In Defense of Idealizations: A Comment on Stokhof & van Lambalgen

Manfred Krifka

The target article of Stokhof and van Lambalgen (2011) discusses what the authors take to be flawed underlying assumptions of modern linguistics, assumptions that may have aided linguistic research but now turn out to hinder its progress and threaten it to become a “failed” discipline.

I think that some of the arguments in this article are themselves flawed, or are based on an understanding of linguistics that is too narrowly focused on certain versions of generative grammar. For example, the argument that in computational applications purely statistical approaches are in general more successful than rule-based approaches has to be qualified: It holds, or may have hold, for certain applications like machine translation, but not for others, like the generation of text to answer queries to databases. Furthermore, statistical methods have been integrated in certain linguistic theories themselves, like stochastic optimality theory (Boersma & Hayes 2001). The authors also claim that linguistics is not able to come up with leading questions for the cognitive and neurophysiological investigation of language processing. This statement is even more puzzling, as it is difficult to name serious research in psycholinguistics or in neurophysiological aspects of language processing that is *not* informed by theoretical notions rooted in linguistics, many of them derived from generative grammar. To cite just one case: Recursion has been proposed – perhaps unjustly so – as the single property that distinguishes human language processing from other animal communication systems; this has led to the identification of special brain regions and pathways of recursive language processing (cf. Friederici 2009).

But not all the points of this article can be dismissed so lightly. Beyond its rhetorics, it challenges one fundamental tenet of generative linguistics, the justification of distinguishing between competence and performance. The argument goes as follows: Linguistics takes on linguistic competence as the object of study, embodied in the notion of an ideal speaker/hearer or the notion of an I-language. In doing this, it dismisses many aspects of actual linguistic performance as uninteresting or intractable. Thus, it ends up with an idealization that tells us little if anything about language, and that is irrelevant for sciences that are concerned with the way how language is embodied in individuals or functions in society. This differs, so the argument goes, from other, successful cases of abstraction in the natural sciences, as e.g. the abstraction from friction in classical mechanics. In such cases, we temporally disregard certain aspects of reality, e.g. by setting certain values in our calculations to zero, in order to be able to get to grips with other aspects of reality or to do any predictive calculations at all. But we recognize these aspects to be of importance for the object of our study, and hope that we can deal with them in the future. The situation in linguistics is different, it is claimed. The kind of abstraction from phenomena of performance that is characteristic for linguistics leads to the hypostatization of the competence of the ideal speaker-hearer. This hypothetical entity interacts with other factors such as memory limitations or communicative goals, which then results in the phenomena that we can observe in linguistic performance.

This is a characterization of a central methodological assumption of generative linguistics that is not quite fair. For example, take the *locus classicus*, Chomsky’s *Aspects* (1965). It develops something like the view described above in its first chapter, *Generative grammars as theories of linguistic competence*, but shows in its second chapter, titled *Theories of linguistic performance,* how the idealized descriptions of grammars may help to describe certain features of performances. For example, it is speculated that direct center-embedding of a category X within a category X is restricted not only by general memory limitations, but by the difficulty to call a procedure to compute a category X while already computing X.

But the characterization of Stokhof and van Lambalgen is not completely unfair either. Chomsky's distinction between competence and performance, or between I-language and E-language, has led to a culture of disinterest in sound empirical evidence beyond grammaticality judgements in some schools of linguistics. It would be easy to cite influential linguistic work that is based on subtle grammaticality judgements by the author alone, often to be contradicted by judgements by other authors. Luckily, there are many different intellectual strands within linguistics, and whether this said culture has indeed dominated linguistic research, and for how long, would itself be an interesting study in the history of our field. It certainly would not be possible to substantiate this claim by counting the number of positions that generative grammarians have in academic departments that are dedicated to the study of language!

