author's attitude. Place one OR multiple ratings from I. in respective columns.

III. Determine the school of thought of an article and write in column 'school'.

Below we provide our 1-page decision diagram and the complete rating-protocol as it was used by the independent rater to replicate our qualitative assessment. After a training-trial consisting of 33 articles, the subsequent reliability rating (n=33) was very high (R's 'kappa2' from package {irr v0.84} (Gamer, Lemon, Fellows, & Singh, 2012): school: $\kappa = 0.935$; opposing: $\kappa = 0.875$; neutral: $\kappa = 0.858$; supportive: $\kappa = 0.9$; supportive-evidence: $\kappa = 0.937$).

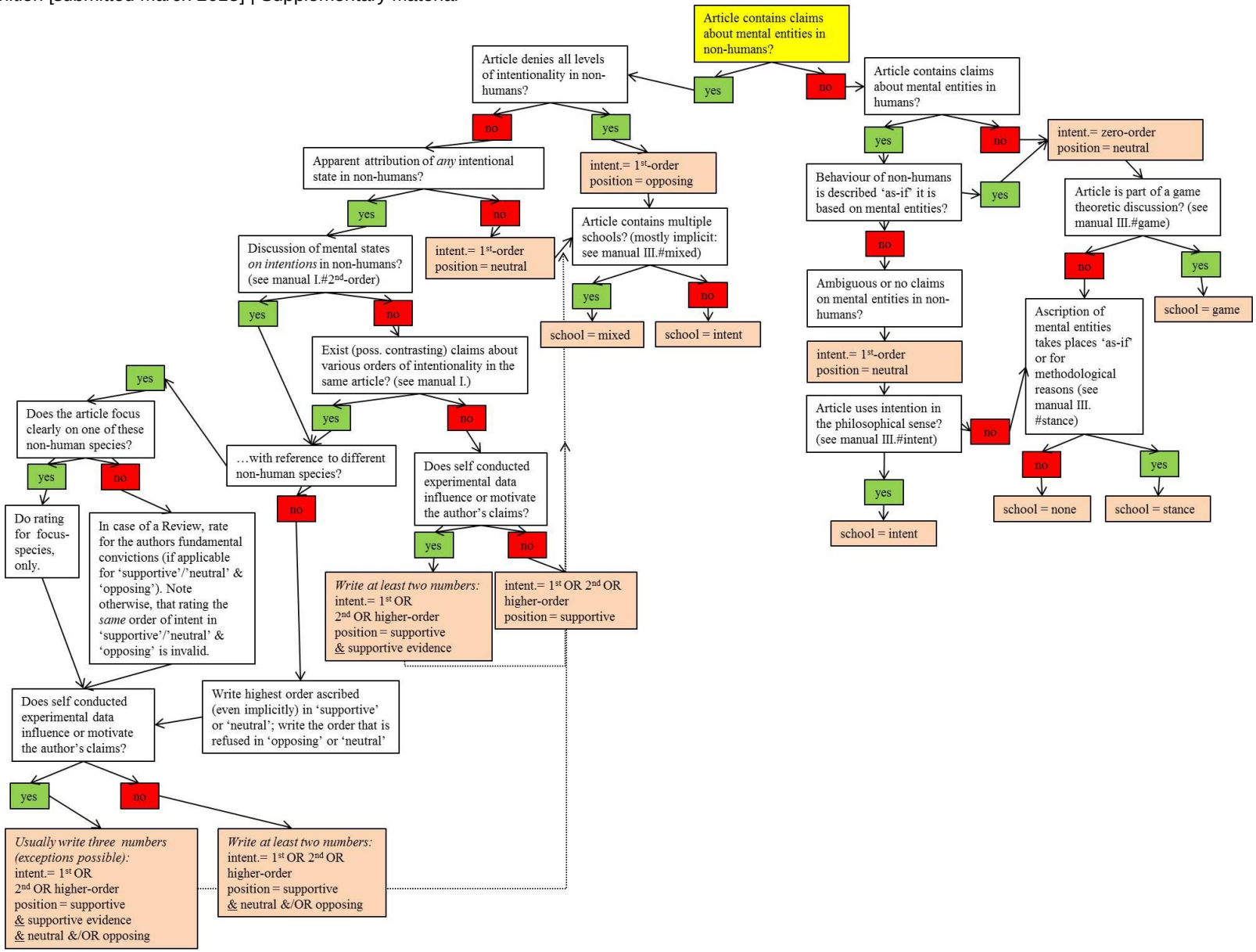

Figure 5 The decision diagram used for rating the articles. Ratings are conducted for the school of thought ('game'; 'stance'; 'intention'; 'mixed'), the order of intentionality (1st-; 2nd-; 3rd-order), and the author's position (supportive; supportive-evidence; neutral; opposing).

Full Rating Manual

I. Determining ORDERS of INTENT

Categorisation based on (Dennett, 1983). Note, an article can entail *multiple* orders of intention; i.e. when authors support 1st-order-intention, but oppose 2nd-order (see II.)

ZERO-ORDER-INTENTIONALITY: rate as '0'

>> "an account that attributes no mentality, no intelligence, no communication, no intentionality at all" (Dennett, 1983, p. 346)
>> also known as "biological purposiveness" (Millikan, 1997)
>> rests upon stimulus and response mechanisms (e.g. the heart beats, because it's purpose is to provide the body with oxygen; however, the heart does not 'believe' that it beats because of 'this')
>> hints that help to identify an article as zero-order:
(i) intention is equated with motivation
(ii) key word: 'aggressive intent'
(iii) key word: 'benign intent'
(iv) key word OR concept: 'game theory' (there is a debate about 'honest/deceptive signalling')
(v) key word OR concept: 'intention movement' (= movements that let the researcher predict what the animal is going to do next)

1ST-ORDER-INTENTIONALITY: rate as '1'

>> "A first-order intentional system has beliefs and desires [...] but no beliefs and desires about beliefs and desires." (Dennett, 1983, p. 345)

>> no understanding of the mind of 'x'

>> "If a thinking subject is not thinking about intentional states but is thinking about something else – anything else – then that subject is in a first-order intentional state" (Browne, 2004, p. 634)

>> hints that help to identify an article as 1st-order:

(i) consciousness can be 1st-order as long as it does not reflect intentional-states by itself

(ii) if an animal 'reads' 'attentional states' it does not always entail reading 'intentions', but reading behaviour (= zero-order)

(iii) the term 'functional reference' might hint to 1st-order - BUT check if authors accept concepts OR representations in general

(iv) articles that accept so called 1.5-order-intentionality, must be rated as 1st-order; i.e. authors that agree in reading 'goals', but not 'minds'

(v) 'social agents' are not 'intentional agents' - former must be rated as 1st-order

(vi) Note; if an author uses 'intent*' but does not specify the order OR it is not clear from text, assume 1st-order. Better rate too low, than too high.

(vii) Rating '1' in column 'neutral' is the minimum for an article from 'school' = 'intention' (see III.) (Even, if the entire article is about refuting 2nd-order-intentionality in a species. But if the authors would refute mental states entirely, they presumably prefer other theoretical frameworks. If so, revise rating in 'school').

2ND-ORDER-INTENTIONALITY: rate as '2'

>> "A second-order intentional system [...] has beliefs and desires [...] about beliefs and desires [...] - both those of others and its own." (Dennett, 1983, p. 345)

>> "If a subject is thinking about intentional states, then that subject is in a higher-order intentional state." (Browne, 2004, p. 634)

>> hints that help to identify an article as 2nd-order:

(i) subjects are described to read minds / intentions (no matter if 'rudimentary', 'basic' or 'simple')

(ii) subjects are described to perceive others as 'intentional agents'

(iii) subjects are described to deceive others with the knowledge of the others intent

(iv) some authors claim that 'joint intentionality' rests on 2nd-order

(v) also, some authors claim 'communication' rests on 2nd-order. BUT careful, others do not share that assumption.

HIGHER-ORDER-INTENTIONALITY: rate as '3'

>> "A third-order intentional system is one that is capable of such states as 'x wants y to believe that x believes he is all alone'" (Dennett, 1983, p. 345)

>> hints that help to identify an article as higher-order:

(i) an animal x knows that animal y 'knows' that animal x wants the banana

(ii) some authors claim that to tactically deceive others rests on higher-order-intentions

(iii) also, some authors mention, that passing the false believe test is an example for higher-order-intentions

II. Determining the author's POSITION

Authors can 'support', 'oppose' or stay 'neutral' on a topic. They can present 'supportive evidence' or support lower orders of intention, while opposing higher ones. According to this, articles can receive one or multiple ratings depending on the author's convictions as apparent alone from the text. Use columns 'supportive-evidence', 'supportive', 'neutral', or 'opposing' for ratings.

 # Articles with one rating

NEUTRAL

>> Place the rating made under 'Determining ORDERS of INTENT' in column 'neutral'.

>> hints that help to identify an article as 'neutral':

(i) author mentions intention, but does not focus on it

(ii) no clear supportive or opposing statement within the text

(iii) by definition all articles from 'school' (see III.) = 'game' OR 'stance' receive a '0'-rating in column 'neutral'

(iv) articles that make assertions about intention in humans, but leave it open for non-humans

(v) authors that do not make conclusive statements at all, because they await more evidence

OPPOSING

>> Place the rating made under 'Determining ORDERS of INTENT' in column 'opposing'.

>> hints that help to identify an article as 'opposing':

(i) author denies the idea of 1st, 2nd, or higher-order intention in non-humans

(ii) author presents evidence against any order of intention

(iii) author finds evidence in human children but believes it's 'unlikely' in non-humans

SUPPORTIVE

>> Place the rating made under 'Determining ORDERS of INTENT' in column 'supportive'.

>> hints that help to identify an article as 'supportive':

(i) author promotes the idea of intention in non-humans

(ii) author reviews supportive literature on that topic without discussing, limiting or rejecting the cited results

SUPPORTIVE-EVIDENCE

>> Place the rating made under 'Determining ORDERS of INTENT' in column 'supportive-evidence'.

>> hints that help to identify an article as 'supportive-evidence':

(i) author promotes intention AND the experimental/observational data influences the authors interpretation on that issue

(ii) based on the experimental results, authors consider intentional interpretations of behaviour

(iii) Note, the experimental results themselves must not directly support evidence for 'intention' in non-humans, but merely motivate researchers to use an 'intentional'-explanation

Articles with multiple ratings

>> Articles can receive more than one rating if multiple positions are promoted within the same article.

>> Giving an article the same rating in column 'supportive'/'neutral' and 'opposing' is invalid.

>> If there are multiple positions on multiple species, only make your rating for the species in focus of that article.

>> If there are multiple positions on the same species (e.g. juvenile and adults), rate for the highest order of intention ascribed or denied by the author.

>> In case of a review, or where experiments are done on multiple species (e.g. dogs and apes), rate for the highest order of intention ascribed or denied by the author.

>> *Example 1*: If an author promotes 1st-order, but opposes 2nd-order-intentionality, you must rate '1' at column 'supportive' and '2' at column 'opposing'

>> *Example 2*: If an author presents experimental evidence for 1st-order-intentional-behaviour, but leaves it open if results can be interpreted as 2nd-order, you must rate '1' at column 'supportive-evidence' & 'supportive' and '2' at column 'neutral'

>> *Example 3*: Authors conduct experiments on monkeys. The results for them refute 2nd-order, but allow the ascription of 1st-order intentionality. However, the authors cite studies that find 2nd-order-intentionality in apes. Rate only for the monkeys: '1' in column 'supportive' & 'supportive evidence', but '2' in column 'opposing'.

>> *Example 4*: An author conducts same or similar experiments on dogs and apes. The result for dogs refutes 2nd-order-intentionality, but allows it for apes. Take the highest ascription of intention and rate '2' in column 'supportive-evidence' & 'supportive'. Do not rate for the dog's negative outcome.

III. Determining theoretical framework - THE SCHOOL

Rating 'The School' means to categorise which theoretical framework authors use. Use column 'school' for ratings.

write: STANCE

>> When authors use 'intention' in order to describe zero-order-intentional behaviour.

>> hints that help to identify an article as 'stance':

(i) 'intention' often is used 'as if' the animal would have beliefs

(ii) describing a behaviour as intentional is done for methodological reasons and will be re-translated into zero-order-intention (= 'intentional stance')

(iii) typical indication: use of 'aggressive intent' in bird song (= birds do not 'intent', but respond to stimulus)

write: GAME

>> When authors use 'intention' in order to describe zero-order-intentional behaviour AND

>> When their discussion circulates about the question how 'honest signalling could have evolved'.

>> hints that help to identify an article as 'game theory':

(i) reference to 'game theory'

(ii) usage of: 'selective advantage'

(iii) usage of: 'accurate information'

(iv) citation of work from John Maynard Smith

write: INTENT

>> When authors use 'intention' in the philosophical sense as psychological phenomenon.

>> They do interpret behaviour from 1st-order upwards.

>> They do not use it as mere 'theory' or 'stance' or 'methodological issue', but as a certain perspective to ascribe mental states.

