In Podesva, R. J. & Sharma, D. (Eds.) Research Methods in Linguistics, Cambridge: Cambridge University press, pp. 428-445

Chapter 19: Modelling in the language sciences

Willem Zuidema
Bart de Boer

1. Introduction[1]

Computers can be used for many different purposes in linguistic research. They can be used for data storage and search. They can be used as devices for speech analysis or synthesis. They can be used to present linguistic stimuli to subjects and record their responses. In all these applications, computers are used as sophisticated tools, and they are programmed according to purely practical criteria: as long as it gets the job done, the internal workings of the software are not the subject of the research.

However, computing can also become the focus of linguistic inquiry. Computers can be used to operationalize linguistic theories by implementing them as computer programs. This is done because linguistic theories may be so complex that their predictions can no longer be derived using verbal reasoning or pen-and-paper analysis. Moreover, turning a linguistic theory into a computer program forces the researcher to make her assumptions explicit. By running the program, and studying its behaviour under a variety of circumstances, the researcher can test the theory against empirical findings and often discover unexpected consequences.

In this chapter, we discuss the use of computational models in the language sciences. Although formalization has had a central place since the 1950s in syntax, semantics and phonetics in particular, the last two decades have seen an explosion of interest in mathematical and computational models in almost all linguistic subfields: from typology to language acquisition, from discourse to phonology, linguists are increasingly viewing formal modelling as an approach that ensures the internal consistency of theories (e.g., Steedman 2001; Wang et al. 2005). However, although many proponents of modelling believe it makes their field more scientific and objective, it seems fair to say that the introduction of formal models has so far not led to a broad consensus among language researchers. On the contrary, models have often been at the heart of fundamental controversies (e.g., those about formalisms vs. functionalism, nativism vs. empiricism, single- vs. dual-mechanism accounts of verb morphology; see, e.g., Pinker & Prince 1988).

One reason, we believe, that modelling has played more of a divisive than a unifying role is that there has been little attention to questions about modelling methodology: what kind of lessons can we expect to learn from a model? What makes a good or a bad model? How may different models of the same linguistic phenomenon relate to each other? How could models of different phenomena fit together? Thinking about such questions leads one to systematically consider the role of specific models in a given subfield: Are they consistent with and complementary to each other? Are the assumptions that go into a particular model, if not (yet) supported by empirical findings, made plausible by results from other models?

The situation is not uniform across all linguistic subfields, of course, but we observe that

---

[1] An earlier version of this chapter was distributed among participants of the *Models of Language Evolution: Does the Math Add Up?*-workshop at the International Conference on the Evolution of Language, April 2010, Utrecht.

in fields where 1 or 2 of these questions have received a lot of attention, the others tend to be ignored even more. For instance, in syntactic theory there has been an enormous amount of work (often of impressive mathematical sophistication) on comparing different syntactic frameworks and their ability to model native speaker intuitions about the grammaticality of carefully selected (but often highly contrived) sentences. However, in our view, this field has paid much too little attention to questions about whether that is really the most important criterion for evaluating models of language and about relations with cognitive and neural models. As we will emphasize in this chapter, the ability to reproduce a selected set of empirical phenomena is certainly not the only criterion for a good model.

Because it is impossible to cover all linguistic subfields, we will make our general points about methodology concrete using examples from two particular domains: the evolution of speech and the learnability of syntax. In both fields computational modelling has played an important role, but in both we also believe progress has been hampered by lack of attention to modelling methodology and the questions one immediately asks about the relation between existing models when taking the view on modelling that we develop in this chapter.

For sustaining the success of modelling approaches in linguistic research, it is crucial that models start living up to their promise: modellers must make explicit how their models fit in with other modelling and empirical work, and how their modelling results affect judgments of plausibility of existing hypotheses in the field to which they wish to make a contribution. Moreover, they must do so based on careful consideration of other work, without overstating their results and misusing the prestige that comes with mathematical and computational approaches.

In section 2 we will start with some considerations about the methodology of modelling in linguistics, and introduce the concepts of model sequencing and model parallelization (the latter is described in more detail in Section 4). In sections 3 and 5 we will illustrate these concepts with two case studies on modelling in the evolution of speech and the learnability of syntax respectively. In section 5 we will then draw some general lessons from these case studies, and sketch an agenda for future research in computational modelling of language.

2. Goals of modelling and the model circuitry

From the great many distinctions one can make between different model studies, there are three particularly useful ones that also allow us to establish some common terminology and formulate our view of the field. The first is a distinction based on function, between predictive models and explanatory models (Gilbert & Troitzsch 2005). Predictive models try to model a system as accurately as possible, and to make accurate predictions about the real system's behaviour, as in weather forecasts for example. Predictive models can also be used to reconstruct behaviour in the past, and could for example be used in reconstructing the spread of language families or of particular instances of language change (e.g., Landsbergen 2009). Explanatory models, in contrast, aim to increase insight in a phenomenon. Explanatory models are generally much more abstract and further removed from reality than predictive models. The phenomenon under study is not modelled in all its detail, but instead only its essentials are modelled. Crucially, what counts as 'essential' very much depends on the research question, and simplifications that are appropriate for one question can be totally indefensible for another. Good explanatory models, moreover, explain the phenomenon of interest in terms of more fundamental phenomena that can, at least in principle, be independently motivated (models that simply reproduce the phenomenon of interest without providing such an explanation are sometimes called *phenomenological models*).

The second important distinction is one based on form, between mathematical and computational models. The distinction is not always strict, but mathematical models tend to be the most abstract and to strip down phenomena to their barest essentials. Typically (but not exclusively), mathematical modelling papers provide both a formalization of a phenomenon (e.g., using matrix algebra, logic, differential equations) and proofs about properties of the formal system. Such proofs are, by definition, universally valid and allow inferences about specific cases (deduction), although the simplifications necessary to arrive at a proof often greatly limit the applicability.

