A Model of Models*

Stuart Glennan

Department of Philosophy and Religious Studies
Butler University
4600 Sunset Avenue
Indianapolis, IN 46228

sglennan@butler.edu

Version 1.4

August 2000

Abstract

Although many philosophers of science have recognized the importance of modeling in contemporary science, relatively little work has been done in developing a general account of models. The most widely accepted account, put forth by advocates of the semantic conception of theories, misleadingly identifies scientific models with the models of mathematical logic. I present an alternative theory of scientific models in which models are defined by their representational relation to a physical system. I explore in some detail a particular sort of model called a 'mechanical model' I illustrate the applicability of my approach by applying it to a problem in contemporary speech perception research. The model of models is used to analyze how competing models of the mechanisms of vowel normalization are constructed, tested, and revised.

* Thanks to Lindley Darden, Erich Reck, and J.D. Trout for valuable comments on an earlier draft of this paper. Portions of this paper are condensed and revised versions of several chapters of my dissertation (Glennan 1992), and I am grateful to the members of my committee, especially Bill Wimsatt, for their help. Parts of this paper were read at a meeting of the Indiana Philosophical Association, and I thank James Mattingly for his useful commentary. Finally, I would like to thank Howard Nusbaum, who introduced me to research in speech perception and provided me with many insights into modeling techniques in cognitive science.

The “received view of scientific theories” is no longer the received view of scientific theories. The so-called “received view” has been supplanted by the “semantic conception” of theories, according to which theories are really collections of models, where models in turn are understood to be instances of the kind of mathematical structures familiar to philosophers from their study of contemporary logic. In this paper I shall raise some doubts about some of the central tenets of this now not-so-new view. In particular, I shall seek to cast doubt upon Patrick Suppes’ (1960) claim that the meaning of the term ‘model’ as it is used in the sciences is essentially the same as the meaning of that term as it is used in mathematical logic. I shall also raise some objections to the semantic conception’s account of the relationship between theories and models. While these criticisms of the semantic conception are serious, they are not meant to argue in favor the older syntactic conception of theories. Moreover, despite these criticisms, I am persuaded that much that has been said by advocates of the semantic conception about the nature of scientific theories and models is deeply illuminating. I will however suggest that acceptance of these insights does not require us to admit that scientific models and mathematical models are essentially the same kinds of things.

Advocates of the semantic conception have developed a “theory of theories”, but their belief that scientific models are just mathematical models put to another use has led them not to offer a “theory of models”, since the nature of mathematical models is (we’ll suppose) well understood. But, if my objection to Suppes’ thesis is correct, there is indeed a need for a theory of scientific models. There are, however, grave difficulties in developing a true <u>theory</u> of models. As I shall argue in Section II of this paper, the hallmark of theories (as opposed to models) is generality, and thus a theory of scientific models would have to be a set of principles characterizing scientific models in general. It seems likely, however, that the sorts of entities scientists call models are so heterogeneous as to defy any informative unifying description worthy of the name ‘theory’. The best one can hope for is to develop models of particular kinds of scientific models. The majority of this paper is devoted to developing one such model—a model of what I shall call “mechanical models”. While I make no claim that all scientific models are of this form, I shall argue that many scientific models are models of mechanisms. To illustrate what a mechanical model is, I shall describe in some detail two competing models of a phenomenon in speech perception called “vowel normalization”. A case study of these models will illustrate some general principles about how mechanical models are constructed, tested and revised.

I. Scientific Models versus Mathematical Models

There are a number of variants of the semantic conception but most semantic theorists would probably accept Fred Suppe’s formulation:

> Theories are extralinguistic entities which can be described by their linguistic formulations. The propositions in a formulation of a theory thus provide true descriptions of the theory, and so the theory qualifies as a model for each of its formulations. This suggests that the semantic techniques of model theory … will be useful in analyzing the structure of scientific theories. This suggestion gains further plausibility when it is noted that in actual practice the presentation of a scientific theory often takes the form of specifying an intended model….(Suppe 1974, 222)

The semantic conception is so named for two reasons. First it claims generally that theories are semantic rather than syntactic entities. Second, and more particularly, it claims that the apparatus of formal semantics and model theory can illuminate questions about the nature of scientific theories and models.