>> hints that help to identify an article as 'intent':

(i) usage of: 'believe', 'deception', 'representation'

(ii) attribution of not exactly observable mental states

(iii) author refers to the theoretical framework of Michael Tomasello

write: MIXED

>> When authors use 'intention' in the philosophical sense as psychological phenomenon AND

>> When authors use 'intention' in order to describe zero-order-intentional behaviour.

>> hints that help to identify an article as 'mixed':

(i) theories from several schools are discussed; e.g. 'intentional stance', behaviouristic explanations, and concepts of intention in the philosophical rich sense

(ii) one and the same article cites positions from e.g. John Maynard Smith, Daniel Dennett, Ruth Milikan

(iii) authors describe ‘intentional behaviour’ as an outcome of anthropocentric interpretation

write: NONE

>> Rate article as 'none' when no school is applicable.

>> hints that help to identify an article as 'none':

(i) 'intention' is used only few times within the articles in vague context

(ii) underlying definition is unclear

write: TECHNICAL or DELETE

>> Rate article as technical / delete when 'intention' is used less than two times in entire article

>> Rate article as technical / delete when 'intention' is used in linguistic sense, only.

Supplementary Results

Table 2 List of the 20 most cited articles within the corpus (Pos. 1-20). In addition to the absolute number of citations, authority scores and hub scores are provided as computed by Kleinberg's authority centrality scores, using the R package {igraph}. Authority scores from 1 (high) to 0 (low) represent articles that are cited by many hubs. Hub scores from 1 (high) to 0 (low) represent articles that cite many authorities. The 10 articles with the highest authority score are marked in black. The following 10 articles with high authority scores are marked in grey. Seven articles with less high authority scores are shown in white.

Pos.	Authors, Year	Title	Cites	Authority Score	Hub Score
1	Premack & Woodruff, 1978	Does the chimpanzee have a theory of mind?	100	0.71	1.6E-16
2	Call & Tomasello, 1994	Production and comprehension of referential pointing by orangutans	79	1	0.37
3	Woodruff & Premack, 1979	Intentional communication in the Chimpanzee	76	0.64	0.05
4	Tomasello et al, 1994	The learning and use of gestural signals by young Chimpanzees	68	0.89	0.11
5	Hare, Call, Bryan, & Tomasello, 2000	Chimpanzees know what conspecifics do and do not see	63	0.59	0.49
6	Povinelli & Eddy, 1996c	Reconstructing the evolution of psychological development	61	0.63	0.57
7	Hare, Call, & Tomasello, 2001	Do chimpanzees know what conspecifics know?	55	0.53	0.34
8	Byrne & Whiten, 1988	Tactical deception in primates	55	0.42	0.13
9	Leavens, Hopkins, & Bard, 1996	Indexical and referential pointing in chimpanzees	53	0.74	0.31
10	Povinelli, Nelson, & Boysen, 1990	Inferences about guessing and knowing by chimpanzees	53	0.54	0.16
11	Povinelli, Parks, & Novak, 1992	Role reversal by rhesus-monkeys, but no evidence of empathy	52	0.3	0.21
12	Leavens & Hopkins, 1998	Intentional communication by Chimpanzees	50	0.71	0.47
13	Leavens, Russell, & Hopkins, 2005	Intentionality as measured in the persistence and elaboration of communication by chimpanzees	50	0.64	0.68
14	Tomasello et al, 1985	The development of gestural communication in young Chimpanzees	46	0.48	0.02
15	Tomasello et al, 2005	Understanding and sharing intentions	46	0.34	0.38
16	Tomasello, Call, & Hare, 2003	Chimpanzees understand psychological states - the question is which ones and to what extent	42	0.40	0.29
17	Genty, Breuer, Hobaiter, & Byrne, 2009	Gestural communication of the gorilla	38	0.51	0.33
18	Dennett, 1983	Intentional Systems in Cognitive Ethology	35	0.16	0.11
19	Hare & Tomasello, 2004	Chimpanzees are more skilful in competitive than in cooperative cognitive tasks	35	0.36	0.23
20	Heyes, 1998	Theory of mind in nonhuman primates	35	0.32	0.36
21	Leavens, Hopkins, & Thomas, 2004	Referential Communication by Chimpanzees (Pan Troglodytes)	35	0.53	0.52
22	Hostetter, Cantero, & Hopkins, 2001	Differential Use of Vocal and Gestural Communication by Chimpanzees (Pan troglodytes) in Response to the Attentional Status of a Human (Homo sapiens)	34	0.54	0.38
23	Bard, 1992	Intentional Behavior and Intentional Communication in Young Free-Ranging Orangutans	34	0.46	0.06
24	Liebal, Call, & Tomasello, 2004	Use of gesture sequences in chimpanzees	34	0.46	0.36
25	Leavens, Hostetter, Wesley, & Hopkins, 2004	Tactical use of unimodal and bimodal communication by chimpanzees	33	0.51	0.52
26	Povinelli, Rulf, & Bierschwale, 1994	Absence of knowledge attribution and self-recognition in young chimpanzees (Pan troglodytes)	32	0.44	0.17
27	Krause & Fouts, 1997	Chimpanzee (Pan troglodytes) Pointing	30	0.45	0.3

Table 3 List of the top 10 articles that, while not part of the corpus, are cited by articles in the corpus. The reasons why each publication did not meet the criteria to find its way into the corpus are provided in the 'Comment' column.

	1st author	Title; if applicable journal abbreviation	Year	Cites	Comment
1	Tomasello, M.	Primate Cognition	1997	83	Book
2	Goodall, J.	The Chimpanzees of Gombe.	1986	68	Book
3	Altmann, J	Observational study of behavior; BEHAVIOUR	1974	55	no use of intent*
4	Tomasello, M.	Five primate species follow the visual gaze of conspecifics; ANIM BEHAV	1998	51	no use of intent*
5	de Waal, F.	Chimpanzee Politics	1982	48	Book
6	Tomasello, M.	Origins of Human Communication	2008	47	Book
7	Seyfarth, R.M.	Vervet monkey alarm calls; ANIM BEHAV	1980	44	no use of intent*
8	Bates, E.	The acquisition of performatives prior to speech; MERRILL PALMER QUART	1976	42	not comparative
9	Tomasello, M.	Chimpanzees, Pan troglodytes, follow gaze direction geometrically, ANIM BEHAV	1999	42	no use of intent*
10	Dawkins, R.	Animal Signals: Information or Manipulation?´; in: Animal signals	1978	40	book chapter

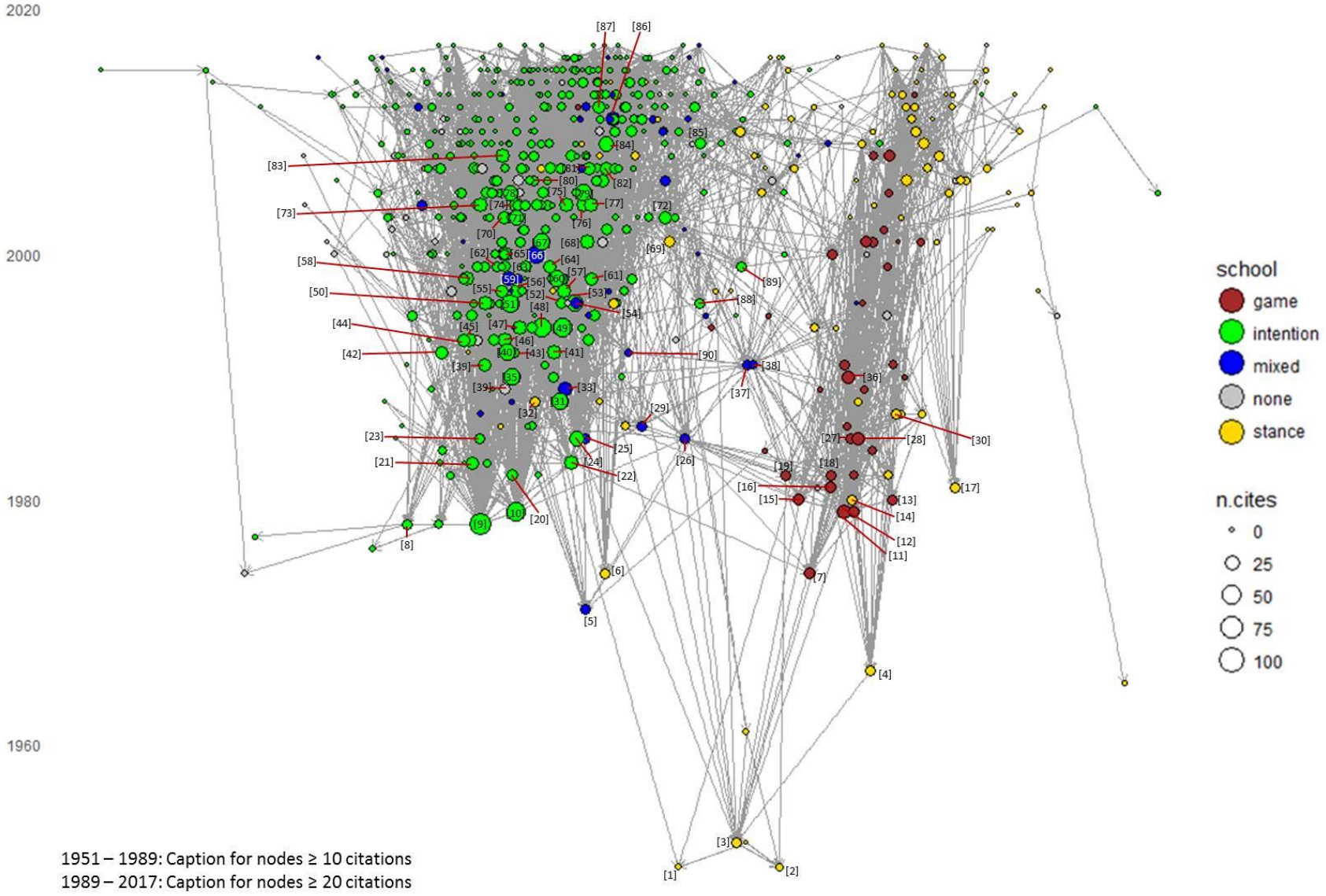

Figure 6 Annotating 90 of the most cited articles in the citation map. See next page for legend of the numbers.

Legend for Figure 6. Listed below are authors and year of publication. For full bibliographical information, see literature section.

[01] Lorenz, 1950
[02] Daanje, 1950
[03] Tinbergen, 1952
[04] Dunham, 1966
[05] Trivers, 1971
[06] Bekoff, 1974
[07] Maynard Smith, 1974
[08] Savage-Rumbaugh, Rumbaugh, & Boysen, 1978
[09] Premack & Woodruff, 1978
[10] Woodruff & Premack, 1979
[11] Caryl, 1979
[12] Maynard Smith, 1979
[13] Van Rhijn, 1980
[14] Andersson, 1980
[15] van Rhijn & Vodegel, 1980
[16] Hinde, 1981
[17] Krebs, Ashcroft, & van Orsdol, 1981
[18] Maynard Smith, 1982
[19] Moynihan, 1982
[20] Gallup, 1982
[21] Wimmer & Perner, 1983
[22] Dennett, 1983
[23] Gallup, 1985
[24] Tomasello, George, Kruger, Jeffrey, & Evans, 1985
[25] Hayaki, 1985
[26] Cheney & Seyfarth, 1985
[27] Enquist, Plane, & Röed, 1985
[28] Enquist, 1985
[29] Marler, Dufty, & Pickert, 1986
[30] Popp, 1987
[31] Byrne & Whiten, 1988
[32] de Waal, 1988
[33] Tomasello, Gust, & Frost, 1989
[34] Boesch & Boesch, 1989
[35] Povinelli, Nelson, & Boysen, 1990
[36] Grafen, 1990
[37] Guilford & Dawkins, 1991, 13 cites
[38] Hauser & Nelson, 1991; 8 cites
[39] Povinelli, Parks, & Novak, 1991
[40] Povinelli, Parks, & Novak, 1992
[41] Bard, 1992
[42] Whiten & Ham, 1992
[43] Byrne & Whiten, 1992
[44] Tomasello, Savage-Rumbaugh, & Kruger, 1993
[45] Tomasello, Kruger, & Ratner, 1993
[46] Heyes, 1993
[47] Povinelli, Rulf, & Bierschwale, 1994
[48] Call & Tomasello, 1994
[49] Tomasello, Call, Nagell, Olguin, & Carpenter, 1994
[50] Povinelli & Eddy, 1996a
[51] Povinelli & Eddy, 1996b, 1996c[158]
[52] Franco & Butterworth, 1996
[53] Leavens, Hopkins, & Bard, 1996
[54] Tanner & Byrne, 1996

[158] Note, in one case there were two articles from Povinelli that could not kept apart, because 1st-author, journal name and year of publication were identical. The articles are: Povinelli, 1996, MONOGR SOC RES CHILD DEV; DOI 10.1111/j.1540-5834.1996.tb00481.x AND Povinelli, 1996, MONOGR SOC RES CHILD DEV; DOI 10.2307/1166159. We decided to delete the article with DOI 10.1111/j.1540-5834.1996.tb00481.x. Instead the article with DOI 10.2307/1166159 received the number of citations from both publications. In the current case the node representing DOI 10.2307/1166159, becomes additionally overlaid from a third Povinelli article: Povinelli, 1996, J COMP PSYCHOL. For that reason the node labelled [51] in fact represents two, in a narrow sense even three Povinelli articles.