Another, related shortcoming in the tradition of generative grammar is that it has encouraged an *a-priori* dismissal of the idea that the inner language faculty is structured in such a way as to accomodate the possible use of language. For example, to counter the plausible idea that the syntax of languages may have developed in such a way as to allow for rapid parsing, Chomsky & Lasnik (1993) state that "In general, it is not the case that language is readily 'usable' or 'designed for use.'" In particular, it is denied that I-language accomodates the use of language as a means of communication, by pointing out that there are apparently non-communicative uses of language, as in soliloquy (which, to my mind, can be seen as derivative uses of communcative language). By following this strategy, generative grammar managed to maneuver itself in a kind of *cul de sac* with regards to the evolutionary explanations of cognitive abilities, as in evolutionary game theory, which have been extremely successful elsewhere. This is one of the reasons why certain versions of it may not be taken as serious by other sciences as they deserve. It is worthwhile to contrast this situation to semantics. There, communicative aspects have been central in the various versions of dynamic interpretation that originated soon after the conception of Montague grammar. Interestingly, in the tradition of generative grammar, communication-related notions like topichood arose early (cf. Chomsky 1970), even if this was not properly addressed in the more general theoretical underpinnings.

In short, even if Stokhof and van Lambalgen overstate their points, they are not completely off the mark. But in this short comment on their article, I would like to turn to the central notion of *idealization,* which they contrast with *abstraction.* The question is: Are idealizations justified, or do they block theoretical progress and practical relevance of our field? I would like to argue, first, that idealizations of the sort performed in linguistics are indeed justified. But I would also like to argue, secondly, that this does not mean that they should be thought of as being embodied in a dedicated module of our cognition, a module that interacts with other modules, resulting in the linguistic behavior that we actually observe. Rather, idealizations can be thought of as high-level computational descriptions that are realized by specific algorithms and embodied in the neural circuitry. This line of thought is not new; it refers to Marr and Poggio's (1976) distinction between different levels of description. I think that Stokhof and van Lambalgen's article would have profited from taking up this distinction, and addressing the value and importance of purely computational descriptions of processes in the brain.

There are many human practices that can be described as idealized systems and for which the issue of whether or not this is justified would never occur. The prime examples are measuring and counting, which have led to one of the best-known formal systems, arithmetics. It is worthwhile to consider this case in detail, as it shows similarities to language.

Research in the past two decades, especially by Stanislas Dehaene and collaborators, has shown that the arithmetical abilities of humans and other species rest on certain elementary abilities (cf. the overviews of Nieder & Dehaene 2009 and Dehaene 2009). First, there is a sense of numerosity, an approximate representation of quantity, that is at least in part independent of the sensory nature of stimuli (like visual, aural, or tactile) and of their symbolic representation (like tally marks, Chinese counting rods, abacus beads, or the Indian-Arab number notation). Like other sensory perceptions, it is logarithmic in nature: it has a more fine-grained representation of smaller numerical quantities than of larger numerical quantities. This is an evolutionarily old sense shared with other mammals and even vertebrates, like birds. Secondly, there is the ability to match elements belonging to two different sets, and in particular to match entities with symbolic expressions. This is the root of counting. Some of the symbolic systems, like tally marks, the number words in certain languages, or specialized graphical systems like the Indian-Arab number notation, can represent very large numbers, even if they fall shy of potentially infinite numbers – the number words even of English are limited,[1] and the amount of available time and paper makes it impossible to write down any given number. True counting and symbolic number representation is linear in its nature, it is evolutionary new and probably represented only in humans, and it has to be formally learned. Thirdly, there is the sense of ordinality, the ability of perceiving ranks, which is not only distinct from the ability to estimate quantities, but also from the ability to count. This shows up in behavioral test: When subjects are asked to judge which of two one-digit numbers is bigger, they are faster when the two numbers are further apart; when asked to judge of two numbers which one follows the other one, they are faster when the numbers are closer together. Little is known about the sense of ordinality, except that it is an evolutionary old sense that is essential, for example, for the establishment of social ranks, as in pecking orders.