To simplify exposition, I use the term ‘scientific model’ to refer to entities which practicing scientists refer to as models — recognizing that the class of such entities is quite heterogeneous. Similarly, I use the terms ‘semantic model’ or ‘mathematical model’ to refer to the kind of model used by mathematicians and logicians. Using these terms, the core of the semantic conception can be presented in the following two propositions:

SM: The class of scientific models is a proper subset of the class of semantic models. In particular, the set of scientific models of a scientific theory is just the set of intended models of a formulation of that theory.

ST: A scientific theory is just the class of intended models of one of the (equivalent) formulations of that theory.

The first proposition tells us what scientific models are—semantic models—while the second proposition tells us what scientific theories are—collections of semantic models.

To give SM more content, one must spell out what is meant by a semantic model. Advocates of the semantic conception have offered two alternatives—the set-theoretic approach and the state space approach. The set theoretic approach, developed chiefly by Suppes (1960, 1967), takes scientific models to be quite literally semantic models as they are ordinarily defined in presentations of the semantics of predicate logic. In such presentations, a model is defined as an interpretation of a set of statements of predicate logic under which all members of that set are true. An interpretation in turn is understood to be a function from non-logical symbols of the language onto individuals or sets of individuals of a given domain. If we consider the image of this function, we have a set-theoretic structure that is said to satisfy the set of statements. The relationship that obtains between theories, models and the systems they model is shown in figure 1:

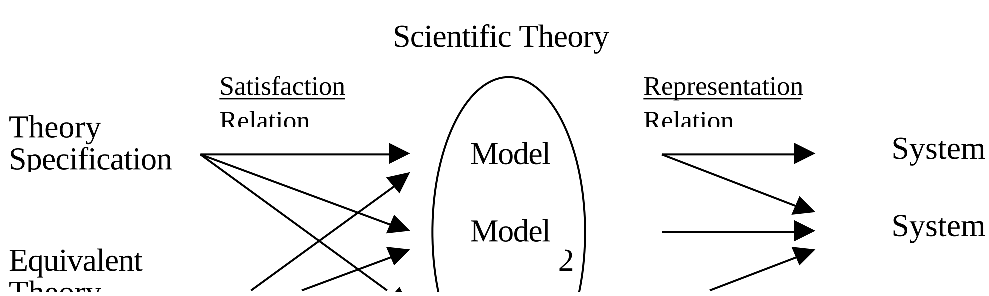

Fig. 1: Models on Theories on the Semantic Conception

A major difficulty with the set-theoretic approach is that it represents models in a way that is, on the surface at least, radically different from the way that models are typically represented by scientists. Models in population genetics, for instance, just aren't represented as set theoretic structures. An advocate of the set-theoretic approach can point to some notable exceptions to this claim. For instance, models of general relativity are naturally understood as set-theoretic structures satisfying Einstein's equation. Moreover, it could be argued that other models that are not typically specified in this way could be so specified. While such specifications may be possible, they would typically take us quite a way from typical scientific practice. A theory of models of this kind would thus fail to satisfy a desideratum of any such theory, that it be closely enough connected to ordinary scientific modeling practice as to make the philosophical analysis intelligible to scientists.

On this score, the second variant of the semantic conception, the state space approach, has clear advantages. According to the state space approach, the state of a physical system can be characterized by a set of state variables—variables measuring the values of various physical magnitudes. The set of logically possible states of the system can be identified with the set of all possible combinations of values of each of the variables, and these combinations in turn are treated as vectors in the state space. The dynamical behavior of a modeled system can be characterized in terms of the trajectory of the system through this vector space over time. Physically possible changes in state of a system may be characterized by laws of succession, while physically possible combinations of values of state variables can be characterized by laws of coexistence (Suppe 1989).

The state space representation of models is powerful and flexible, and it is hard to imagine anything that scientists might call a model that could not be represented in this way. Indeed, this type of representational format is borrowed from physics, and has been used by scientists in other disciplines as well (e.g., Lewontin 1974). This suggests that the state space version of the semantic conception provides the beginning of an adequate theory of scientific models.