[55] Povinelli, Reaux, Bierschwale, Allain, & Simon, 1997
[56] Tomasello, Call, & Gluckman, 1997
[57] Krause & Fouts, 1997
[58] Call & Tomasello, 1998
[59] Heyes, 1998
[60] Leavens & Hopkins, 1998
[61] Veà & Sabater-Pi, 1998
[62] Povinelli, Bierschwale, & Čech, 1999
[63] Hare & Tomasello, 1999
[64] Leavens & Hopkins, 1999
[65] Emery, 2000
[66] Hare, Call, Bryan, & Tomasello, 2000
[67] Hare, Call, & Tomasello, 2001
[68] Hostetter, Cantero, & Hopkins, 2001
[69] Owren & Rendall, 2001
[70] Povinelli & Vonk, 2003
[71] Tomasello, Call, & Hare, 2003
[72] Seyfarth & Cheney, 2003
[73] Call, Hare, Carpenter, & Tomasello, 2004
[74] Hare & Tomasello, 2004
[75] Liebal, Pika, Call, & Tomasello, 2004; Leavens, Hopkins, et al., 2004[159]
[76] Leavens, Hostetter, Wesley, & Hopkins, 2004
[77] Liebal, Call, & Tomasello, 2004
[78] Tomasello, Carpenter, Call, Behne, & Moll, 2005
[79] Leavens, Russell, & Hopkins, 2005
[80] Liebal, Pika, & Tomasello, 2006
[81] Pollick & de Waal, 2007
[82] Cartmill & Byrne, 2007
[83] Call & Tomasello, 2008
[84] Genty, Breuer, Hobaiter, & Byrne, 2009
[85] Rendall, Owren, & Ryan, 2009
[86] Hobaiter & Byrne, 2011
[87] Crockford, Wittig, Mundry, & Zuberbühler, 2012
[88] Cheney, Seyfarth, & Palombit, 1996, 11 cites
[89] Rendall, Seyfarth, Cheney, & Owren, 1999, 14 cites
[90] Bekoff & Allen, 1992, 7 cites

[159] The node for Leavens, Hopkins, et al, 2004 is masked by Liebal et al, 2004

Supplementary Literature

Andersson, M. (1980). Why are there so many threat displays? *Journal of Theoretical Biology*, *86*(4), 773–781. http://doi.org/10.1016/0022-5193(80)90310-0

Bard, K. A. (1992). Intentional Behavior and Intentional Communication in Young Free-Ranging Orangutans. *Child Development*, *63*(5), 1186–1197. http://doi.org/10.1111/j.1467-8624.1992.tb01688.x

Bekoff, M. (1974). Social Play and Play-Soliciting by Infant Canids. *American Zoologist*, *14*(1), 323–340.

Bekoff, M., & Allen, C. (1992). Intenional Icons: Towards an Evolutionary Cognitive Ethology. *Ethology*, *91*(1), 89–107. http://doi.org/10.1111/j.1439-0310.1992.tb00846.x

Boesch, C., & Boesch, H. (1989). Hunting behavior of wild chimpanzees in the Tai National Park. *Am J Phys Anthropol*, *78*(4), 547–573. http://doi.org/10.1002/ajpa.1330780410

Browne, D. (2004). Do dolphins know their own minds? *Biology and Philosophy*, *19*(4), 633–653. http://doi.org/10.1007/sBIPH-004-0928-1

Byrne, R. W., & Whiten, A. (1988). Tactical deception in primates. *Behavioral and Brain Sciences*, *11*(2), 233–273. http://doi.org/10.1017/S0140525X00049682

Byrne, R. W., & Whiten, A. (1992). Cognitive Evolution in Primates: Evidence from Tactical Deception. *Man*, *27*(3), 609–627. http://doi.org/10.2307/2803931

Call, J., Hare, B., Carpenter, M., & Tomasello, M. (2004). "Unwilling" versus "unable": chimpanzees' understanding of human intentional action. *Developmental Science*, *7*(4), 488–498. http://doi.org/10.1111/j.1467-7687.2004.00368.x

Call, J., & Tomasello, M. (1994). Production and comprehension of referential pointing by Orangutans (Pongo pygmaeus). *Journal of Comparative Psychology*, *108*(4), 307–317. http://doi.org/10.1037/0735-7036.108.4.307

Call, J., & Tomasello, M. (1998). Distinguishing intentional from accidental actions in orangutans (Pongo pygmaeus). *Journal of Comparative Psychology*, *112*(2), 15p. http://doi.org/10.1037/0735-7036.112.2.192

Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences*, *12*(5), 187–192. http://doi.org/10.1016/j.tics.2008.02.010

Cartmill, E. A., & Byrne, R. W. (2007). Orangutans Modify Their Gestural Signaling According to Their Audience's Comprehension. *Current Biology*, *17*(15), 1345–1348. http://doi.org/10.1016/j.cub.2007.06.069

Caryl, P. G. (1979). Communication by Agonistic Displays: What Can Games Theory Contribute to Ethology? *Behaviour*, *68*(1), 136–169. http://doi.org/10.1163/156853979X00287

Cheney, D. L., & Seyfarth, R. M. (1985). Vervet Monkey Alarm Calls: Manipulation Through Shared Information? *Behaviour*, *94*(1), 150–166. http://doi.org/10.1163/156853985X00316

Cheney, D. L., Seyfarth, R. M., & Palombit, R. (1996). The function and mechanisms underlying baboon "contact" barks. *Animal Behaviour*, *52*, 507–518. http://doi.org/10.1006/anbe.1996.0193

Crockford, C., Wittig, R. M., Mundry, R., & Zuberbühler, K. (2012). Wild chimpanzees inform ignorant group members of danger. *Current Biology*, *22*(2), 142–146. http://doi.org/10.1016/j.cub.2011.11.053

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal*, *Complex Systems*, (5), 1695. Retrieved from http://igraph.org

Daanje, A. (1950). On Locomotory Movements in Birds and the Intention Movements Derived from Them. *Behaviour*, *3*(1), 48–98. http://doi.org/10.1163/156853951X00214

de Waal, F. B. M. (1988). The communicative repertoire of capitve bonobos (Pan paniscus) compared to that of chimpanzees. *Behaviour*, *106*(3), 183–251. http://doi.org/10.1163/156853988X00269

Dennett, D. C. (1983). Intentional systems in cognitive ethology: The "Panglossian paradigm" defended. *Behavioral and Brain Sciences*, *6*(3), 343. http://doi.org/10.1017/S0140525X00016393

Dunham, D. W. (1966). Agonistic Behavior in Captive Rose-Breasted Grosbeaks, Pheucticus ludovicianus. *Behaviour*, *27*(1), 160–173. http://doi.org/10.1163/156853966X00146

Emery, N. J. (2000). The eye have it: The neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews*, *24*(6), 581–604. http://doi.org/10.1016/S0149-

7634(00)00025-7
Enquist, M. (1985). Communication during aggresive interactions with particular reference to variation in choice of behavior. *Animal Behaviour*, *33*(NOV), 1152–1161. http://doi.org/10.1016/s0003-3472(85)80175-5
Enquist, M., Plane, E., & Röed, J. (1985). Aggressive communication in fulmars (Fulmarus glacialis) competing for food. *Animal Behaviour*, *33*(AUG), 1007–1020. http://doi.org/10.1016/S0003-3472(85)80035-X
Franco, F., & Butterworth, G. (1996). Pointing and social awareness: declaring and requesting in the second year. *Journal of Child Language*, *23*(2), 307–336. http://doi.org/10.1017/S0305000900008813
Gallup, G. G. (1982). Self-awareness and the emergence of mind in primates. *American Journal of Primatology*, *2*(3), 237–248. http://doi.org/10.1002/ajp.1350020302
Gallup, G. G. (1985). Do minds exist in species other than our own? *Neuroscience and Biobehavioral Reviews*, *9*(4), 631–641. http://doi.org/10.1016/0149-7634(85)90010-7
Gamer, M., Lemon, J., Fellows, I., & Singh, P. (2012). irr: Various Coefficients of Interrater Reliability and Agreement. Retrieved from https://cran.r-project.org/package=irr
Genty, E., Breuer, T., Hobaiter, C., & Byrne, R. W. (2009). Gestural communication of the gorilla (Gorilla gorilla): Repertoire, intentionality and possible origins. *Animal Cognition*, *12*(3), 527–546. http://doi.org/10.1007/s10071-009-0213-4
Grafen, A. (1990). Biological signals as handicap. *Journal of Theoretical Biology*, *144*(4), 517–546. http://doi.org/10.1016/S0022-5193(05)80088-8
Guilford, T., & Dawkins, M. S. (1991). Receiver psychology and the evolution of animal signals. *Animal Behaviour*, *42*(1), 1–14. http://doi.org/10.1016/S0003-3472(05)80600-1
Hare, B., Call, J., Bryan, A., & Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour*, *59*, 771–785.
Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour*, *61*(1), 139–151. http://doi.org/10.1006/anbe.2000.1518
Hare, B., & Tomasello, M. (1999). Domestic dogs (Canis familiaris) use human and conspecific social cues to locate hidden food. *Journal of Comparative Psychology*, *113*(2), 173–177. http://doi.org/10.1037//0735-7036.113.2.173
Hare, B., & Tomasello, M. (2004). Chimpanzees are more skilful in competitive than in cooperative cognitive tasks. *Animal Behaviour*, *68*(3), 571–581. http://doi.org/10.1016/j.anbehav.2003.11.011
Hauser, M. D., & Nelson, D. A. (1991). "Intentional" Signaling in Animal Communication. *Trends in Ecology & Evolution*, *6*(6), 186–189. http://doi.org/10.1016/0169-5347(91)90211-F
Hayaki, H. (1985). Social play of juvenile and adolescent chimpanzees in the Mahale Mountain National Park, Tanzania. *Primates*, *26*(4), 343–360. http://doi.org/10.1007/BF02382452
Heyes, C. M. (1993). Anecdotes, training, trapping and triangulating: do animals attribute mental states? *Animal Behaviour*, *461*(1), 177–188. http://doi.org/10.1006/anbe.1993.1173
Heyes, C. M. (1998). Theory of mind in nonhuman primates. *Behavioral and Brain Sciences*, *21*(1998), 101–134. http://doi.org/10.1017/S0140525X98000703
Hinde, R. A. (1981). Animal signals: Ethological and games-theory approaches are not incompatible. *Animal Behaviour*, *29*(2), 535–542. http://doi.org/10.1016/S0003-3472(81)80116-9
Hobaiter, C., & Byrne, R. W. (2011). The gestural repertoire of the wild chimpanzee. *Animal Cognition*, *14*(5), 745–767. http://doi.org/10.1007/s10071-011-0409-2
Hostetter, A. B., Cantero, M., & Hopkins, W. D. (2001). Differential Use of Vocal and Gestural Communication by Chimpanzees (Pan troglodytes) in Response to the Attentional Status of a Human (Homo sapiens). *Journal of Comparative Psychology*, *115*(4), 337–343. http://doi.org/10.1037//0735-7036.115.4.337
Krause, M. A., & Fouts, R. S. (1997). Chimpanzee (Pan troglodytes) Pointing: Hand Shapes, Accuracy, and the Role of Eye Gaze. *Journal of Comparative Psychology*, *111*(4), 330–336. http://doi.org/10.1037/0735-7036.111.4.330
Krebs, J. R., Ashcroft, R., & van Orsdol, K. (1981). Song matching in the great tit Parus major L. *Animal Behaviour*, *29*(3), 918–923. http://doi.org/10.1016/S0003-3472(81)80029-2
Leavens, D. A., & Hopkins, W. D. (1998). Intentional communication by Chimpanzees: A Cross-

Sectional Study of the Use of Referential Gestures. *Developmental Psychology*, *34*(5), 813–822. http://doi.org/10.1037/0012-1649.34.5.813

Leavens, D. A., & Hopkins, W. D. (1999). The Whole-Hand Point: The Structure and Function of Pointing From a Comparative Perspective. *Journal of Comparative Psychology*, *113*(4), 417–425. http://doi.org/10.1016/j.bbi.2008.05.010

Leavens, D. A., Hopkins, W. D., & Bard, K. A. (1996). Indexical and Referential Pointing in Chimpanzees (Pan troglodytes). *Journal of Comparative Psychology*, *4*(110), 346–353. http://doi.org/10.1037/0735-7036.110.4.346

Leavens, D. A., Hopkins, W. D., & Thomas, R. K. (2004). Referential Communication by Chimpanzees (Pan Troglodytes). *Journal of Comparative Psychology*, *118*(1), 48–57. http://doi.org/10.1037/0735-7036.118.1.48

Leavens, D. A., Hostetter, A. B., Wesley, M. J., & Hopkins, W. D. (2004). Tactical use of unimodal and bimodal communication by chimpanzees, Pan troglodytes. *Animal Behaviour*, *67*(3), 467–476. http://doi.org/10.1016/j.anbehav.2003.04.007

Leavens, D. A., Russell, J. L., & Hopkins, W. D. (2005). Intentionality as measured in the persistence and elaboration of communication by chimpanzees (Pan troglodytes). *Child Development*, *76*(1), 291–306. http://doi.org/10.1111/j.1467-8624.2005.00845.x

Liebal, K., Call, J., & Tomasello, M. (2004). Use of gesture sequences in chimpanzees. *American Journal of Primatology*, *64*(4), 377–396. http://doi.org/10.1002/ajp.20087

Liebal, K., Pika, S., Call, J., & Tomasello, M. (2004). To move or not to move. How apes adjust to the attentional state of others. *Interaction Studies*, *5*(2), 199–219. http://doi.org/10.1075/is.5.2.03lie

Liebal, K., Pika, S., & Tomasello, M. (2006). Gestural communication of orangutans (Pongo pygmaeus). *Gesture*, *6*(1), 1–38. http://doi.org/10.1075/gest.6.1.02lie

Lorenz, K. (1950). The comparative method in studying innate behavior patterns. In *Society for Experimental Biology, Physiological mechanisms in animal behavior. (Society's Symposium IV.)* (pp. 221–254). Oxford: Academic Press.