Computational models tend to be much more concrete and complex. Phenomena are formalized in a programming language, and the resulting programs studied experimentally. From different runs with different parameter settings, the modeller tries to infer general properties of the formal system (induction). The programs can be very complex, allowing for models with fewer abstractions but often barring analytic proofs. In some cases, computational models are used to investigate versions of a mathematical model that are too complicated to study analytically (including *numerical models*, that are defined algebraically but studied using numerical methods on the computer).

A third major distinction concerns the validation of models: we distinguish between internal validation and external validation (also discussed in Chapter 7). Internal validation is about demonstrating that the phenomenon of interest indeed follows from the stated assumptions, and mathematical proof provides its most powerful form. This is much harder to achieve with computer models, although extensive testing and systematic exploration of the parameter space of a computational model can lead to a great degree of confidence. External validation is about checking whether the stated and unstated assumptions are supported by empirical evidence, or by the outcome of other, independent models, and whether the model's predictions are confirmed in the real world. As computational models are often formulated in more concrete terms, it tends to be easier to achieve external validation.

In the language sciences, we are mainly concerned with the external validation of explanatory models, which in all cases requires an interpretative step: explanatory models have, by definition, abstracted away many details of the phenomenon of interest, making it a matter of judgement whether abstractly formulated assumptions and predictions are supported by concrete evidence. In many fields external validation is further complicated by the fact that there is little direct evidence about which assumptions and predictions are valid, because many of the causal events are unobservable because they happened in a distant past (as in historical and evolutionary linguistics), inside the brain or distributed over millions of language users. External validation is thus only achievable by *model sequencing*: assumptions and prediction of any particular model are validated mainly by results from other models, and only at various points in a string of models do empirical results come into play.

Moreover, because linguistics deals with complicated phenomena for which the appropriate simplifications have not necessarily been established, modelling research should employ *model parallelisation*: for any particular phenomenon, researchers should develop multiple formalisations, compare results and relate observed differences to explicit and implicit assumptions embodied in these alternative models.

Modellers in language research must thus work out relations between different models, whether they stand in sequence or in parallel to each other. This terminology is, of course, based on the metaphor of electronic circuits; we will therefore refer to our perspective on modelling as the 'model circuitry view'. (See Chapter 18 for a discussion of the importance of independent verification in linguistic argumentation more generally.)

3. Model sequencing in practice: A case study on the evolution of speech

To make the ideas about different types of models, and in particular model sequencing, concrete, we will now discuss in some detail the use of models in one particular subfield of linguistics: the evolution of speech. This field is not only one that we have been active in ourselves, but it also offers a particularly good example of a field where modelling can make all the difference because of the paucity of empirical data, but where opportunities have perhaps been missed because of lack of attention to modelling methodology. We will start by briefly discussing some background to this field, and then survey the role of models in answering the key questions of the field.

In the research on how the speech abilities of humans evolved, the focus is usually on the differences between modern humans and the hypothetical latest common ancestor (henceforth, LCA) of humans, chimpanzees and bonobos. Modern humans, as every linguist knows, have a descended larynx, have voluntary control over speech (but much less so over emotional utterances), and have a large learned repertoire of linguistic utterances. Moreover, those utterances have complex internal structure that is used productively, and there are regularities in the repertoires of speech sounds that humans use (phonological universals). The vocal abilities of the LCA are inferred from the abilities that humans, chimpanzees and other apes share or do not share. From such comparisons, it can be derived that the LCA had a repertoire of calls for communicative purposes, and therefore a limited ability to modulate the vocal tract. However, it most likely had a vocal anatomy more comparable to that of chimpanzees and vocal folds comparable to those of chimpanzees and gorillas. The LCA did not, it seems, have modern human's descended larynx, it had less voluntary control over breathing (MacLarnon & Hewitt 1999), and probably did have supralaryngeal air sacs. Finally, it is generally assumed that the LCA, like all modern apes except humans, had only limited voluntary control over vocalizations, learned its vocalizations only to a very limited extent and lacked internal (combinatorial) structure in its calls.

The challenge for research of the evolution of speech is to give an account of how the modern phenotype evolved from the LCA's phenotype: i.e., how did the descended larynx, voluntary control, vocal learning, combinatorial phonology and phonological universals evolve? A key issue here is to what extent the evolutionary changes should be considered adaptations for language, or to what extent they evolved for other reasons. Computer models (and to some extent mathematical models) have been used for a long time to investigate such issues – but in the existing literature (e. g. de Boer 2005; de Boer & Fitch 2010) there are some striking gaps in the range of topics considered and some disturbing confusions about the role of various models. The most studied topics are the evolution of the vocal tract (Lieberman & Crelin 1971; Boë *et al.* 2002; de Boer 2009) and the emergence of phonological universals (de Boer 2000b; Oudeyer 2005; Zuidema & de Boer 2009); the evolution of voluntary control, vocal learning and combinatoriality have received much less attention in the modelling literature, and the issue of how models of these different aspects fit together has been almost completely ignored.

The starting point for many models of how speech evolved is existing models of how speech perception and production work in human adults. Surveying the literature, we quickly find that many models that have been developed for the study of human speech are not necessarily directly usable in the study of the evolution of speech. Illustrative examples from modelling the acoustic production of speech are the 3-parameter model of the vocal tract (Stevens & House 1955; Fant 1960), the coupled mass-spring model of the vocal folds (Dudgeon 1970; Ishizaka & Flanagan 1972), and the source-filter model of speech production (Fant 1960). These are simplified, explanatory models of the human vocal tract, the human vocal folds and the (lack of) interaction between the human vocal folds and the vocal tract, respectively.