We are beginning to understand the neural underpinnings of how these basic arithmetic abilities work. Recent findings point to an area in the intraparietal sulcus that is involved in number estimation and calculation, and has precursors with similar functions in monkeys; in addition, activations associated with grasping, pointing, eye-movement, and language-related activities are invoked. These activations are not necessarily specific for numerosities, but are intervowen with representations of size and location. There are neurons that have firing peaks around particular numerosities; these peaks narrow from infants to adults, allowing for a more precise representation of numerosities. There are also neurons with firing rates that vary monotonically with numerosities in the visual field, which points to two distinct neural codes for numerosities. For very small numbers there is another representation system, subitization, that allows us to identify such numbers without relying on counting or estimation of magnitude. Numbers like one, two and three can be identified before it is "discovered" that two is the result of adding one to one, and three is the result of adding one to two. Numerosities can be added and subtracted too, but these operations typically overshoot, that is, added numerosities are estimated too large, and if one numerosity is subtracted from another, larger one, the result is estimated too small. This has been explained by assuming that these processes are associated with the representation of movement in space.

Symbolic number representation, as with number words or Indian-Arabic numbers, relates to these pre-symbolic representations. Symbolic representation is clearly related to language, which shows up, among other things, in an increasing left-lateralization of neural activiation related to numerosities in children. It generally has the effect of sharpening representations of numerosities, in a way similar to color terms and color recognition. It also leads to a greater "linearization" of the originally logarithmic representations of numerosities: While young children or subjects of cultures with minimal number word lexicon place numbers from 1 to 9 on a line in a way that places 3 in the middle, older children and adults place them with equal distance to each other. It is as if the older system of the representation of numerosities that is subject to Weber's law of logarithmic representations is calibrated by another system of representations. Symbolic representation also integrates subitization, the sense of small numbers, with the sense of numerosity. Young children

---

1 Note that *the greatest number not expressible in English, plus one* is not a number word!

easily learn to apply number words for one, two, and three objects correctly, and it takes some time to generalize this to higher numbers.

It has been shown that the understanding of approximate numerosity and of the distance between different numerosities is essential for the acquisition of symbolic number representation and basic arithmetic in school children. These cultural activities, which include algorithms like addition and subtraction of large numbers, multiplication and long division, the observation of bracketing rules and laws like associativity and distributivity, are quite different in their nature, and the way how they work is still little understood. What we understand even less is the activity of mathematicians. How is an insight like a theorem of arithmetics related to the representation of numerosities in pigeons, macaque monkeys, and young infants? It seems clear that the acquisition of numbers and operations on numbers, in particular addition and its generalized form, multiplication, is necessary to grasp the realisation of the theorem of distributivity, $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$. But we can imagine different ways in which intelligent beings could come up with this law, without relying on more basic senses like the approximate sense of numerosity, and the sense of subitizing small numbers. So this theorem is independent of the more basic senses that have to do with number and quantity.

I would like to turn here to the activity of *meta*-mathematicians that lay the formal grounds of arithmetics. What is the relation of systems such as Peano's axioms and inference rules that are the formal basis of arithmetics to the counting and calculating behavior of humans that we can actually observe? It is not an abstraction in the usual sense. As an example of an abstraction, take the following: When people estimate the number of objects or count them, the nature and spatial distribution of the objects may influence their behavior, and for specific purposes we may want to abstract away from this influence by the quality of objects, and perhaps represent them by dots with equal distances from each other, just as in mechanics we represent three-dimensional objects by their pointlike centers of gravity. Peano's axioms make this kind of abstraction towards pure quantities, as they just talk about one kind of objects, natural numbers. But they clearly go beyond that. They were not developed in order to provide a cognitive or psychological model for how people actually perform calculations. They rather provide a formal model of arithmetic operations that sometimes can be immediately tested for its correctness (such as when we derive the sum of the numbers 2 and 3) and that can be trusted in case of arithmetic operations that go beyond limitations of memory and mental stamina (such as when we derive the sum of 2849 and 4085). Furthermore, they form the basis of rules that can be used to judge the right solution in case two persons come up with different results.