Marler, P., Dufty, A., & Pickert, R. (1986). Vocal communication in the domestic chicken: II. Is a sender sensitive to the presence and nature of a receiver? *Animal Behaviour*, *34*(1), 194–198. http://doi.org/10.1016/0003-3472(86)90023-0

Maynard Smith, J. (1974). The theory of games and the evolution of animal conflict. *Journal of Theoretical Biology*, *47*(1), 209–221. http://doi.org/10.1016/0022-5193(74)90110-6

Maynard Smith, J. (1979). Game Theory and the Evolution of Behaviour. *Proceedings of the Royal Society B: Biological Sciences*, *205*(1161), 475–488. http://doi.org/10.1098/rspb.1979.0080

Maynard Smith, J. (1982). Do animals convey information about their intentions? *Journal of Theoretical Biology*, *97*(1), 1–5. http://doi.org/10.1016/0022-5193(82)90271-5

Millikan, R. G. (1997). Varieties of Purposive Behavior. In R. W. Mitchell, N. S. Thompson, & H. L. Miles (Eds.), *Anthropomorphism, Anecdotes, and Animals* (pp. 189–197). Albany: SUNY Press.

Moynihan, M. (1982). Why is lying about intentions rare during some kinds of contests? *Journal of Theoretical Biology*, *97*(1), 7–12. http://doi.org/10.1016/0022-5193(82)90272-7

Neuwirth, E. (2014). RColorBrewer: ColorBrewer Palettes. Retrieved from https://cran.r-project.org/package=RColorBrewer

Owren, M. J., & Rendall, D. (2001). Sound on the rebound: Bringing form and function back to the forefront in understanding nonhuman primate vocal signaling. *Evolutionary Anthropology*, *10*(2), 58–71. http://doi.org/10.1002/evan.1014

Pedersen, T. L. (2017). ggraph: An Implementation of Grammar of Graphics for Graphs and Networks. Retrieved from https://cran.r-project.org/package=ggraph

Pollick, A. S., & de Waal, F. B. M. (2007). Ape gestures and language evolution. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(19), 8184–9. http://doi.org/10.1073/pnas.0702624104

Popp, J. W. (1987). Risk and Effectiveness in the Use of Agonistic Displays by American Goldfinches. *Behaviour*, *103*(1), 141–156. http://doi.org/10.1163/156853987X00314

Povinelli, D. J., Bierschwale, D. T., & Čech, C. G. (1999). Comprehension of seeing as a referential act in young children, but not juvenile chimpanzees. *British Journal of Developmental Psychology*, *17*(1), 37–60. http://doi.org/10.1348/026151099165140

Povinelli, D. J., & Eddy, T. J. (1996a). Chimpanzees: Joint Visual Attention. *Psychological Science*,

7(3), 129–136. http://doi.org/10.1111/j.1467-9280.1996.tb00345.x

Povinelli, D. J., & Eddy, T. J. (1996b). Factors influencing young chimpanzees' (Pan troglodytes) recognition of attention. *Journal of Comparative Psychology*, *110*(4), 336–345.

Povinelli, D. J., & Eddy, T. J. (1996c). Reconstructing the evolution of psychological development. *Monographs of the Society for Research in Child Development*, *61*(3), 1–16. http://doi.org/10.1111/j.1540-5834.1996.tb00481.x

Povinelli, D. J., Nelson, K. E., & Boysen, S. T. (1990). Inferences about guessing and knowing by chimpanzees (Pan troglodytes). *Journal of Comparative Psychology*, *104*(3), 203–210. http://doi.org/10.1037//0735-7036.104.3.203

Povinelli, D. J., Parks, K. A., & Novak, M. A. (1991). Do Rhesus Monkeys (Macaca mulatto) Attribute Knowledge and Ignorance to Others? *Journal of Comparative Psychology*, *105*(4), 318–325. http://doi.org/10.1037/0735-7036.105.4.318

Povinelli, D. J., Parks, K. A., & Novak, M. A. (1992). Role reversal by rhesus-monkeys, but no evidence of empathy. *Animal Behaviour*, *44*(2), 269–281. http://doi.org/10.1016/0003-3472(92)90033-6

Povinelli, D. J., Reaux, J. E., Bierschwale, D. T., Allain, A. D., & Simon, B. B. (1997). Exploitation of pointing as a referential gesture in young children, but not adolescent chimpanzees. *Cognitive Development*, *12*(4), 423–461. http://doi.org/10.1016/S0885-2014(97)90017-4

Povinelli, D. J., Rulf, A. B., & Bierschwale, D. T. (1994). Absence of knowledge attribution and self-recognition in young chimpanzees (Pan troglodytes). *Journal of Comparative Psychology*, *108*(1), 74–80. http://doi.org/10.1037//0735-7036.108.1.74

Povinelli, D. J., & Vonk, J. (2003). Chimpanzee minds: Suspiciously human? *Trends in Cognitive Sciences*, *7*(4), 157–160. http://doi.org/10.1016/S1364-6613(03)00053-6

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *The Behavioral and Brain Sciences*, *4*, 515–526.

R Development Core Team, R. (2016). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.r-project.org/

Rendall, D., Owren, M. J., & Ryan, M. J. (2009). What do animal signals mean? *Animal Behaviour*, *78*(2), 233–240. http://doi.org/10.1016/j.anbehav.2009.06.007

Rendall, D., Seyfarth, R. M., Cheney, D. L., & Owren, M. J. (1999). The meaning and function of grunt variants in baboons. *Animal Behaviour*, *57*(3), 583–592. http://doi.org/10.1006/anbe.1998.1031

Savage-Rumbaugh, S. E., Rumbaugh, D. M., & Boysen, S. (1978). Linguistically mediated tool-use and exchange by chimpanzees (Pan troglodytes). *Behavioral and Brain Sciences*, *4*(1), 539–554.

Seyfarth, R. M., & Cheney, D. L. (2003). Signalers and receivers in animal communication. *Annual Review of Psychology*, *54*, 145–173. http://doi.org/10.1146/annurev.psych.54.101601.145121

Sievert, C., Parmer, C., Hocking, T., Chamberlain, S., & Karthik, R. (2017). plotly: Create Interactive Web Graphics via "plotly.js". Retrieved from https://cran.r-project.org/package=plotly

Tanner, J. E., & Byrne, R. W. (1996). Representation of Action Through Iconic Gesture in a Captive Lowland Gorilla. *Current Anthropology*, *37*(1), 162–173. http://doi.org/10.1086/204484

Tinbergen, N. (1952). "Derived" Activities; Their Causation, Biological Significance, Origin, and Emancipation During Evolution. *The Quarterly Review of Biology*, *27*(1), 1–32. http://doi.org/10.1086/398642

Tomasello, M., Call, J., & Gluckman, A. (1997). Comprehension of Novel Communicative Signs by Apes and Human Children. *Child Development*, *68*(6), 1067–1080. http://doi.org/10.2307/1132292

Tomasello, M., Call, J., & Hare, B. (2003). Chimpanzees understand psychological states - The question is which ones and to what extent. *Trends in Cognitive Sciences*, *7*(4), 153–156. http://doi.org/10.1016/S1364-6613(03)00035-4

Tomasello, M., Call, J., Nagell, K., Olguin, R., & Carpenter, M. (1994). The Learning and Use of Gestural Signals by Young Chimpanzees: A Trans-generational Study. *Primates*, *35*(2), 137–154. http://doi.org/10.1007/BF02382050

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, *28*(5), 675–735.

http://doi.org/10.1017/S0140525X05000129
Tomasello, M., George, B. L., Kruger, A. C., Jeffrey, M. F., & Evans, A. (1985). The development of gestural communication in young chimpanzees. *Journal of Human Evolution*, *14*(2), 175–186. http://doi.org/10.1016/S0047-2484(85)80005-1
Tomasello, M., Gust, D., & Frost, G. T. (1989). A longitudinal investigation of gestural communication in young chimpanzees. *Primates*, *30*(1), 35–50. http://doi.org/10.1007/BF02381209
Tomasello, M., Kruger, A. C., & Ratner, H. H. (1993). Cultural learning. *Behavioral and Brain Sciences*, *16*(3), 495. http://doi.org/10.1017/S0140525X0003123X
Tomasello, M., Savage-Rumbaugh, S., & Kruger, A. C. (1993). Imitative Learning of Actions on Objects by Children , Chimpanzees, and Enculturated Chimpanzees. *Child Development*, *64*(6), 1688–1705. http://doi.org/10.1111/j.1467-8624.1993.tb04207.x
Trivers, R. L. (1971). The Evolution of Reciprocal Altruism. *The Quarterly Review of Biology*, *46*(1), 35–57. http://doi.org/10.1086/406755
Van Rhijn, J. G. (1980). Communication by agonistic displays: A discussion. *Behaviour*, *74*(3), 284–293. http://doi.org/10.1163/156853980X00500
van Rhijn, J. G., & Vodegel, R. (1980). Being honest about one's intentions: An evolutionary stable strategy for animal conflicts. *Journal of Theoretical Biology*, *85*(4), 623–641. http://doi.org/10.1016/0022-5193(80)90261-1
Veà, J. J., & Sabater-Pi, J. (1998). Spontaneous pointing behahaviour in the Wild Pygmy Chimpanzee (Pan paniscus). *Folia Primatologica*, *69*(5), 289–290. http://doi.org/10.1159/000021640
Whiten, A., & Ham, R. (1992). *On the Nature and Evolution of Imitation in the Animal Kingdom: Reappraisal of a Century of Research*. *Advances in the Study of Behavior* (Vol. 21). Elsevier Masson SAS. http://doi.org/10.1016/S0065-3454(08)60146-1
Wickham, H. (2009). ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag. Retrieved from http://had.co.nz/ggplot2/book
Wickham, H. (2016). tidyr: Easily Tidy Data with `spread()` and `gather()` Functions. Retrieved from https://cran.r-project.org/package=tidyr
Wickham, H., & Francois, R. (2016). dplyr: A Grammar of Data Manipulation. Retrieved from https://cran.r-project.org/package=dplyr
Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*, 103–128. http://doi.org/10.1016/0010-0277(83)90004-5
Woodruff, G., & Premack, D. (1979). Intentional communication in the chimpanzee: The development of deception. *Cognition*, *7*(4), 333–362. http://doi.org/10.1016/0010-0277(79)90021-0

General Discussion

The dissertation's aim was to investigate values and norms that influence constructions of language in human and non-human animals. In order to avoid misunderstandings note that the dissertation was not meant to answer prominent questions such as: 'What is language?' or 'How did language evolve?', although these questions certainly motivated the project in the first place. Instead, the target of investigation was not *language* as such, but is placed in *researchers* and *articles.* They were examined for implicit background assumptions, in order to detect and describe systemic patterns. Those patterns might influence recent and past constructions of *language* and thereby experimental designs, the choice of study species and the weighting of evidence. While investigating three chosen values in the course of the dissertation project, some questions turned up repeatedly: (1) Why do researchers, interested in language evolution, predominantly investigate primates, more so than birds? (2) When do they use the term *communication* and when do they refer to *language*? (3) How did various constructions of *language* evolve over time; and (4) what were the reasons for shifting them?

Those apparent questions emerged from the beginning of the project and were present throughout all publications. While question (1) was directly addressed in *publication 2*, the more general questions (2)-(4) are only implicitly answered in the course of the project. For that reason this section will discuss them more directly in order to collect possible answers on more far-reaching issues. No single visualisation, chart or analysis can provide a clear answer to them. Nonetheless, collectively, the results help to approach an overall conclusion.