These models are well established in phonetics, and provide valuable insights in the process of speech production. However, some researchers in the evolution of speech – erroneously, in our view – reuse these models to represent properties of vocal tracts of our evolutionary ancestors or of other species (see the discussion about Riede *et al.* 2005 in Lieberman 2006). But this is based on a misunderstanding of the *explanatory nature* of the existing models, that involved simplifications which were very helpful for understanding speech production but are specific to human adult vocal tracts. It is, in fact, unlikely that ape-like vocal tracts can make the deformations of the vocal tract that are assumed by the 3-parameter model, and it is clear that the acoustic effects of supralaryngeal air sacs are not captured by it. It is further unknown whether chimpanzee-like vocal folds work in the same way as human vocal folds, and whether in chimpanzee-like vocalizations the vocal folds can really be considered acoustically independent of the vocal tract. Simplifications made in building these models must thus be re-evaluated in the light of what is known about ape and fossil vocal anatomy.

A second problem with existing models of the evolution of speech anatomy concerns its relation to models of the biological and cultural evolution of communication, i.e., its external validation through model sequencing. Even if we could establish a sequence of vocal tracts, leading from ape-like to human-like shapes in gradual steps, that in itself, although an important step, would not provide an evolutionary explanation. As we and others have argued elsewhere (Parker & Maynard Smith 1990; Zuidema & de Boer 2003, 2009), evolutionary explanations must provide a 'path of ever increasing fitness', where every new variant provides a fitness advantage in a population where the previous variant is still common. In the case of vocal tract evolution, it is unclear what the appropriate fitness function is. Existing models tend to assume that it is a simple function of the size of the acoustic space allowed by a particular vocal tract configuration. But fitness due to speech must be a function of how well an individual communicates with others in a population, which in turn depends on the communication system the population uses. However, the relation between the repertoire of speech sounds that emerges in a population and the anatomical and neurocognitive features of individuals is far from trivial.

Models that study the emergence of such repertoires have focused on vowel inventories, and on a role for self-organization in shaping them (Glotin 1995; Berrah & Laboissière 1999; de Boer 2000a; Oudeyer 2005), given constraints on the vowel space formalised by existing models of vowel perception and production. This group of models is a good example of model parallelization: different models making different simplifications modelling the same phenomenon. They are not a good example of model sequencing, however: although these models have yielded a beautiful connection between empirical data on vowel systems and biophysical constraints, it is clear that they only scratch the surface of the full set of phonological universals: they have, for instance, little to say about consonants, syllable-structure or supra-segmental speech patterns.

Ultimately, the connection between phonology and anatomical and neurocognitive features needs to become clear to allow us to evaluate particular scenarios of the evolution of speech. However, despite the progress in modelling vocal tract evolution and vowel universals, we are still quite far from a model-based understanding of the evolution of speech. In the required sequence of explanatory models we still observe, for a variety of reasons, many gaps.

One reason is that, when addressing these more complex issues, the limits of what is at present possible with computer models are reached quickly. It is then tempting to use high-level abstractions (such as distinctive features, constraints and rule-based phonological explanations). However, making use of such abstractions, which have after all been derived for description of modern human language, and are in general not based on direct observation of neurocognitive mechanisms, incurs the risk of implicitly including the phenomena to be explained in the model – and thus resorting to phenomenological rather than explanatory modelling. For example, from typological studies it is known which consonants are unusual (for example uvular plosive [q]) and which are common (for example velar plosive [k]), but there is no language-independent biophysical and neurocognitive model that reliably predicts which articulations are more difficult to produce than others. Thus research into more complex aspects of speech is not only hampered by the computational complexity of such models, but also by our lack of knowledge about the underlying phenomena.

Likewise, we have no models of the evolution of the vocal folds. Although there are many models for human vocal folds (Dudgeon 1970; Ishizaka & Flanagan 1972; Titze 1973, 1974, 2008) and some models of the interaction between the vocal folds and the vocal tract (Flanagan & Meinhart 1964; Titze 2002, 2008) as far as we are aware, no models exist of either chimpanzee vocal folds or of hypothetical ancestral vocal folds. This has undoubtedly to do with the lack of anatomical data (although some has recently been presented Demolin & Delvaux 2006) but also with the fact that vocal folds (and their interaction with the vocal tract) are much more difficult to model than the acoustics of the vocal tract itself.

Another reason is that in spite of much parallel modelling effort, in some domains no consensus is reached. There is, for example strong controversy in the study of the articulatory abilities of Neanderthals and the role of modern human vocal anatomy (with its descended larynx). In this debate, Lieberman (Lieberman & Crelin 1971) and Carré et al. (Carré *et al.* 1995) propose that vocal anatomy has evolved for speech, while Boë et al. (2002) propose that it has not evolved for speech, because (neural) control is more important. They reach opposite conclusions, even though they use very similar modelling techniques. The debate has lead to a rather heated exchange (Boë *et al.* 2007; Lieberman 2007).

Finally, some topics seem to be simply overlooked. For instance, important innovations in the cognitive adaptations for using speech that occurred between the LCA and modern humans have not been addressed by modelling. These include the ability to productively use combinatorial structure of speech and the (related) ability to learn large sets of complex utterances. Such models would be quite complex computationally, but their results might be transferable to other aspects of language, most notably syntax. After all, it has been proposed that the sequential processing and learning that are necessary for using syntax are based on adaptations for the sequential processing and learning mechanisms that are necessary for using combinatorial utterances (Carstairs-McCarthy 1999).