One thing that is odd about the state space approach is that it is not obvious in what sense state-space models are semantic—other than the obvious and uninformative sense that models are not individuated syntactically. Scientific models cannot be a subclass of semantic models unless state space models are semantic models. Trajectories through state spaces are entities of a different kind than the set theoretic structures which are models of the traditional Tarskian semantic kind. Of course it can be argued that nothing requires us to use a Tarskian approach in providing a semantics for a formal language. It is indeed possible to develop a semantics for a formal language in such a way that one can construct a definition of truth under an interpretation for statements of this language and in which the interpretations are state space models.[1] But once one grants that there are many different ways to develop formal semantics, the claim that scientific models are semantic models becomes uninformative. Models are semantic in the sense that the same model can have alternative, syntactically distinct formulations, but this fact does not allow us to infer that models really are sets, trajectories through a state space, or any other kind of abstract structure. The relationship between scientific models and state space or other types of semantic models is analogous to the relationship between numbers and their representations as sets. It is possible to characterize scientific models as semantic models just as it is possible to characterize numbers as sets, but from this fact we are not justified in claiming that scientific models really are semantic models any more than we are justified in claiming that numbers really are just sets of a certain kind.

The real difficulty here lies in the attempt to identify models by a satisfaction relationship that obtains between them and linguistic theory specifications. Let us consider an alternative definition that avoids this identification:

SM': Scientific models are structures that (if correct) represent particular physical systems.

Referring to figure 1, note that SM' simply characterizes models in terms of the representational relations they bear to physical systems rather than the satisfaction relations they may bear to linguistic specifications. Thus, there is not real contradiction between SM and SM'. SM' simply makes the representational rather than the satisfaction relation primary. What is represented is what I have called here a "physical system". By this I mean an actual collection of entities in the physical world whose properties (presumably measurable) bear causal/nomological relations to one another. Examples of physical systems include the solar system, a brain cell, or a CD player.[2]

Nothing about SM' prevents one from using either set-theoretic or state space representations of scientific models. One could argue that the state-space approach is a flexible and powerful way to describe scientific models. Similarly, one could argue that in some instances the set-theoretic approach provides a better way to describe a model than does the state space approach. If this is true, it might seem that Suppes was exactly right to say that the difference between scientific and mathematical models is only one of use, not of meaning. But the distinction between meaning and use cannot be so easily made. Once one has severed the connection between the mathematical and scientific uses of a model, there exists no necessary connection between mathematical and scientific models at all. It may well be that a scientific model used to represent a physical system can also stand in a satisfaction relation to a set of statements in a formal language, but it is irrelevant to the success of a scientific model whether or not it can perform this semantic function.

II. Models versus Theories

So far I have been concerned with the adequacy of the semantic conception's identification of scientific and semantic models. Let us now turn to the semantic conception's account of the relationship between models and theories, as expressed in proposition ST. According to ST, the relationship between theories and models is mediated by a theory specification, which is a linguistic representation of a theory. The theory is understood as a semantic entity, namely the class of models that satisfy the theory specification. Thus, for instance, Newtonian gravitational theory is characterized by an axiomatization of Newton's laws of motion and the law of universal gravitation. There is a set of (mathematical) models of this theory. The members of this set differ in the size of their domains (i.e., the number of bodies) and the properties of entities within this domain (masses and velocities), but in each of the models the trajectories of each of the bodies satisfies Newton's laws. To assert that Newtonian mechanics adequately describes some particular physical system is to assert that one of these models is an adequate representation of that system.

There is certainly much to recommend this as a "theory of theories". In particular, the idea that theories <u>define</u> a class of ideal systems which are then held (via a theoretical hypothesis) to represent actual physical systems seems an especially fruitful way to characterize the relationship between theory and reality. I am concerned, however, that the semantic approach has lost track here of a significant distinction between scientific theories and scientific models. On the semantic approach, semantic models form part of the technical apparatus that helps explicate the concept of a scientific theory. But there is no theory of scientific models. At least if one reads Suppe, one seldom sees reference to the word 'model' except in the context of this theory of theories and he seems more or less to equate models with theories (see especially, Suppe 1989, Chapter 8).