In concrete terms, relating to (1), one could answer that researchers actually do investigate birds – mostly songbirds – when they are interested in language evolution. They investigate them, because songbirds are vocal learners, like humans. However, researcher often ask proximate questions on bird signalling (e.g. *How* do they learn their vocalisations?) instead of ultimate ones (e.g. What do their vocalisations mean?). When it comes to ultimate questions, it seems to be a common practice to predominantly investigate primates. Admittedly, that habit is changing in more recent research programmes (e.g. Fitch, Huber, & Bugnyar, 2010; Güntürkün, Ströckens, Scarf, & Colombo, 2017; Pepperberg, 2018). However, for the older articles which were in focus for the dissertation, the primate bias could be observed frequently. One obvious reason to ask ultimate instead of proximate questions is that the underlying neuronal and cognitive mechanisms on which language is built remain poorly identified. Another reason is that primates share more human characteristics and behaviours than birds (e.g. they have two frontal eyes, arms, and breast-feed their offspring). These similarities seem to motivate research on primates. However, intuitive similarities do not allow conclusions about possible cases of co-evolution and deep homology. Both do not rest on close relationship. For instance, tool-use (Bird & Emery, 2009) or the role of the FoXP2 transcription factor for vocal learning (Scharff & Petri, 2011) in birds are examples of co-evolution and deep-homology in distantly

related species. Still, they were less frequently investigated as compared to primates. The reason that primates in the past are more often investigated to better understand scenarios of language evolution cannot be explained by their outstanding abilities. Quite the contrary, as soon as researchers started investigating the communicative and cognitive abilities of songbirds, they have found that the abilities are not qualitatively different to primates, but comparable in degree (Güntürkün et al., 2017). The results of the dissertation provide evidence that the often observed primate-bias in language evolution research might express historical and social values, such as the idea that humans and their primate-relatives must represent 'higher species'. The consequent implications of *Scala-Naturae*-thinking were investigated and discussed in *publication 2* and *3* (Ullrich & Liebal, 2018; Ullrich et al., 2018).

Question (2) asks at which point researchers make use of the terms *language* and *communication*. The motivation for asking this question was the observation that some researchers categorically reject a comparison of human language and animal communication (e.g. Scott-Phillips, 2014). Scientists who use the term *language* might want to stress the human unique aspect of it. Conversely, researchers who refer to *communication* might want to stress the mere quantitative difference between human and non-human communication. Suggestive evidence for that interpretation can be found in the frequency of usage of the term *unique* and the selection of adjectives in corpus 'language' and 'communication' in *publication 2* (Ullrich et al., 2018). As all partial studies illustrate, the academic dispute is often around the question if human *language* differs qualitatively to animal communication. The results of the partial studies show that the answer to that dispute is a conceptual / philosophical one. As Ray Jackendoff once noted: "Your theory of language evolution depends on your theory of language" (Jackendoff, 2010).

Constructions underlying the theory of language evolved and still evolve. This concerns question (3) from above. For instance, during the era of behaviourism the cognitive infrastructure of language was only of little interest for scientific research programmes. Linguists, like Leonard Bloomfield were informed by the *behaviouristic norm* and consequently focussed on phonology, syntax, or morphology instead of intentionality, ambiguity, or context (Bloomfield, 1943; Levelt, 2013). In general, research on animal communication under the influence of behaviourism is mostly interested in form and structure of the *code* that underlies language (e.g. Doupe & Kuhl, 1999). This concerns also the often cited design features for language (e.g. Fitch, 2010, p. 19) as formulated by Charles Hockett (Hockett, 1960; Hockett & Altmann, 1968). As Wacewicz and Żywiczyński write, behaviourists like Hockett "focus on the means at the expense of content and focus on the code itself rather than the cognitive abilities of its users." (Wacewicz & Żywiczyński, 2015, p. 29). Because of that, the design features appear "radically unfit" for modern day research on language evolution which "needs a primarily 'internalistic' perspective" (Wacewicz & Żywiczyński, 2015, p. 42). The reason why Hockett's design features appear incompatible with recent developments in language evolution is based on the 'cognitive turn' that happened between 1960 and 1990. The dissertation put a focus on

that shift by investigating influential and prominent authors and articles of that period of time. However, the *cognitive shift* – as investigated in *publication 3* and *4* (Ullrich & Liebal, 2018; Ullrich, Mittelbach, & Liebal, submitted) – was only one of several shifts that influenced definitions of language. Wacewicz and Żywiczyński mention additional 'turns' to the cognitive, such as the 'biosemiotic turn' and the 'adaptive turn' (Wacewicz & Żywiczyński, 2015). While the first offers a biologically grounded perspective on how signs are produced and interpreted in a pre-linguistic manner (Favareau, 2008), the latter stresses that language must be understood by studying the natural selection and evolutionary history of it (Gontier & Pina, 2014, p. 6). There are still more shifts and more epistemological reassessments as mentioned here. Definitions of language and their theoretical groundings are ever changing. That observation motivated question (4) from above which asks why *language* constructions change anyway?

Constructions of *language* seem to serve a narrative function for the human self-description. Again, that might be illustrated by reference to the cognitive shift. As scientists started to investigate language evolution from the standpoint of behaviourism, more similarities between human language and animal communication were revealed (e.g. vocal learning, phonological syntax, etc.). Biologists like Marc Naguib concluded: "Most of the traits that were believed for decades to be unique to human language have been shown to be part of one or another communication system in animals." (Naguib, 2006). As many of Charles Hockett's design features of language were described in other non-human species as well, the overall conclusion was not to reject the unique aspect of human language. Instead some researcher argued that definitions of language were erroneous, because they were missing out an intrinsic perspective.

> *"There is, then, a critical difference between the codes used in code model communication, and the codes used in language"* (Scott-Phillips, 2014, p. 18)
>
> *They are made possible by different internal mechanisms. That is, the mechanisms that are causally responsible for the existence of each type of communication are different in each case: associations on the one hand, metapsychology on the other."* (Scott-Phillips, 2014, p. 12)
>
> *Instead, what I want to emphasize is that we cannot simply assume that the essential difference between the linguistic code and the code used in non-human primate communication is a difference of degree, even a large degree. It may instead be a difference of kind.* (Scott-Phillips, 2014, p. 18)

What is interesting from that section is that the author provides a conceptual / philosophical narrative of why and how *language* differs qualitatively from animal communication: While *language* supposedly rests on metapsychology, communication rests on associations. However, although the author stresses that the theoretical conception needs to be tested empirically and conclusive evidence does not yet exist, he continues speculating about consequences of the idea. Such is observable in the language evolution discourse frequently. Authors often complain that "the richness of ideas is accompanied by a poverty of evidence" (Hauser et al., 2014, p. 1). Still, the same authors

subsequently conclude that "For now, the evidence from comparative animal behavior provides little insight into how *our* language phenotype evolved. The gap between *us and them* is simply *too great* to provide any understanding of evolutionary precursors or the evolutionary processes (e.g., selection) that led to change over time." (Hauser et al., 2014, p. 5 italics R.U.)´. The question arises, how authors can note that there is a poverty of evidence, but subsequently conclude on that basis that language creates a gap "too great"?

The examples illustrate a left out problem in the evolution of language research: Researchers are an immanent part of their own investigations. This is revealed in an exemplary way by the use of pronouns highlighted in the above quotation. Any outcome and subsequent interpretation might rattle on long held personal convictions and thus may result in a personal conflict of interest. Conflicts of interest are well acknowledged in relation to financial interest (e.g. Thompson, 1993). This is why scientists are asked to declare financial interest when publishing in various journals. Similar to that case, researchers might be biased by their own self-perception. This point was observed earlier by various researchers such as Ulrich Frey[160] or Woolard and Schiefflein[161]. It is for that reason that researcher working on language evolution should make transparent their personal academic background and motivation for working on that topic. Knowing the background of a researcher can help explaining some epistemological decisions of a research programme and might help in evaluating the evidence provided in a research paper. This is also the reason why section 1 (*General introduction,* p. 2) provides information on the background and motivation of the author of these lines. Transparency is important to monitor science and its unintended influences. On that note, it is not only important to keep open and transparent raw-data and protocols, but also work flow and personal motivation.

To make transparent the work flow is the aim of that last part of the dissertation's discussion. It will explain some changes in the conception of the project. Perhaps the most apparent change concerns the term 'norm' which became increasingly replaced by the alternative notion 'value' or 'social and historical influence'. The reason is that there are too many influences affecting the constructions of *language*. Only few of them fit the definition of a norm as explained in *publication 1* (Ullrich, 2016). Other influences might be called 'non-epistemological', 'cognitive attitudes', 'biases', 'prejudices', 'attributions', or 'motivations'. Authors from more socio-historical research fields often subsume those influences under the umbrella term 'value' (for review see Elliott & Willmes, 2014; Kunda, 1990). In order to harmonize the terminology, later publications of the dissertation employ the same term.

[160] "Some scientific discoveries contradict core beliefs of our perception of ourselves and the world. […] It is argued that conceptual revolutions imply restructuring one's beliefs." (Frey, 2010, p. 4)

[161] "Ideologies of language are significant for social as well as linguistic analysis because they are not only about language. Rather, such ideologies envision and enact links of language to group and personal identity, to aesthetics, to morality, and to epistemology.[…] As R. Williams observed, 'a definition of language is always, implicitly or explicitly, a definition of human beings in the world'." (Woolard & Schiefflein, 1994, p. 56)

The advent of meta-research was another factor which had influenced the overall dissertation project. Meta-research emerged when a number of spectacular cases of scientific fraud initiated a crisis in science (see *General introduction*; section 3; p.10). The deeper reason for that crisis was the conclusion that the majority of studies are not intentionally fabricated, but still hard to replicate and likely to contain false data (Ioannidis, 2012; Open Science Collaboration, 2015). That put the focus on *questionable research practices* which are supposed to serve as real cause of the so called 'replication crisis' (Munafò et al., 2017; Pashler & Harris, 2012) and which were described and quantified in order to evaluate, reassess, or improve scientific practice.

For the sake of the argument, just assume that it might be possible to remove all *questionable research practices* from science. Still, the influence of values would continue to exist (cf. Elliott & Willmes, 2014). This is where the dissertation project enters the meta-research discourse. It extends the current attempt of quantifying *questionable research practices*, by putting the focus on conceptual / philosophical assumptions.

As most efforts to investigate the socio-historical influences on science were made by historians or sociologists in a qualitative manner, the current dissertation wanted to employ methods from the quantitative spectrum of meta-research. The dissertation developed a mixed-methods strategy to reveal and describe values that leave their mark in published papers. In times of discussions on 'fake news' and 'alternative facts', the proclamation and description of values might cast a poor light on science. However, to deny those influences will have worse effects (Douglas, 2016; Elliott, 2013). In describing systemic instead of individual biases, the dissertation wants to contribute to the process of scientific self-correction.

The field of meta-research is developing fast. Frequently, new meta-research tools for analysing data are published and developed, i.e. for the coding language R. One example is the R-package 'bibliometrix' which creates an automatic comprehensive citation-network (Aria & Cuccurullo, 2017). The package was published after a manually-written script was just finished for the dissertation project. The publication of 'bibliometrix' came too late to be used for the current project. Similar technical developments influenced the project's scope and created new ideas, which were not planned from the very beginning - e.g. the text-mining procedure. However, the overall project might serve as an idea generator for further packages which have just started to develop. The quantification of *questionable research practices* was a famous issue for meta-research in the last 10 years. In comparison, the quantification of *values* in science has hardly begun. The dissertation project is one attempt to introduce more quantitative methods into that research field. Investigating academic relationships (e.g. completing the fragmentary database *academictree.org*), investigating the influences between co-authors (e.g. who cooperated and published with whom) and analysing citation habits of researchers (e.g. which authors and articles are *not* mentioned), might be productive new research attempts for future projects. The increasing digitalisation of research papers and meta-

information will allow a number of new ideas and projects in the future. Meta-research is still in its early stages. There is much more to come.

All things considered, the dissertation has shown that various constructions of *language* existed in the past and continue to exist. Many constructions became obsolete as soon as their value-laden foundations were revealed: Language is not anymore constructed as a modality, a measure for advancement, mere code or structure, or an organ. Nowadays language is widely constructed as a biological entity that rests on an evolutionary shared substrate, but it is unclear which parts of the substrate are human unique, crucial, or genetically determined. Therewith the dissertation reached the most recent trend in constructing *language*. Instead of focussing "on the pinnacles of mental evolution, [and] asking all-or-nothing questions" (de Waal & Ferrari, 2010) researchers increasingly put their attention towards a 'bottom-up perspective'. That perspective investigates basic building blocks of not only language, but cognition in general. Current research is interested in neuroscientific perspectives and evolutionary biology, likewise. Future meta-research will have to devote attention to those recent constructions of language as they are certainly affected by values and socio- historical influences. This is not bad news, but simply the reality of doing science. Scientists are part of a society and herewith influenced by values. Values by themselves are not the problem. They become a problem as soon as they are unintended and implicit and by this means spread without limitation through the discourse. Continuous reflection, meta-research, and avoidance of historical amnesia can prevent the negative impacts, and instead unfold their positive effects for cross-species comparative research on language evolution.