Given these gaps in our understanding of the evolution of speech, the possibilities for external validation are at present limited and we should guard against over-interpreting modelling results. A case in point is the reception of Nowak *et al.* (1999), who presented an information-theoretic model and a mathematical proof of the conditions for combinatorial coding to have a fitness advantage. This proof is an elegant example of internal validation. The model fits into a larger research program in which a number of proofs of mathematical models related to the evolution of language have been presented (Nowak & Krakauer 1999; Nowak *et al.* 2001, 2002). These models have been interpreted by other researchers as having '...demonstrated the evolvability of the most striking features of language...' (Pinker 2000). However this confuses internal validation (the models are internally consistent) with external validation (the models correspond to reality). The latter is unfortunately far from established, given the many simplifying assumptions in Nowak et al.'s (1999) model, as we have pointed out elsewhere (Zuidema & de Boer 2009).

In conclusion, the evolution of speech offers us a good example of a field in which models have played a central role in making progress, but also of a field where it pays to step back a little a consider the relations between all the different models proposed. Such a 'model circuitry' point of view quickly reveals a numbers of important gaps in the existing research and helps both to set an agenda for future research and to put overly optimistic assessments of the state of the art into perspective.

4. Model parallelisation

There are of course infinitely many ways in which models of the same phenomenon can differ. However, we are not talking about small differences between models that are best captured with different settings of one or several parameters. Rather, 'model parallelisation' is about studying models that differ *qualitatively* in the way they approximate reality, i.e., in the simplification that they make. We will briefly discuss two dimensions in which models may differ, one concerning the *ontological status* of language, the other the *linguistic representation* used.

We regard to the ontological status, the issue of what kind of 'thing' a language is, we observe that many models of linguistic phenomena abstract away from individual linguistic cognition and individual differences, and treat a natural language as an independently existing entity. We therefore refer to this approach as 'Platonic'. For many questions in syntax, semantics or phonology the Platonic approach offers a reasonable abstraction. For instance, when providing a formal account of non-constituent coordination in English, as in *Mary wrote and Peter read the book* (Steedman 2001), it makes sense to ignore variation within the English language, differences between production and reception, performance constraints or neural correlates of knowledge of language. It is still useful to develop alternative models for such a phenomenon, and evaluate the different predictions they make (for instance, those on information structure, Steedman 2001), but those alternative models will likely (and reasonably) share the simplifications from the Platonic approach.

For many important linguistic phenomena, however, the Platonic abstraction is not so obviously justified. In accounting for language universals, language change, language acquisition and language evolution, it is crucial that models of several types are studied and compared. A possible conclusion of such a comparison might of course be that particular Platonic models do suffice, but the very fact that possible causal factors (such as heterogeneity in a population, or non-linguistic cognition) get abstracted out in the Platonic approach necessitates investigating these issues.

Fortunately, there are modelling traditions such as 'agent-based modelling' (e.g., Hogeweg 1988; Kirby 2002; Gilbert & Troitzsch 2005) that allow for language users and their interactions to be modelled directly. In such models two or more agents are selected from the population to interact linguistically. Usually in such interactions, one agent is the speaker and another agent is the hearer, but it is also possible that both agents have the role of speaker and hearer during the interaction. Agents generally update their linguistic knowledge in reaction to an interaction. In this way, linguistic knowledge can be transferred from one agent to another and spread in a population. The exact nature of interactions and how the agents react to them depends completely on what the researcher wants to investigate and achieve with the model.

Many different schemes for selecting agents from the population are possible. It is possible that all agents have an equal probability to participate in each interaction (this is called a random mixing population) but it is also possible that certain subgroups of agents have a higher probability of interacting with each other than with other subgroups. This can be due to the modelled spatial location of agents, their social status, their age or any other factor a researcher wishes to model. A scheme that is often used is that the population is divided into two subpopulations: one of teachers and one of learners. Teachers only interact with learners, and neither learners nor teachers interact among themselves. In addition, in such a scheme often the learners are the only ones that update their linguistic knowledge. This is the simplest possible model of transfer of language from one generation to the next. Note that populations do not need to be static: agents can enter (this models immigration or birth) or leave the population (emigration or death), and the agents' behaviour can change over time.

An approach different from both Platonic and agent-based models, is where languages are described at the population level, without modelling details of individual behaviour. One could model, for example the proportion of the population that speaks a variant of the language as a number and then model the way this changes over time using a dynamical system. A dynamical system is a system of mathematical equations that describe changes over time. This can be done with difference equations or with differential equations. Such equations can sometimes be solved analytically, but more often than not, they can only be investigated numerically, with a computer model. Such numerical simulations have been used, for example, to model language change (Wang & Minett 2005). The advantage of such models is that they make use of existing mathematical formalisms and may therefore be easier to read and interpret than computer models. A disadvantage is that it is not always easy to model mathematically what can be modelled straightforwardly with agent-based models.

The second key dimension in which models in linguistics tend to differ is in the linguistic representation used. In the brain of the individual language user, knowledge of language is represented in a complex network of neurons, connections, electrical currents and chemical gradients. Models of language – thankfully – abstract out many of the complexities involved. Many models ignore the inherently continuous and stochastic aspects of the brain, and represent language with discrete, categorical variables and rules. Other models make other simplifications, though, and a true understanding of many phenomena in language again requires comparing these different models.