The axiom system of arithmetics, as well as the algorithms developed to perform arithmetic operations, are clear cases of idealizations. Their objects are of a different ontological nature, namely abstract entities – the natural numbers. The rules that govern them, like the successor function, are different from the rules governing numerosities. They are also different from number word systems, which never allow for completely productive successor-based number words like *one and one and one...* (cf. Hammarström 2010), and often are restricted in the sense that they cannot represent a potential infinity of numbers, but have maximal numbers that they can name. In a sense, the axiom system of arithmetics is an *extrapolation* of the simple arithmetical operations that we can perform in our heads or with the assistance of our fingers. It reduces the number of notions and operations to a few basic ones (in Peano-style systems, the concept of a first number, 0, and the notion of a successor), and applies these notions and operations in unlimited ways. For a theory of how this extrapolation might have developed, cf. Lakoff & Núñez (2000), who assume that mathematical notions spring from the application of methaphorical thought, which themselves are rooted in an embodied mind.

The question is: Would Stokhof and van Lambalgen argue against arithmetics because it is not an abstraction from our cognitive abilities to deal with numbers, but an idealization? Presumably not.

It is of great interest to study the nature of arithmetics in its own right, regardless whether we conceive of the objects of arithmetics as Platonic entities or abstract objects that emerge from the repeated application of the successor function. Pure arithmetics is of great interest for intrinsic reasons, and for numerous practical applications in society and engineering. But it is also of interest when we want to study the way how humans perform arithmetical computations. Without this idealization, we would not even be able to understand the difference between the sense of numerosity, which is logarithmic in its nature, and the linear nature of counting. Also, we would not understand that the symbolic representation of numbers does influence the way how we deal with them – for example, why the addition task 24 + 5 is harder than the addition task 26 + 5 in a decimal (but not in a vigesimal) system of representation. For all this, formal arithmetics is crucial.

Now, one can argue that formal grammatical descriptions relate to linguistic performance of humans just as formal arithmetics relates to the ways humans deal with numerosities, counting systems, and subitization. That is, our linguistic performance reveals patterns that can be captured by an idealized formal system in just the same way as our dealing with numbers and quantities can be captured by an idealized formal system. This understanding of grammar is evident in the writings of Richard Montague, who considered the syntax and semantics of natural languages as "branches of mathematics" (see the introduction to Montague 1974 by Richmond Thomason; cf. also the discussion in Partee 1980). But there is an important difference between the notion of idealization as the emerging generalized patterns behind the way how humans behave and how their brains work, and the way how generative grammar is understood by many of its practicioners. For Chomsky, the study of language is not a branch of mathematics, but of psychology. That is, the rules of generative grammar are meant to be embodied, in one way or other, in the human brain. In the analogy between our linguistic and arithmetic abilities that I have tried to draw, this is as if we would claim that somewhere in our brain, there is a module representing Peano's axioms and the definitions of addition, subtraction, multiplication etc., from which the rules of arithmetics follow. We would assume a recursive rule of addition in the mind, stating that a + 0 = a and a + succ(b) = succ(a + b), as this is the simplest way addition can be defined.

There is a certain tension in the attempt to formulate rules of the highest generality on the one hand, and assuming their cognitive reality on the other. This is particularly obvious in recent minimalist work, which favors rationalist arguments and tries to identify features of language that have to be assumed by "virtual conceptual necessity," yet at the same time takes the resulting theory as cognitively real. By the same reasoning, we could make an argument for Peano's axioms, which certainly can be assumed by virtual conceptual necessity, to underly the human cognition of natural numbers. Also, it is not necessary, for the explanation of the structure of arithmetics, to assume a separate sense of numerosity, and a separate faculty of subitization of small numbers. A purely rationalist view would stipulate that they could not exist because they need not exist – regardless what our empirical findings may tell us.