References

Abe, K., & Watanabe, D. (2011). Songbirds possess the spontaneous ability to discriminate syntactic rules. *Nature Neuroscience*. http://doi.org/10.1038/nn.2869

Alasuutari, P. (2010). The rise and relevance of qualitative research. *International Journal of Social Research Methodology*, *13*(2), 139–155. http://doi.org/10.1080/13645570902966056

Allchin, D. (1999). Values in Science: An Educational Perspective. *Science Education*, *8*(1), 1–12. http://doi.org/10.1023/A:1008600230536

Allen, T. A., & Fortin, N. J. (2013). The evolution of episodic memory. *PNAS*, *110 Suppl*, 10379–86. http://doi.org/10.1073/pnas.1301199110

Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, *11*(4), 959–975. http://doi.org/10.1016/j.joi.2017.08.007

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature*, *533*(7604), 452–454. http://doi.org/10.1038/533452a

Beasley, M. R., Datta, S., Kogelnik, H., Kroemer, H., & Monroe, D. (2002). *Report of the Investigation Committee on the possibility of Scientific Misconduct in the work of Hendrik Schon and Coauthors*. *Bell Labs*.

Begley, C. G., & Ioannidis, J. P. A. (2015). Reproducibility in science: Improving the standard for basic and preclinical research. *Circulation Research*, *116*(1), 116–126. http://doi.org/10.1161/CIRCRESAHA.114.303819

Bentley-Condit, V. K., & Smith, E. O. (2010). Animal tool use: current definitions and an updated comprehensive catalog. *Behaviour*, *147*(2), 185–221. http://doi.org/10.1163/000579509X12512865686555

Betz, G. (2013). In defence of the value free ideal. *European Journal for Philosophy of Science*, *3*(2), 207–220. http://doi.org/10.1007/s13194-012-0062-x

Bickerton, D. (2009). *Adam's Tongue. How Humans Made Language, How Language Made Humans.* New York: Hill and Wang.

Bird, C. D., & Emery, N. J. (2009). Rooks Use Stones to Raise the Water Level to Reach a Floating Worm. *Current Biology*, *19*(16), 1410–1414. http://doi.org/10.1016/j.cub.2009.07.033

Bloomfield, L. (1943). Meaning. *Monatshefte Für Deutschen Unterricht*, *35*(3), 101–106.

Bluff, L., Weir, A., & Rutz, C. (2007). Tool-related cognition in new Caledonian crows. *Comparative Cognition & Behavior Reviews*, *2*, 1–25.

Boesch, C., & Boesch, H. (1990). Tool Use and Tool Making in Wild Chimpanzees. *Folia Primatologica*, *54*, 86–99.

Bolhuis, J. J., Okanoya, K., & Scharff, C. (2010). Twitter evolution: converging mechanisms in birdsong and human speech. *Nature Reviews. Neuroscience*, *11*(11), 747–59. http://doi.org/10.1038/nrn2931

Bornmann, L., & Mutz, R. (2015). Growth Rates of Modern Science: A Bibliometric Analysis Based on the Number of Publications and Cited References. *Journal of the Association for Information Science and Technology*, *66*(11), 2215–2222. http://doi.org/10.1002/asi.23329

Botha, R. (2000). Discussing the evolution of the assorted beasts called language. *Language & Communication*, *20*, 149–160. http://doi.org/10.1016/S0271-5309(99)00022-1

Brembs, B., Button, K., & Munafò, M. (2013). Deep impact: unintended consequences of journal rank. *Frontiers in Human Neuroscience*, *7*(June), 1–12. http://doi.org/10.3389/fnhum.2013.00291

Brown, C. (2012). Tool use in fishes. *Fish and Fisheries*. http://doi.org/10.1111/j.1467-2979.2011.00451.x

Brumfiel, G. (2002). Misconduct finding at bell labs shakes physics community. *Nature*, *419*(6906), 419–421. http://doi.org/10.1038/419419a

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*(5), 365–376. http://doi.org/10.1038/nrn3475

Chalmers, I., & Glasziou, P. (2009). Avoidable waste in the production and reporting of research

evidence. *The Lancet*, *374*(9683), 86–89. http://doi.org/10.1016/S0140-6736(09)60329-9
Chan, A.-W., Hróbjartsson, A., Haahr, M. T., Gøtzsche, P. C., & Altman, D. G. (2004). Empirical Evidence for Selective Reporting of Outcomes in Randomized Trials. *The Journal of the American Medical Association*, *291*(20), 2457–2465. http://doi.org/10.1001/jama.291.20.2457
Chater, N., Reali, F., & Christiansen, M. H. (2009). Restrictions on biological adaptation in language evolution. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(4), 1015–1020. http://doi.org/10.1073/pnas.0807191106
Chomsky, N. (2011). Language and Other Cognitive Systems. What Is Special About Language? *Language Learning and Development*, *7*(4), 263–278. http://doi.org/10.1080/15475441.2011.584041
Chomsky, N. (2017). The language capacity: architecture and evolution. *Psychonomic Bulletin and Review*, *24*(1), 200–203. http://doi.org/10.3758/s13423-016-1078-6
Christiansen, M. H., & Kirby, S. (2003). Language evolution: consensus and controversies. *Trends in Cognitive Sciences*, *7*(7), 300–307. http://doi.org/10.1016/S1364-6613(03)00136-0
Commitee on Science Engineering and Public Policy. (1992). *Responsible science: Ensuring the integrity of the research process* (Vol. I). Washin. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK234523/pdf/Bookshelf_NBK234523.pdf
Dally, J. M., Emery, N. J., & Clayton, N. S. (2006). Food-caching western scrub-jays keep track of who was watching when. *Science (New York, N.Y.)*, *312*(5780), 1662–5. http://doi.org/10.1126/science.1126539
Darwin, C. (1963). *The Origin Of Species And The Descent Of Man [1871]*. New York: The Modern Library published by Random House.
de Vries, R., Anderson, M. S., & Martinson, B. C. (2006). Normal Misbehavior: Scientists Talk about the Ethics of Research. *Journal of Empirical Research on Human Research Ethics*, *1*(1), 43–50. http://doi.org/10.1525/jer.2006.1.1.43
de Waal, F. B. M., & Ferrari, P. F. (2010). Towards a bottom-up perspective on animal and human cognition. *Trends in Cognitive Sciences*, *14*(5), 201–207. http://doi.org/10.1016/j.tics.2010.03.003
Dorato, M. (2004). Epistemic and nonepistemic values in science. In P. Machamer & G. Wolters (Eds.), *Science, Values, and Objectivity* (pp. 52–77). Pittsburgh: University of Pittsburg Press.
Douglas, H. (2009). *Science, policy, and the value-free ideal*. Pittsburgh: University of Pittsburgh Press.
Douglas, H. (2016). Values in Science. In P. Humphreys (Ed.), *The Oxford Handbook of Philosophy of Science* (pp. 609–632). Oxford, New York: Oxford University Press. http://doi.org/10.1093/oxfordhb/9780199368815.013.28
Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Annual Review of Neuroscience*, *22*, 567–631. http://doi.org/10.1146/annurev.neuro.22.1.567
Edinger, L. (1908). The Relations of Comparative Anatomy to Comparative Psychology. *The Journal of Comparative Neurology and Psychology*, *18*(5), 437–457. http://doi.org/10.1002/cne.920180502
Elliott, K. C. (2013). Douglas on values: From indirect roles to multiple goals. *Studies in History and Philosophy of Science Part A*, *44*(3), 375–383. http://doi.org/10.1016/j.shpsa.2013.06.003
Elliott, K. C. (2017). *A Tapestry of Values. An Introduction to Values in Science*. New York: Oxford University Press.
Elliott, K. C., McCright, A. M., Allen, S., & Dietz, T. (2017). Values in environmental research: Citizens' views of scientists who acknowledge values. *PLoS ONE*, *12*(10), 1–18. http://doi.org/10.1371/journal.pone.0186049
Elliott, K. C., & McKaughan, D. J. (2009). How Values in Scientific Discovery and Pursuit Alter Theory Appraisal. *Philosophy of Science*, *76*(5), 598–611. http://doi.org/10.1086/605807
Elliott, K. C., & Willmes, D. (2014). Cognitive Attitudes and Values in Science. *Philosophy of Science*, *80*(5), 807–817. http://doi.org/10.1086/673719
Fanelli, D. (2009). How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PLoS ONE*, *4*(5). http://doi.org/10.1371/journal.pone.0005738
Fanelli, D. (2010). "Positive" results increase down the hierarchy of the sciences. *PLoS ONE*, *5*(4). http://doi.org/10.1371/journal.pone.0010068

Favareau, D. (2008). The Biosemiotic Turn. *Biosemiotics*, *1*(1), 5–23. http://doi.org/10.1007/s12304-008-9010-8
Feyerabend, P. (1978). *Science in a Free Society*. London: NLB.
Fiedler, K. (2011). Voodoo Correlations Are Everywhere - Not Only in Neuroscience. *Perspectives on Psychological Science*, *6*(2), 163–171. http://doi.org/10.1177/1745691611400237
Fiedler, K., & Schwarz, N. (2016). Questionable Research Practices Revisited. *Social Psychological and Personality Science*, *7*(1), 45–52. http://doi.org/10.1177/1948550615612150
Fitch, T. W. (2010). *The Evolution of Language*. Cambridge: Cambridge University Press.
Fitch, T. W. (2017a). Empirical approaches to the study of language evolution. *Psychonomic Bulletin & Review*. http://doi.org/10.3758/s13423-017-1236-5
Fitch, T. W. (2017b). On externalization and cognitive continuity in language evolution. *Mind & Language*, *32*(5), 597–606. http://doi.org/10.1111/mila.12162
Fitch, T. W., Huber, L., & Bugnyar, T. (2010). Social Cognition and the Evolution of Language: Constructing Cognitive Phylogenies. *Neuron*, *65*(6), 795–814. http://doi.org/10.1016/j.neuron.2010.03.011
Frederick, R. (2015). The search for what sets humans apart. *Proceedings of the National Academy of Sciences*, *112*(2), 299–301. http://doi.org/10.1073/pnas.1423320112
Frey, U. J. (2010). Modern Illusions of Humankind. In U. J. Frey, C. Störmer, & K. P. Willführ (Eds.), *Homo Novus – A Human Without Illusions* (pp. 263–288). Berlin, Heidelberg: Springer Verlag. http://doi.org/10.1007/978-3-642-12142-5
Gardner, A., & Gardner, B. T. (1969). Teaching sign language to a chimpanzee. *Science*, *165*(3894), 664–72.
Ghiselin, M. T. (2005). The Darwinian revolution as viewed by a philosophical biologist. *Journal of the History of Biology*, *38*(1), 123–136. http://doi.org/10.1007/s10739-004-6513-2
Gill, S. A., & Bierema, A. M. K. (2013). On the Meaning of Alarm Calls: A Review of Functional Reference in Avian Alarm Calling. *Ethology*, *119*(6), 449–461. http://doi.org/10.1111/eth.12097
Gontier, N., & Pina, M. (2014). The Evolution of Social Communication in Primates. In M. Pina & N. Gontier (Eds.), *The Evolution of Social Communication in Primates* (pp. 1–28). Heidelberg: Springer Cham. http://doi.org/10.1007/978-3-319-02669-5_1
Goodall, J. (1964). Tool-Using and Aimed Throwing in a Community of Free-Living Chimpanzees. *Nature*, *201*(4926), 1264–1266.
Güntürkün, O., & Bugnyar, T. (2016). Cognition without Cortex. *Trends in Cognitive Sciences*, *20*(4), 291–303. http://doi.org/10.1016/j.tics.2016.02.001
Güntürkün, O., Ströckens, F., Scarf, D., & Colombo, M. (2017). Apes, feathered apes, and pigeons: differences and similarities. *Current Opinion in Behavioral Sciences*, *16*, 35–40. http://doi.org/10.1016/j.cobeha.2017.03.003
Hauser. (1997). *The Evolution of Communication* (4th print.). Cambridge, Massachusetts: A Bradford Book.
Hauser, M. D., Chomsky, N., & Fitch, T. W. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science (New York, N.Y.)*, *298*(5598), 1569–79. http://doi.org/10.1126/science.298.5598.1569
Hauser, M. D., Yang, C., Berwick, R. C., Tattersall, I., Ryan, M. J., Watumull, J., … Lewontin, R. C. (2014). The mystery of language evolution. *Frontiers in Psychology*, *5*(MAY), 1–12. http://doi.org/10.3389/fpsyg.2014.00401
Head, M. L., Holman, L., Lanfear, R., Kahn, A. T., & Jennions, M. D. (2015). The Extent and Consequences of P-Hacking in Science. *PLoS Biology*, *13*(3), 1–15. http://doi.org/10.1371/journal.pbio.1002106
Hewes, G. (1973). Primate communication and the gestural origin of language. *Current Anthropology*, *14*(1), 5–24.
Hill, J. H. (1972). On the evolutionary foundations of language. *American Anthropologist*, *74*(3), 308–317.
Hobaiter, C., & Byrne, R. W. (2011). The gestural repertoire of the wild chimpanzee. *Animal Cognition*, *14*(5), 745–767. http://doi.org/10.1007/s10071-011-0409-2
Hockett, C. F. (1960). The Origin of Speech. *Scientific American*, *203*, 88–111.
Hockett, C. F., & Altmann, S. A. (1968). A note on design features. In T. A. Sebeok (Ed.), *Animal*