We identify four classes of representations of language: symbolic models, memory-based models, statistical models and connectionist models; these are schematized in Figure 1. Symbolic models implement linguistic items as abstract, symbolic entities. Typically, no information is represented about how often a certain linguistic object occurs, nor is there a way of representing degrees of acceptability of different linguistic utterances. A linguistic utterance is either possible or not. This makes it relatively easy to analyse and understand the working of these models. Also, these models are usually usable for both production and perception (processing) of language. However, they may have difficulties learning: given that they have a hard time dealing with variation, they tend not to be very robust to noise (speech errors, linguistic variation).

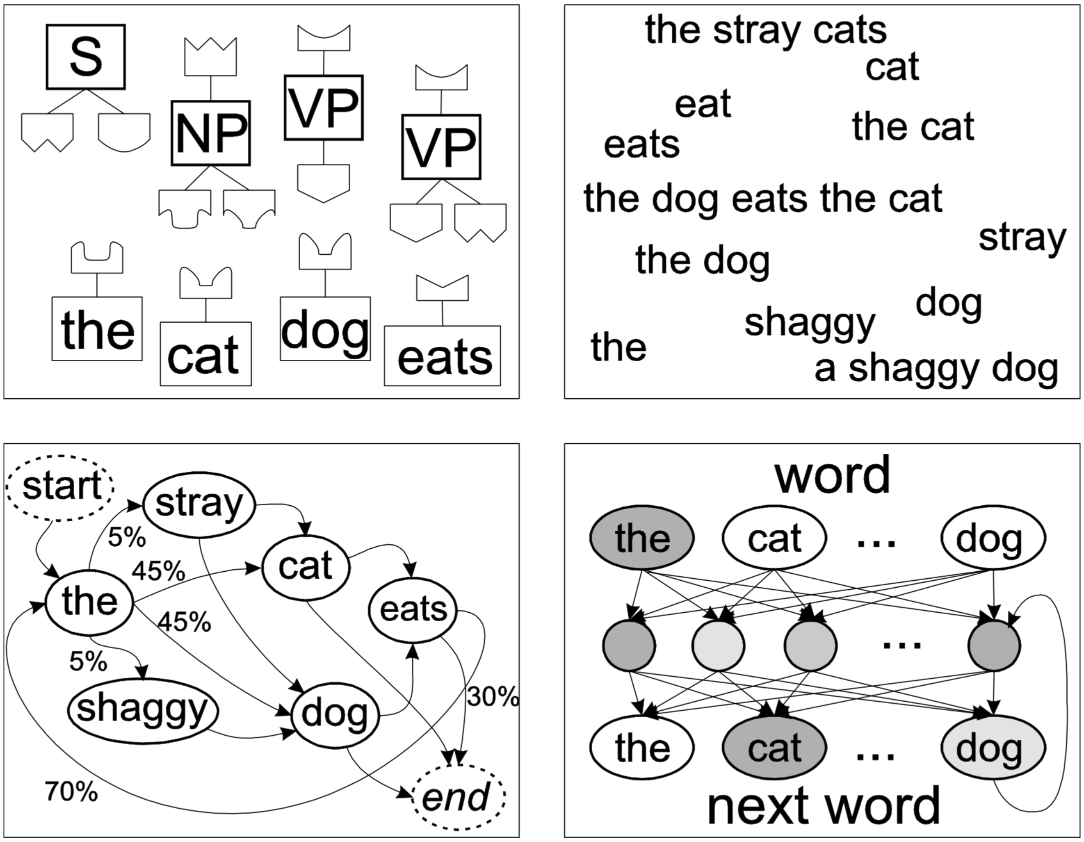

Figure 1: Classes of representation of language

Memory-based models do not generally represent higher level abstractions than that which is observable. Also, they are fundamentally learning systems that can deal with complex and noisy input. The most extreme memory-based system stores all information it observes. By defining a distance function on the items that are stored, the system can retrieve items that are close to a previously unobserved item. In a complete memory-based system, information about meaning, pronunciation and other aspects of the utterance will be stored as well. This allows for generalizations about previously unobserved utterances: forms are expected to have meanings that are close to closely related forms, and meanings are expected to have corresponding forms that are close to closely related meanings.

Memory-based models can be highly successful in modelling human behaviour that involves lots of rote learning, such as acquisition of large lexicons, of irregular stress assignment and of irregular verbs. They are robust to errors in the input, and to predictable variation, such as dialectal variation. With a good distance function they can even generalize well. It is often relatively easy to get an idea of what a memory-based model has learned. However, they have a hard time dealing with the combinatorial nature of human language, without some pre-programmed notion of what the basic elements that are being combined are. For example, it would be difficult for a purely memory based system to figure out how to apply the different morphemes –s in 'the cat**s** bite the dog' versus 'the cat bite**s** the dog'. In order to do this, some notion of words, word classes and morphemes is required.

A third class of models is statistical models. These do not store everything they observe, but store statistical information about how often linguistic items are observed. In one of the earliest (but still often used) instantiations (Shannon 1948), such models represent how likely words are to follow each other. Such models can be trained by counting the co-occurrence of words in large corpora of text (see Chapter 13). Many aspects of human language can be modelled to a reasonable extent by such non-hierarchical statistical models (known as Markov models). They can even deal to some extent with the combinatorial structure of human language. However, they have a hard time dealing with the long-distance dependencies that exist in human languages. This problem can be solved augmenting it with components from symbolic models, such as representations of phrase structure (the result would be a *probabilistic grammar*, an approach that combines the strengths of the symbolic and statistical models we discussed).

The final class of models that we mention are connectionist models. These are also called neural networks, and are inspired by the way the brain is organized. They consist of nodes (modelling neurons) and connections (modelling axons). The nodes each have a level of activation. Connections go from one node to another node and have a weight associated with them. The activation of a node is a function of the sum over the products of the weight of each incoming connection multiplied by the activation of the node from which it originates. Input to the system consists of setting the right activations of the input nodes, and output of the system can be read from the activation of the output nodes. Nodes that are neither input nodes nor output nodes are called hidden nodes. It should be noted that there can be loops in the neural network: connections going ‘back’ in a neural network are called recurrent connections.