In research on vision, David Marr and Tomaso Poggio (1976) argued for a distinction between four levels of description of cognitive abilities (cf. also Marr 1982, who incidentally used the example of arithmetics and a cash register to illustrate the idea): The level of the basic components, like the single neuron (call it level 0); the level of how neurons actually work together (call it level 1); the level of the algorithm which describes in a general way how neurons work together (call it level 2; level 1 can be seen as a particular implementation of level 2); and finally the level of computation (level 3), the level at which the computational task can be stated and which the algorithms at level 2 should solve, sometimes imperfectly and not in a general way. Marr and Poggio state: "Each level of description has its place in the eventual understanding of perceptual information processing, and it is important to keep them separate. Too often in attempts to relate psychophysical problems to physiology there is confusion about the level at which a probem arises." They add: "More disturbingly, although the top level is the most neglected, it is also the most important. This is because the structure of the computations that underly perception depend more upon the computational problems that have to be solved than on the particular hardware in which their solutions are implemented." Chomsky has dismissed the relevance of Marr and Poggio's levels of representation (cf. Chomsky 1995); the reason given is that Marr's use in vision research relates to input/output systems, and the language faculty is a differnt kind of system. I do not see why representation levels should only be applicable to computation arising in input/output systems. In particular, one could see level 3 descriptions as idealizations – for example, the Peano axioms for our integrated arithmetic abilities, or the rules postulated by a generative grammar for our linguistic abilities.[2]

But there may be an important difference between level 3 descriptions and idealizations. According to Marr and Poggio, level 3 descriptions formulate the computational problem that the algorithms that are implemented in neural circuitry solve, sometimes imperfectly. Yet the idealizations of our arithmetic as well as of our linguistic abilities go beyond that. They allow for an unlimited set of natural numbers and for unlimited syntactic embedding, for example – things that are not required by the actual tasks that the organism has to perform, which always will be confronted with a limited set of numbers and embeddings. Yet a computational description that would integrate such potential limitations, whose precise values are not even known, would be more complex, and as a matter of fact worse, if it were captured by algorithms that integrate such limitations. This is the justification for stating the computational problem in full generality, and perhaps also, minimalistic simplicity.

Perhaps Marr and Poggio were wrong to assume the priority of the computational description of a problem, for which then algorithms are designed that satisfy the computational description to a satisfying degree and that can be implemented in appropriate hardware. This might be right from an engineering perspective, e.g. when we design a cash register. From an evolutionary perspective, it appears that the algorithms have priority over computational descriptions. Natural or cultural selection exerts pressure on algorithms under which they develop and optimize, and the generalized computational description of what they actually compute only emerges as a result: Those algorithms turn out to be the best that are closest in computing the emergent simple computational description. But why should this be so? For arithmetics, and other mathematical structures, one could make the point that the physical world itself is governed by rules that are expressed in mathematical structures (cf. Penrose 2004, chapter 3: *Kinds of number in the physical world*). The development of algorithms that optimize the computation of such structures would then result in evolutionary advantages, and as there is only one physical world, they would converge on computing identical mathematical structures. In this way, idealizations could indeed exert their influence on algorithms and the organisms in which they are implemented.

It is unclear, however, whether the same reasoning can be invoked for I-language because it does not seem to assist, at first sight, in computing anything in the physical world, as arithmetics and its derivatives, like calculus and complex analysis, do. If this is so, and if it is still possible to give a simple, idealized, computational-level description of language, then one could resort to the thesis of I-language as an evolutionary by-product – an exaptation or "spandrel," to use the terms of Stephen Jay Gould – which has been suggested by Chomsky at various places (e.g., Chomsky 1988). Chomsky has proposed that the "number faculty" of humans, the ability to deal with "discrete infinity", is another such by-product, and one that rests on the faculty of language. Conversely, we can speculate that the language faculty is a by-product of the sense of numerosity that is developing to a sense of number, and a capacity to count, which converges on the mathematical structure of arithmetics because this reflects features of the physical world. A problem with that account, however, is that there are known language communities that did not develop number concepts beyond numerosities and subitization, yet seem to have a fully articulated language (the hedge in this sentence is due to the ongoing discussion of languages like Pirahã, which clearly do not have

2 After finishing this reply, I was made aware of Kolb (1997), who points out different conceptual levels in Chomsky's argumentation, which he likens to a logical calculus (the idealization) and an implemented theorem prover (an algorithm that can be implemented).

number words but for which it is debated whether or not they are fully recursive; see Everett 2009 and Nevins e.a. 2009).