communication. Techniques of Study and Results of Research. (pp. 61–72). Bloomington, London: Indiana University Press.
Hodos, W., & Campbell, C. G. B. (1969). Scala Naturae: Why There Is No Theory In Comparative Psychology. *Psychological Review*, *76*(4), 337–350.
Hoffmann, R. (2014). The tensions of scientific storytelling. *American Scientist*, *102*(4), 250–253. http://doi.org/10.1511/2014.109.250
Hudson, R. (2016). Why We Should Not Reject the Value-Free Ideal of Science. *Perspectives on Science*, *24*(2), 167–191. http://doi.org/10.1162/POSC_a_00199
Hurford, J. R. (2014). *Origins of Language: A Slim Guide*. Oxford: Oxford University Press.
Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, *2*(8), 0696–0701. http://doi.org/10.1371/journal.pmed.0020124
Ioannidis, J. P. A. (2012). Why Science Is Not Necessarily Self-Correcting. *Perspectives on Psychological Science*, *7*(6), 645–654. http://doi.org/10.1177/1745691612464056
Ioannidis, J. P. A., Boyack, K. W., & Klavans, R. (2014). Estimates of the continuously publishing core in the scientific workforce. *PLoS ONE*, *9*(7). http://doi.org/10.1371/journal.pone.0101698
Ioannidis, J. P. A., Munafò, M. R., Fusar-Poli, P., Nosek, B. A., & David, S. P. (2014). Publication and other reporting biasesin cognitive sciences: detection,prevalence, and prevention. *Trends in Cognitive Sciences*, *18*(5), 235–241. http://doi.org/10.1016/j.tics.2014.02.010
Jackendoff, R. (2010). Your theory of language evolution depends on your theory of language. In R. K. Larson, V. Dèprez, & H. Yamakido (Eds.), *The Evolution of Human Languages: Biolinguistic Perspectives* (pp. 63–72). Cambridge: Cambridge Univ Press.
Jackendoff, R. (2011). What is the human language faculty?: Two views. *Language*, *87*(3), 586–624. http://doi.org/10.1353/lan.2011.0063
Jackson, J. H. (1884). The Croonian Lectures On Evolution And Dissolution Of The Nervous System. *The British Medical Journal*, *1*(1213), 591–593.
John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the Prevalence of Questionable Research Practices With Incentives for Truth Telling. *Psychological Science*, *23*(5), 524–532. http://doi.org/10.1177/0956797611430953
Kant, I. (2003). *Metaphysische Anfangsgründe der Naturwissenschaft [1786]*. (W. Weischedel, Ed.). Frankfurt am Main: Suhrkamp.
Kellogg, W. N. (1968). Communication and Language in the Home-Raised Chimpanzee. *Science*, *162*(3852), 423–427. http://doi.org/10.1126/science.162.3852.423
Kerr, N. (1998). HARKing: Hypnothesizing After the Results are Known. *Personality and Social Psychology Review*, *2*(3), 196–217. http://doi.org/10.1207/s15327957pspr0203_4
Knetsch, G. (2017, July). Faked Science – Ein Feature über Manipulation und Betrug in der Wissenschaft. Germany: WDR 5. Retrieved from https://www1.wdr.de/radio/wdr5/sendungen/dok5/manipulation-wissenschaft-102.pdf
Kousta, S., Ferguson, C., & Ganley, E. (2016). Meta-Research: Broadening the Scope of PLOS Biology. *PLoS Biology*, *14*(1), 1–2. http://doi.org/10.1371/journal.pbio.1002334
Kunda, Z. (1990). The Case for Motivated Reasoning. *Psychological Bulletin*, *108*(3), 480–498. http://doi.org/10.1037/0033-2909.108.3.480
Lehrer, J. (2010, December). The Truth Wears Off: Is there something wrong with the scientific method? *The New Yorker*, 52–57. Retrieved from https://www.newyorker.com/magazine/2010/12/13/the-truth-wears-off
Levelt, W. (2013). *A History of Psycholinguistics: The Pre-Chomskyan Era*. Oxford: Oxford University Press. http://doi.org/10.1093/acprof:oso/9780199653669.001.0001
Longino, H. E. (1990). *Science as Social Knowledge: Value and Objectivity in Scientific Inquiry*. Princeton, New Jersey: Princeton University Press.
Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in Psychology Research: How Often Do They Really Occur? *Perspectives on Psychological Science*, *7*(6), 537–542. http://doi.org/10.1177/1745691612460688
Martins, M. D. J. D., Muršič, Z., Oh, J., & Fitch, W. T. (2015). Representing visual recursion does not require verbal or motor resources. *Cognitive Psychology*, *77*(June), 20–41. http://doi.org/10.1016/j.cogpsych.2015.01.004
Martinson, B. C., Anderson, M. S., & de Vries, R. (2005). Scientists behaving badly. *Nature*,

435(7043), 737–8. http://doi.org/10.1038/435737a
Maynard-Smith, J., & Harper, D. (2003). Animal Signals. Oxford: Oxford University Press.
McGrew, W. C. (2013). Is primate tool use special? Chimpanzee and New Caledonian crow compared. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1630), 20120422–20120422. http://doi.org/10.1098/rstb.2012.0422
McKaughan, D. J., & Elliott, K. C. (2015). Introduction: Cognitive attitudes and values in science. *Studies in History and Philosophy of Science Part A*, *53*, 57–61. http://doi.org/10.1016/j.shpsa.2015.05.011
Modrak, D. K. W. (2001). *Aristotle's Theory of Language and Meaning*. Cambridge: Cambridge University Press.
Mogie, M. (2000). Historical baggage in biology: The case of "higher" and "lower" species. *BioEssays*, *22*(9), 868–869. http://doi.org/10.1002/1521-1878(200009)22:9<868::AID-BIES13>3.0.CO;2-A
Müller, F. M. (1885). *Lectures on the science of language*. London: Longmans, Roberts, and Green.
Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie Du Sert, N., … Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(1). http://doi.org/10.1038/s41562-016-0021
Murat, A. (2010). The Language of Thought Hypothesis. In *The Stanford Encyclopedia of Philosophy* (Fall 2010). Edwart N. Zalta (ed.). Retrieved from http://plato.stanford.edu/archives/fall2010/entries/language-thought/
Naguib, M. (2006). Animal Communication: An Overview. In *The encyclopedia of language & linguistics - Second edition* (Keith Brow, Vol. 1, pp. 276–280). Elsevier.
National Science Board. (2018). *Science & Engineering Indicators*. Retrieved from https://www.nsf.gov/statistics/2018/nsb20181/assets/nsb20181.pdf
Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., … Yarkoni, T. (2015). Promoting an open research culture. *Science*, *348*(6242), 1422–1425. http://doi.org/10.1126/science.aab2374
Nosek, B. A., & Bar-Anan, Y. (2012). Scientific Utopia: I. Opening Scientific Communication. *Psychological Inquiry*, *23*(3), 217–243. http://doi.org/10.1080/1047840X.2012.692215
Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific Utopia: II. Restructuring Incentives and Practices to Promote Truth Over Publishability. *Perspectives on Psychological Science*, *7*(6), 615–631. http://doi.org/10.1177/1745691612459058
Nottebohm, F. (1972). The origins of vocal learning. *American Naturalist*, *106*(947), 116–140.
Nuijten, M. B., Hartgerink, C. H. J., Assen, M. A. L. M. van, Epskamp, S., & Wicherts, J. M. (2015). The prevalence of statistical reporting errors in psychology (1985–2013). *Behavior Research Methods*, (2011), 1–22. http://doi.org/10.3758/s13428-015-0664-2
O'Boyle, E. H., Banks, G. C., & Gonzalez-Mulé, E. (2017). The Chrysalis Effect: How Ugly Initial Results Metamorphosize Into Beautiful Articles. *Journal of Management*, *43*(2), 376–399. http://doi.org/10.1177/0149206314527133
Oakley, K. (1956). The Earliest Tool-makers. *Antiquity*, *30*(117), 4–8. http://doi.org/10.1017/S0003598X00026351
Odling-Smee, L., Giles, J., Fuyuno, I., Cyranoski, D., & Marris, E. (2007). Where are they now? *Nature*, *445*(April), 244–245. http://doi.org/10.1038/445244a
Oesch, N. (2016). Deception as a derived function of language. *Frontiers in Psychology*, *7*(SEP), 1–7. http://doi.org/10.3389/fpsyg.2016.01485
Okanoya, K. (2017). Sexual communication and domestication may give rise to the signal complexity necessary for the emergence of language: An indication from songbird studies. *Psychonomic Bulletin & Review*, *24*(1), 106–110. http://doi.org/10.3758/s13423-016-1165-8
Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716-1-aac4716-6. http://doi.org/10.1126/science.aac4716
Palus, S. (2015). Diederik Stapel now has 58 retractions. Retrieved March 20, 2018, from https://retractionwatch.com/2015/12/08/diederik-stapel-now-has-58-retractions/
Pashler, H., & Harris, C. R. (2012). Is the Replicability Crisis Overblown? Three Arguments Examined. *Perspectives on Psychological Science*, *7*(6), 531–536. http://doi.org/10.1177/1745691612463401

Patel, A. D. (2006). Musical rhythm, linguistic rhythm, and human evolution. *Music Perception*, *24*(1), 99–104. http://doi.org/10.1525/mp.2006.24.1.99

Patel, A. D., Iversen, J. R., Bregman, M. R., & Schulz, I. (2009). Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Current Biology : CB*, *19*(10), 827–30. http://doi.org/10.1016/j.cub.2009.03.038

Pepperberg, I. M. (2018). Tool use in Goffin's cockatoos: Shape/frame matching. *Learning and Behavior*. http://doi.org/10.3758/s13420-018-0317-2

Plos Biology. (2018). The importance of being second. *PLoS Biology*, *16*(1). http://doi.org/10.1371/journal.pbio.2005203

Pluye, P., & Hong, Q. N. (2014). Combining the Power of Stories and the Power of Numbers: Mixed Methods Research and Mixed Studies Reviews. *Annual Review of Public Health*, *35*(1), 29–45. http://doi.org/10.1146/annurev-publhealth-032013-182440

Poincaré, H. (1905). *Science and Hypothesis*. New York: London W. Scott. Retrieved from https://archive.org/details/scienceandhypoth00poinuoft

Premack, D. (1971). Language in chimpanzee? *Science (New York, N.Y.)*, *172*(985), 808–822. http://doi.org/10.1126/science.172.3985.808

Premack, D. (1983). The codes of man and beasts. *Behavioral & Brain Sciences*, *6*, 125–167. http://doi.org/10.1017/S0140525X00015077

Price, M. (2010). Sins against science. *Monitor on Psychology*, *41*(7), 44.

Price, T., Wadewitz, P., Cheney, D., Seyfarth, R., Hammerschmidt, K., & Fischer, J. (2015). Vervets revisited: A quantitative analysis of alarm call structure and context specificity. *Scientific Reports*, *5*, 13220. http://doi.org/10.1038/srep13220

Radick, G. (2007). *The simian tongue: the long debate about animal language*. Chicago, London: The University of Chicago Press.

Reiss, J., & Sprenger, J. (2014). Scientific Objectivity. In *The Stanford Encyclopedia of Philosophy* (Fall 2014). Edward N. Zalta (ed.). Retrieved from http://plato.stanford.edu/entries/scientific-objectivity/

retraction watch. (2018). The Retraction Watch Leaderboard. Retrieved March 20, 2018, from https://retractionwatch.com/the-retraction-watch-leaderboard/

Rigato, E., & Minelli, A. (2013). The great chain of being is still here. *Evolution: Education and Outreach*, *6*(18), 1–6. http://doi.org/10.1186/1936-6434-6-18

Roediger, H. L. I. (2012). Psychology's Woes and a Partial Cure: The Value of Replication. *Observer*, *25*(2), 27–29. Retrieved from https://www.psychologicalscience.org/observer/psychologys-woes-and-a-partial-cure-the-value-of-replication

Rohrmeier, M., Zuidema, W., Wiggins, G. A., & Scharff, C. (2015). Principles of structure building in music, language and animal song. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *370*(1664), 20140097. http://doi.org/10.1098/rstb.2014.0097

Rosati, A. G., Wobber, V. V, Hughes, K., & Santos, L. R. (2014). Comparative developmental psychology: how is human cognitive development unique? *Evolutionary Psychology*, *12*(2), 448–473. http://doi.org/10.1177/147470491401200211

Rosenthal, R. (1979). The "File Drawer Problem" and tolerance for null results. *Psychological Bulletin*, *86*(3), 638–641. http://doi.org/10.1037/0033-2909.86.3.638

Ruse, M. (1996). *Monad to Man. The Concept of Progress in Evolutionary Biology*. Cambridge, Massachusetts, London, England: Harvard University Press.