Most connectionist models learn. This happens through adaptation of the connection weights based on the input (and possibly the output) that is presented to the network. In the example, the network would be presented with an input word and an output word and its weights would be adapted such that the node representing the output word has higher activation. Connectionist models are robust to noise and variation in the input. In addition, because knowledge is represented in a distributed way – it is distributed over the different connection weights and activations – the network is robust to loss of nodes and connections in a way very similar to the way real brains are robust to damage. This can be an advantage when using computer models to study models of brain damage and aphasia. The distributed representations are a disadvantage, however, when one wants to understand what exactly a connectionist model has learned and how it solves problems. It can be hard or impossible to reduce the distributed representation to a more abstract representation that provides insight about the problem.

On both dimensions – the ontological status and the linguistic representation – there is thus an enormous variation in existing and possible models in the language sciences, and there are often fierce debates about what the 'correct' choices are. We argue that we need to move away from questions about the correct level or correct formalism: there is no single best choice that works for all research questions; rather, we need to compare parallel models and use simplifications that are appropriate for the particular issue we are studying.

5. Model parallelisation in practice: a case study on the learnability of syntax

As a case study on the need for model parallelisation we will now briefly discuss several models relating to language learnability. This field provides a good example of a field where models have played a central role, but also of a field where modelling results have been widely misinterpreted. Careful attention to model parallelisation could, we believe, have avoided these misunderstandings.

The seminal model study in this field is by Mark Gold (1967), who proved that several classes of formal languages are not learnable in a technical sense. Gold defined 'learnability' as a property of a class of language, using the notion of 'identification in the limit'. The learning situation can be imagined as follows: a teacher selects a language L from a given class C of languages, and presents the grammatical sentences from L in an arbitrary order to a learner A. From the very start, the learner tries to guess which language the teacher has in mind. A class C is called *learnable* if there exists an algorithm A that is guaranteed to arrive (and stay) at the correct hypothesis in the limit of an infinite amount of examples. Gold went on to show that some popular classes of formal languages, including finite-state, context-free and context-sensitive languages, are not learnable in this sense. These results have been widely interpreted as providing support for a nativist view on language: if the type grammars we need to describe natural language are not learnable, the argument goes, it is reasonable to conclude that they are not learned but in essence innate.

Now, as is already clear from this informal description, Gold made a number of idealizations of the language learning situation, and it is thanks to these simplifying assumptions that his mathematical proofs were possible at all. One of these idealizations is that there is an infinite amount of data; in a sense, Gold is therefore even too lenient, given that actual language acquisition has to happen – and does happen – within a finite and even relatively short period of time. A number of alternative modelling frameworks, including PAC-learning (Valiant, 1984), have been developed that are applicable to learning situations with time constraints, but these don't fundamentally change the analysis we present here and we will not discuss them.

In other idealizations, Gold appears too strict. In the original versions of his proofs, no reference is made to semantics, pragmatics and phonological information, even though some (and perhaps many) cues from each of these domains are obviously available to the language-learning child. Moreover, Gold's best known results are for situations where learners are presented only with positive evidence, but he obtained different learnability results when negative evidence is also available. These observations have led to quite heated debates with researchers critical of nativism denouncing Gold's theorem (e.g., Elman et al. 1996), but others pointing out that the additional cues can also be modelled as strings and that negative evidence is almost absent from learning child’s linguistic input (Marcus 1993).

Many of the claims in these debates about Gold's results are factually incorrect, as reviewed extensively by Johnson (2004). Johnson also shows that the participants in the debate curiously overlooked a much more essential point: that Gold's definition of learnability as ‘identification in the limit’ is fundamentally unpsychological, because it is a property of *predefined* classes, across all possible learning algorithms and all possible learning environments. In contrast, in real language learning there are strong biological constraints on the possible learning algorithms and environments, and the classes of language are not predefined but rather a consequence of a learning cycle. Concretely, this means that Gold's proofs are perfectly consistent with a situation where a domain-general learning algorithm A is successful at learning many different languages $L_1, L_2, \ldots L_n$ from an unlearnable class C (Zuidema 2003). The fact that class C is *unlearnable* only implies that A cannot be (guaranteed to be) successful at learning all languages in C. Negative learnability results about classes like the class of context-sensitive languages are thus only relevant for people that would claim all context-sensitive languages are possible targets for learning, which is an absurd claim, even for the staunchest, *tabula rasa* empiricist.

Zuidema (2003) presented a simple computational model where a toy grammar induction algorithm is successful at learning some target context-free languages (i.e., a subclass of the context-free languages is learnable by the given algorithm). He showed that only languages from that subclass survive the process of cultural evolution where languages are transmitted from one generation to the next. Hence, subclasses of unlearnable superclasses might be perfectly learnable, and those subclasses will emerge in the cultural evolution of language regardless of whether there are language-specific innate constraints. Together, these observations make clear that Gold's work simply has nothing to say about the nativism-empiricism controversy in linguistics. Moreover, all it took to demonstrate this fact is a simple agent-based model that avoided some of the common idealizations from the mathematical tradition of learnability research.

In short, Gold's theorem has played a crucial role in the debate about learnability and about innate specialization for language. Although many alternative models of learnability have been developed and used in the debate, they typically have adopted the conceptualization of the problem as provided by Gold, including notions of learnability as a property of predefined classes across all possible learners and learning environments. Careful comparison of Gold's model with models developed in a different paradigm (such as the learning paradigm of Solomonoff 1963)—as required by model parallelization—would have clarified the confusion about the relevance of Gold's theorem for cognitive science much sooner, and would have spared the field much unhelpful and bitter controversy.