But perhaps language represents certain aspects of the physical world just as numbers do, and perhaps it also represents certain essential properties of communication. Then algorithms that determine communicative and representational behavior would converge on allowing idealized grammatical descriptions, just as algorithms dealing with quantity converge on arithmetics. The idealized grammatical descriptions would allow for the classification of entities and phenomena in the natural world: For the combinations of classes by Boolean operations like conjunction, disjunction, and negation; for the restriction of classes by modifiers; for two- or more-place relations between entities and phenomena; for quantifiers as relations between classes; and perhaps also for communication-related notions like the enrichment of common ground, for presuppositions and for implicatures. Idealized grammatical descriptions would be justified because the algorithms of language would be constrained by the essential property of language as a means to represent and communicate, and by the uniformity of what can be represented and what communication is. The consequences of this would be far-reaching. One particular striking one would be that we would expect that we do not only share essential parts of mathematics with extraterrestrial intelligent other life forms, but that our languages would be similar in structure as well.

Boersma, Paul & Brian Hayes. 2001. Empirical testing of the gradual learning algorithm. *Linguistic Inquiry* 32: 45-86.

Chomsky, Noam & Howard Lasnik. 1993. Principles and parameters theory. In: Jacobs Joachim et.al. (eds), *Syntax: An International Handbook of Contemporary Research(1)*. Berlin: de Gruyter, 506-569.

Chomsky, Noam. 1965. *Aspects of the theory of syntax.* Cambridge, Mass.: The MIT Press.

Chomsky, Noam. 1970. Deep structure, surface structure and semantic interpretation. In: Jakobson R. & S. Kawamoto, (eds), *Studies in general and oriental linguistics*. Tokyo: T.E.C. Corporation.

Chomsky, Noam. 1988. *Language and the problems of knowledge. The Managua lectures.* Cambridge, Mass: The MIT Press.

Chomsky Noam. 1995. Language and nature. *Mind* 104: 1-61.

Dehaene, Stanislas. 2009. Origins of mathematical intuitions. *The Year in Cognitive Neuroscience 2009,* 232-259.

Everett, Daniel L. 2009. Piraha culture and grammar: A response to some criticisms. *Language* 85: 404-442.

Friederici, Angela D. 2009. Pathways to language: fiber tracts in the human brain. *TRENDS in cognitive sciences* 13: 175-181.

Hammarström, Harald. 2010. Rarities in numeral systems. In: Wolgemuth Jan & Michael Cysouw, (eds), *Rethinking universals. How rarities affect linguistic systems*. Berlin: Mouton de Gruyter, 11-60.

Kolb, Hans-Peter. 1997. Is language a generative procedure? *Arbeitspapiere des SFB 340*, Bericht 110, 3-11.

Lakoff, George & Rafael E. Núñez. 2000. *Where Mathematics comes from. How the embodied mind brings mathematics into being.* New York: Basic Books.

Marr, David & Tomaso Poggio. 1976. From understanding computation to understanding neural circuitry. *AI Memo 357*. MIT, Artificial Intelligence Laboratory.

Marr, David. 1982. *Vision.* San Francisco: W.H. Freeman.

Montague, Richard. 1974. *Formal philosophy. Selected papers of Richard Montague.* New Haven/London: Yale University Press.

Nevins, Andrew, David Pesetsky & Cilene Rodrigues. 2009. Piraha exceptionality: a reassessment. *Language* 85: 355-404.

Nieder, Andreas & Stanislas Dehaene. 2009. Representation of number in the brain. *Annual Review of Neuroscience* 32: 185-208.

Partee, Barbara H. 1980. Semantics -- mathematics or psychology? In: Bäuerle Rainer, Urs Egli & Arnim von Stechow (eds), *Semantics from different points of view*. Berlin: Springer, 1-14.

Penrose, Roger. 2004. *The road to reality. A complete guide to the laws of the universe.* New York: Vintage Books.

Stokhof, Martin & Michiel van Lambalgen. 2011. Abstraction and idealization. The construction of modern linguistics. *Theoretical Linguistics: An open peer review journal.* # (This volume.)