Samuel Reich, E. (2012). Misconduct ruling is silent on intent. *Nature*, *489*(7415), 189–190. http://doi.org/10.1038/489189a

Scharff, C., & Petri, J. (2011). Evo-devo, deep homology and FoxP2: implications for the evolution of speech and language. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *366*(1574), 2124–40. http://doi.org/10.1098/rstb.2011.0001

Schooler, J. (2014). Metascience could rescue the "replication crisis." *Nature*, *515*(7525), 9. http://doi.org/doi: 10.1038/515009a

Schooler, J. (2016). The Infancy Of Meta-Science. Retrieved February 14, 2018, from https://www.edge.org/response-detail/26751

Schor, S., & Karten, I. (1966). Statistical Evaluation of Medical Journal Manuscripts. *JAMA*, *195*(13), 145–150. http://doi.org/10.1001/jama.1966.03100130097026

Scott-Phillips, T. C. (2014). *Speaking our Minds: Why Human Communication is Different and How Language Evolved to Make it Special*. New York: Palgrave Macmillan.

Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980). Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science*, *210*(4471), 801–803. http://doi.org/10.1126/science.7433999

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366. http://doi.org/10.1177/0956797611417632

Slocombe, K. E., Waller, B. M., & Liebal, K. (2011). The language void: the need for multimodality in primate communication research. *Animal Behaviour*, *81*(5), 919–924. http://doi.org/10.1016/j.anbehav.2011.02.002

Smaldino, P. E., & McElreath, R. (2016). The Natural Selection of Bad Science. *Royal Society Open Science*, *3*(9), 160384. http://doi.org/10.1098/rsos.160384

Smit, H. (2016). The Transition from Animal to Linguistic Communication. *Biological Theory*, *11*(3), 158–172. http://doi.org/10.1007/s13752-016-0246-2

Snowdon, C. T. (2004). Social processes in the evolution of complex cognition and communication. In K. D. Oller & U. Griebel (Eds.), *Evolution of communication systems: a comparative approach* (pp. 131–150). Cambridge: MIT Press.

Steel, D. (2010). Epistemic Values and the Argument from Inductive Risk*. *Philosophy of Science*, *77*(1), 14–34. http://doi.org/10.1086/650206

Steen, R. G., Casadevall, A., & Fang, F. C. (2013). Why Has the Number of Scientific Retractions Increased? *PLoS ONE*, *8*(7), 1–9. http://doi.org/10.1371/journal.pone.0068397

Sterling, T. D. (1959). Publication Descisions and Their Possible Effects on Inferences Drawn from Tests of Significance -- Or Visa Versa. *Journal of the American Statistical Association*, *54*(285), 30–34. http://doi.org/10.2307/2282137

Sterling, T. D., Rosenbaum, W. L., & Weinkam, J. J. (1995). Publication decisions revisited: The effect of the outcome of statistical tests on the decision to publish and vice versa. *American Statistician*, *49*(1), 108–112. http://doi.org/10.1080/00031305.1995.10476125

Suddendorf, T., & Busby, J. (2003). Mental time travel in animals? *Trends in Cognitive Sciences*, *7*(9), 391–396. http://doi.org/10.1016/S1364-6613(03)00187-6

Taglialatela, J. P., Russell, J. L., Pope, S. M., Morton, T., Bogart, S., Reamer, L. A., … Hopkins, W. D. (2015). Multimodal communication in chimpanzees. *American Journal of Primatology*, *77*(April), 1143–1148. http://doi.org/10.1002/ajp.22449

Taylor, T. J. (1997). The Origin of Language: Why It Never Happened. *Language Sciences*, *19*(1), 67–77.

Tenopir, C., & King, D. W. (2014). The growth of journals publishing. In *The Future of the Academic Journal* (pp. 159–178). Woodhead Publishing Limited. http://doi.org/10.1533/9781780634647.159

Terrace, H. S., Petitto, L. A., Sanders, R. J., & Bever, T. G. (1979). Can an ape create a sentence? *Science*, *206*(4421), 891–902. http://doi.org/10.1126/science.504995

Thompson, D. F. (1993). Understanding Financial Conflicts of Interest. *The New England Journal of Medicine*, *329*(8), 573–576. http://doi.org/10.1056/NEJM199308193290812

Townsend, S. W., Koski, S. E., Byrne, R. W., Slocombe, K. E., Bickel, B., Boeckle, M., … Manser, M. B. (2016). Exorcising Grice's ghost: an empirical approach to studying intentional communication in animals. *Biological Reviews*. http://doi.org/10.1111/BRV.12289

Tulving, E. (1983). *Elements of episodic memory*. London, New York: Oxford University Press.

Ullrich, R. (2016). From "speech" to "gesture": The "oral" as norm in "language" research. *Interdisziplinäre Anthropologie Jahrbuch: Wahrnehmung*, *4*, 179–208. http://doi.org/10.1007/978-3-658-14264-3_15

Ullrich, R., & Liebal, K. (2018). Times change, values change: Criteria for attributing language in species comparative research [in print: expected summer 2018]. In E. Luef & M. M. Marin (Eds.), *The talking species: Perspectives on the evolutionary, neuronal, and cultural foundations of language*. Graz: Unipress Graz-Verlag.

Ullrich, R., Mittelbach, M., & Liebal, K. (n.d.). [submitted] Introducing intention : How an idea has spread within cross-species comparative science. *Animal Cognition*.

Ullrich, R., Mittelbach, M., & Liebal, K. (2018). Scala naturae: the impact of historical values on current "evolution of language" discourse. *Journal of Language Evolution*, *3*(1), 1–12. http://doi.org/10.1093/jole/lzx017

Ullrich, R., Norton, P., & Scharff, C. (2016). Waltzing Taeniopygia: Integration of courtship song and dance in the domesticated Australian zebra finch. *Animal Behaviour*, *112*(February 2016), 285–300. http://doi.org/10.1016/j.anbehav.2015.11.012

Vaesen, K. (2012). The cognitive bases of human tool use. *Behavioral and Brain Sciences*, *35*(4), 203–218. http://doi.org/10.1017/S0140525X11001452

Vinkers, C. H., Tijdink, J. K., & Otte, W. M. (2015). Use of positive and negative words in scientific PubMed abstracts between 1974 and 2014: Retrospective analysis. *BMJ*, *351*(December), 1–6. http://doi.org/10.1136/bmj.h6467

Vygotsky, L. (1986). *Thought and Language [1934]*. (A. Kozulin, Ed.). Cambridge: MIT Press.

Wacewicz, S., & Żywiczyński, P. (2015). Language Evolution: Why Hockett's Design Features are a Non-Starter. *Biosemiotics*, *8*(1), 29–46. http://doi.org/10.1007/s12304-014-9203-2

Wagenmakers, E.-J., Wetzels, R., Borsboom, D., von der Maas, H. L. J., & Kievit, R. A. (2012). An Agenda for Purely Confirmatory Research. *Perspectives on Psychological Science*, *7*(6), 632–638. http://doi.org/10.1177/1745691612463078

Weir, A., Chappell, J., & Kacelnik, A. (2002). Shaping of hooks in New Caledonian crows. *Science*, *297*(5583), 981. http://doi.org/10.1126/science.1073433

Wicherts, J. M., Borsboom, D., Kats, J., & Molenaar, D. (2006). The poor availability of psychological research data for reanalysis. *The American Psychologist*, *61*(7), 726–728. http://doi.org/10.1037/0003-066X.61.7.726

Woolard, K. A., & Schiefflein, B. B. (1994). Language ideology. *Annual Review of Anthropology*, *23*, 55–82.

Yerkes, R. (1925). *Almost Human*. London: J. Cape.

Yong, E. (2012). Replication Studies: Bad Copy. *Nature*, *485*(7398), 298–300. http://doi.org/10.1038/485298a

Zuberbühler, K. (2003). Referential signalling in non-human primates: cognitive presursors and limitations for the evolution of language. *Advances in the Study of Behavior*, *33*(2002), 265–307.

Zuberbühler, K. (2013). Primate Communication. In C. Lefebvre, B. Comrie, & H. Cohen (Eds.), *New Perspectives on the Origins of Language* (pp. 187–210). Amsterdam: John Benjamins Publishing Co. http://doi.org/10.1075/slcs.144

Allgemeine Kurzfassung der Ergebnisse

• *Ausgangslage:* Dass die Sprache den Menschen zum Menschen macht, ist eine häufig diskutierte Hypothese. Spätestens seit artvergleichende Forschung die Sprachevolution nachzuzeichnen versucht, ist die Frage nach den Vorstufen, den biologischen Substraten oder den interspezifischen Gemeinsamkeiten eine viele Fachdisziplinen umspannende Variable. Da es keine Einigkeit über eine fachinterne, oder fächerübergreifende Definition von Sprache gibt, existieren vielfältige Konstruktionen dessen was später als solche bezeichnet wird. Diese Konstruktionen sind unter anderem das Ergebnis historischer und gesellschaftlicher Einflüsse.

• *Ansatz:* Die vorliegende Arbeit hat drei der sich daraus ergebenden impliziten Hintergrundannahmen herausgearbeitet und visualisiert. Die Dissertation folgt einem sogenannten mixed-methods Ansatz. Qualitative, historische Recherche liefert die Grundlage für anschließende evidenzbasierte Text- und Zitationsnetzanalysen. Damit gehört die Dissertation zum jungen Forschungsfeld der Meta-Wissenschaften.

• *Ergebnisse:* Drei Hintergrundannahmen wurden untersucht: (i) Sprache ist eine Lautäußerung (Norm des Oralen); (ii) Sie ist die höchste evolutive Errungenschaft (Scala Naturae); (iii) Sprachevolutionsfragen sind nicht über Introspektion zu beantworten (Behaviourismus).

Die Dissertation illustriert die Problematik der jeweiligen Aussagen. Die erste Teilstudie (i) untersucht den Gehörlosendiskurs und analysiert die unbegründete Voreingenommenheit für verbale Lautäußerungen im Vergleich zu nonverbalen Varianten. Die Ergebnisse aus dem Fallbeispiel erklären, warum in der artvergleichenden Forschung zu Beginn des 20. Jahrhunderts kaum Interesse an der Untersuchung von z.B. Gesten bestand. Teilstudie (ii) untersucht die intuitive Annahme, Sprache sei das Ergebnis gerichteter evolutiver Entwicklung. Zwar gilt die Idee seit Veröffentlichung der Evolutionstheorie als überholt, dennoch finden sich bis heute Vorwürfe zu ihrer fortbestehenden Präsenz. Eine quantitative text-mining Analyse untersucht das Vokabular von 915 publizierten Fachartikeln der letzten 10 Jahre. Ein direkter Beleg für Scala Naturae-Denken wurde nicht gefunden, wohl aber Hinweise darauf, dass Forschungsergebnisse von historisch ‚höher' definierten Arten mit positiveren Begriffen beschrieben werden als umgekehrt. Teilstudie (iii) erhellt, wie sich Forschungsparadigmen über die Zeit ändern. Mittels einer Zitationsnetzwerkanalyse wurde die Einführung des Begriffs der *Intention* über einen Zeitraum von 69 Jahren in 653 Fachartikeln untersucht. Die Analyse ermöglicht wichtige Publikationen und Autoren zu identifizieren, die den Korpus besonders prägten. Darüberhinaus konnte der kognitive Umschwung in der artvergleichenden Forschung, sowie ein Primaten-Bias beschrieben werden.

• *Quintessenz:* Wissenschaft war nie, ist nicht und wird wahrscheinlich nie frei sein von gesellschaftlichen und historischen Einflüssen. Ein Problem ist das nicht. Ein Problem ist das Ignorieren und die fehlende systematische Untersuchung der Einflüsse. Die Dissertation als Meta-Analse trägt daher dazu bei einen methodenkritischen Blick auf den aktuellen Diskurs; auf die Auswahl von Spezies; das Studiendesign; die Absicht von Fragen und die Interpretation von Ergebnissen zu entwickeln. Sie ist ein aktiver Beitrag zur wissenschaftlichen Selbstkontrolle.

Erklärung über die eigenständige Verfassung der vorgelegten Dissertation

Hiermit versichere ich, dass die vorgelegte Dissertation gemäß §7 der Promotionsordnung des Fachbereichs Erziehungswissenschaft und Psychologie der Freien Universität Berlin vom 2. Dezember 2008 eine selbstständig durchgeführte und eigenständig verfasste Forschungsleistung darstellt und ich keine anderen als die angegebenen Hilfsmittel und Quellen benutzt habe. Alle Stellen, die wörtlich oder sinngemäß aus anderen Schriften entnommen sind, habe ich als solche kenntlich gemacht. Die Arbeit hat weder in gleicher noch in ähnlicher Form einem anderen Prüfungsausschuss vorgelegen.

Berlin, 09.04.2018

Ort, Datum Robert Ullrich, MSc