6. Conclusions

We have presented a number of techniques that can be useful in linguistic modelling, but more importantly, we have tried to illustrate how we think models should fit together and how they should relate to empirical evidence. There are a number of lessons we would like to be drawn from our analysis. First of all, it seems modellers should pay more attention to how their models relate to other models, and how they fit the bigger linguistic picture. Although most papers on linguistic modelling do a good job at internal validation and at crediting other researchers' work, authors do not often make explicit how their models fit more broadly into linguistics outside the detailed issue they study and in what way their model provides external validation for other models or how other models provide it for theirs.

Second, we note that there is no lack of models and no lack of data, but there is a rather uneven distribution of modelling effort over relevant questions. It is perhaps not surprising that (as in other fields of scientific inquiry) the majority of papers are concentrated around the easiest questions. Understandable as this is, we have now reached a stage where we should also attempt to tackle the more difficult questions, and consider carefully whether a collection of models together constitute a convincing explanation.

In order to make progress with computational models, a framework in which different models can be situated and compared with each other, and in which gaps in the modelling effort can be identified, would be useful. In the study of the cognitive processes underlying language, human behaviour presents the point of reference. A problem is that non-modelling linguists have not yet reached consensus about how language works in the brain. However, there is at least a wealth of data that can be used for external validation of computer models. Increasingly, through studies of the workings and the genetics of the brain, data are available about the actual way the brain processes language.

Such data are not always available for the modelling of the history and evolutionary dynamics of language – these are historical processes and information is irretrievably lost. However, papers presenting complete scenarios 'verbally' may be very useful in structuring a research program. Jackendoff (2002) is one of the few authors who provides a rather detailed scenario of evolution that may provide a useful framework if handled with care.

If these challenges are taken up by the field, we should have in a few years several models for each issue *in parallel*, as well as a set of models that *in sequence* really speak to the plausibility of a particular theory. Only then are we approaching *external validation* of *explanatory models* of language, and is the modelling approach really proving its worth to the whole field of linguistics.

References

Berrah, A.-R., & Laboissière, R. (1999). Species: An evolutionary model for the emergence of phonetic structures in an artificial society of speech agents. In D. Floreano, J.-D. Nicoud & F. Mondada (Eds.), *Advances in artificial life, lecture notes in artificial intelligence* (Vol. 1674, pp. 674–678). Berlin: Springer, 1999.

Boë, L.-J., Heim, J.-L., Honda, K., & Maeda, S. (2002). The potential Neandertal vowel space was as large as that of modern humans. *Journal of Phonetics, 30*(3), 465–484.

Boë, L.-J., Heim, J.-L., Honda, K., Maeda, S., Badin, P., & Abry, C. (2007). The vocal tract of newborn humans and Neanderthals: Acoustic capabilities and consequences for the debate on the origin of language. A reply to Lieberman (2007a). *Journal of Phonetics, 35*(4), 564–581.

Carré, R., Lindblom, B., & MacNeilage, P. F. (1995). Rôle de l'acoustique dans l'évolution du conduit vocal humain. *Comptes Rendus de l'Académie des Sciences, Série II, 320*(série IIb), 471–476.

Carstairs-McCarthy, A. (1999). *The origins of complex language: An inquiry in the evolutionary beginnings of sentences, syllables, and truth*. Oxford: Oxford University Press.

de Boer, B. (2000a). Emergence of vowel systems through self-organisation. *AI Communications, 13*, 27–39.

de Boer, B. (2000b). Self organization in vowel systems. *Journal of Phonetics, 28*(4), 441–465.

de Boer, B. (2005). Evolution of speech and its acquisition. *Adaptive Behavior, 13*(4), 281–292.

de Boer, B. (2009). Why women speak better than men (and its significance for evolution). In R. Botha & C. Knight (Eds.), *The prehistory of language* (pp. 255–265). Oxford: Oxford University Press.

de Boer, B., & Fitch, W. T. (2010). Computer models of vocal tract evolution: An overview and critique. *Adaptive Behavior, 18*(1), 36-47.

Demolin, D., & Delvaux, V. (2006). A comparison of the articulatory parameters involved in the production of sounds of bonobos and modern humans. In A. Cangelosi, A. D. M. Smith & K. Smith (Eds.), *The evolution of language: Proceedings of the 6th international conference (evolang6)* (pp. 67–74). New Jersey: World Scientific.

Dudgeon, D. E. (1970). Two-mass model of the vocal cords. *Journal of the Acoustical Society of America, 48*(1A), 118.

Eigen, M., & Schuster, P. (1977). The hypercycle: A principle of natural self-organization

part a: Emergence of the hypercycle. *Die Naturwissenschaften, 64*(11), 541–565.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science, 14*, 179-211.

Elman, Jeffrey L., Elizabeth A. Bates, Mark H. Johnson, Annette Karmiloff-Smith, Domenico Parisi, and Kim Plunkett. (1996), Rethinking Innateness. A Connectionist Perspective on Development. Cambridge, MA: MIT Press.

Fant, G. (1960). *Acoustic theory of speech production*. Gravenhage: Mouton.

Flanagan, J. L., & Meinhart, D. I. S. (1964). Source-system interaction in the vocal tract. *Journal of the Acoustical Society of America, 36*(10), 2001-2002.

Gilbert, N., & Troitzsch, K. G. (2005). *Simulation for the social scientist, second edition*. Maidenhead (UK): Open University Press.

Glotin, H. (1995). *La vie artificielle d'une société de robots parlants: Émergence et changement du code phonétique*. Grenoble: DEA sciences cognitives-Institut National Polytechnique de Grenoble.

Gold, E. M. (1967). Language identification in the limit. *Information and control control (now information and computation), 10*, 447–474.

Goldberg, A. E. (2003). Constructions: A new theoretical approach to language. *Trends in cognitive sciences, 7*(5), 219-224.
Hogeweg, P. (1988), MIRROR beyond MIRROR, puddles of Life. *Artificial Life* (C. Langton, ed.). Addison Wesley Publ. Comp., pp. 297-315.
Ishizaka, K., & Flanagan, J. L. (1972). Synthesis of voiced sounds from a two-mass model of the vocal cords. *The Bell system technical journal, 51*(6), 1233-1268.
Jackendoff, R. (2002). *Foundations of language*. Oxford: Oxford University Press.
Johnson, K. (2004), Gold's Theorem and Cognitive Science, *Philosophy of Science.*
Kirby, S. (2002). Natural language from artificial life. *Artificial Life*, 8(2), 185–215.
Landsbergen, F. (2009). *Cultural evolutionary modeling of patterns in language change: Excercises in evolutionary linguistics*. Utrecht: LOT.
Lieberman, P. H. (2006). Limits on tongue deformation - diana monkey formants and the impossible vocal tract shapes proposed by riede et a. (2005). *Journal of Human Evolution, 50*(2), 219-221.
Lieberman, P. H. (2007). Current views on Neanderthal speech capabilities: A reply to Boë et al. (2002). *Journal of Phonetics, 35*(4), 552–563.
Lieberman, P. H., & Crelin, E. S. (1971). On the speech of Neanderthal man. *Linguistic Inquiry, 2*, 203–222.
MacLarnon, A., & Hewitt, G. P. (1999). The evolution of human speech: The role of enhanced breathing control. *American Journal of Physical Anthropology, 109*(3), 341–343.
Marcus, G.F. (1993), Negative Evidence in Language Acquisition. *Cognition* 46, 1:53–85.
Nowak, M. A., Komarova, N. L., & Niyogi, P. (2001). Evolution of universal grammar. *Science, 291*(5501), 114–118.
Nowak, M. A., Komarova, N. L., & Niyogi, P. (2002). Computational and evolutionary aspects of language. *Nature, 417*(6889), 611-617.
Nowak, M. A., & Krakauer, D. (1999). The evolution of language. *Proceedings of the National Academy of Sciences, 96*, 8028–8033.
Nowak, M. A., Krakauer, D., & Dress, A. (1999). An error limit for the evolution of language. *Proceedings of the Royal Society of London, 266*, 2131–2136.
Oudeyer, P.-Y. (2005). The self-organization of speech sounds. *Journal of Theoretical Biology, 233*(3), 435–449.
Parker, G. A., & Maynard Smith, J. (1990). Optimality theory in evolutionary biology. *Nature, 348*, 27–33.
Pinker, S. (2000). Survival of the clearest. *Nature, 404*, 441–442.
Pinker, Steven, and Alan Prince (1988). On Language and Connectionism: Analysis of a Parallel Distributed Processing Model of Language Acquisition. *Cognition* 28: 73–193.
Quinn, M. (2001). Evolving communication without dedicated communication channels. In J. Kelemen & P. Sosik (Eds.), *Ecal 2001, lecture notes in artificial intelligence 2159* (pp. 357-366). Berlin: Springer-Verlag.
Redford, M. A., Chen, C. C., & Miikkulainen, R. (2001). Constrained emergence of universals and variation in syllable systems. *Language and Speech, 44*, 27–56.
Riede, T., Bronson, E., Hatzikirou, H., & Zuberbühler, K. (2005). Vocal production in a non-human primate: Morphological data and a model. *Journal of Human Evolution, 48*(1), 85-96.
Shannon, Claude E. (1948), A Mathematical Theory of Communication. *The Bell Systems Technical Journal* 27: 379–423 and 623–656.
Steedman, M. (2001), *The syntactic process*. Cambirdge MA: MIT Press.

Stevens, K. N., & House, A. S. (1955). Development of a quantitative description of vowel articulation. *Journal of the Acoustical Society of America, 27*(3), 484–493.

Titze, I. R. (1973). The human vocal cords: A mathematical model part i. *Phonetica, 28*(3), 129–170.

Titze, I. R. (1974). The human vocal cords: A mathematical model part ii. *Phonetica, 29*(1), 1-21.

Titze, I. R. (2002). Regulating glottal airflow in phonation: Application of the maximum power transfer theorem to a low dimensional phonation model. *Journal of the Acoustical Society of America, 111*(1 Pt 1), 367-376.

Titze, I. R. (2008). Nonlinear source–filter coupling in phonation: Theory. *Journal of the Acoustical Society of America, 123*(5), 2733-2749.

Wang, W. S.-Y., & Minett, J. W. (2005). The invasion of language: Emergence, change and death. *Trends in Ecology & Evolution, 20*(5), 263–269.

Zuidema, W. (2003). How the poverty of the stimulus solves the poverty of the stimulus. Proceedings in: Suzanna Becker, Sebastian Thrun, and Klaus Obermayer (eds.), *Advances in Neural Information Processing Systems 15 (Proceedings of NIPS'02)*, MIT Press, Cambridge, MA, 51-58.

Zuidema, W., & de Boer, B. (2003). How did we get from there to here in the evolution of language? *Behavioral and Brain Sciences, 26*(6), 694–695.

Zuidema, W., & de Boer, B. (2009). The evolution of combinatorial phonology. *Journal of Phonetics, 37*(2), 125–144.