What then is the difference between a model and a theory? Scientific usage is notoriously non-uniform on this point, and in some situations the terms 'model' and 'theory' are used interchangeably. Still, if we consider a few examples, we may be able to classify them according to whether they may be called theories or models. The best examples of theories are from physics: Newtonian mechanics, Maxwell's theory of electricity and magnetism and the quantum theory. Some representative examples of models include Bohr's model of the hydrogen atom, the <u>lac operon</u> model of gene regulation, or the Copernican model of planetary motion. Do these examples tell us anything about the difference between models and theories? The obvious distinction between these examples at least is that the theories are quite general, while the models are models of particular systems, or at least, of particular kinds of systems. Another interesting difference is this: Each of the theories mentioned can be said to "have models". In other words, there are models that can be constructed in accordance with the laws of these theories which characterize particular systems that are of the general type of which the theory treats.

These considerations suggest that the crux of the distinction between models and theories is the distinction between general and particular. It is this fact that led me in my definition of models (SM') to say that models are representations of particular systems. This fact also suggests the following alternative characterization of scientific theories:

ST': Scientific theories are collections of laws which (if correct) characterize (ceteris paribus) properties of and relations between entities of some physical type.

The idea that theories are collections of laws is hardly novel. It may seem like a simple return to the received view of theories. But ST' should definitely not be construed as a return to the received view. What was wrong with the received view was not the claim that theories consist of sets of scientific laws, but rather the explication of the relationship between statements of the theory and the phenomena with which the theories were concerned. My assertion that theories are collections of laws is not fundamentally inconsistent with the approach to laws given by advocates of the semantic conception.

A number of philosophers, including Margaret Morrison, Mary Morgan and Nancy Cartwright have lately characterized models as "mediators" between laws or theories (Morgan and Morrison 1999). I would claim that the role of models as mediators can be made more sense of given the distinction between models and theories that I have proposed, than it can given the distinction made in the semantic conception. Of course, as figure 1 shows, on the semantic conception models do indeed mediate. But the mediation in this case is between a model and its' linguistic specification. As I have indicated before, I am in agreement with the semantic conception's basic approach to the relationship that obtains between linguistic specifications of theories and actual physical reality, but this particular sort of mediating role is not the role that Morrison et al have in mind, which is mediation between theories and laws (theories and laws are here construed as semantic rather than linguistic entities) on the one hand and real systems in the physical world on the other.

There are at least two different ways in which models can be said to mediate between laws and physical system. On the one hand, models can specify the idealized conditions under which laws can be appropriately applied to a system, while, on the other hand, models can specify how a combination of general laws or principles combine to apply to a particular case. The idea that models specify idealized systems for which laws hold has been argued for by a number of philosophers, including Beatty (1980), Cartwright (1983, 1997), and Giere (1988), and it is arguably an important contribution of the semantic conception to the understanding of the nature of laws. Less attention has been paid to the second sort of mediation, but it is this second sort of mediation that makes sense of the distinction I propose between theories as collections of laws and models as descriptions of particular systems. This second sort of mediation arises when we try to build a model that has any kind of predictive relationship to a complex natural phenomenon. Consider modern meteorological models. Such models are decidedly not specifications of a set of idealizing conditions which, if they obtained, would entail that a physical system behaved in accordance with some simple and general law. These models instead attempt to take into account the heterogeneous causal factors that influence weather patterns. To build such models, one must know things about everything from sunspot activity to seasonal changes in ocean currents. In building such models, a meteorologist uses bits and pieces of many different theories. Shaffner (1993) argues that this feature, which he calls the "reticulateness" of models, is a pervasive feature of models (and relatedly "theories of the middle range) in the biomedical sciences. Cartwright's work on models in physics (e.g., 1983) suggests that the biomedical sciences are not special in this respect, a point about which Suppe (1989, 275) appears to agree. But if models are constructed by piecing together bits of different theories, then it is _prima facie_ problematic to _identify_ theories with models or classes of models, as is done by ST.

My proposal then is that we understand the relationship between theories and models as a relationship between general and particular. In concluding this section, I would like to consider three objections to this thesis:

A. _The relationship of general to particular is exactly that expressed by the semantic conception._

As was noted above, most advocates of the semantic conception identify a theory with a family of models. If the family of models is specified by a formal theory (i.e., a set of statements closed under logical consequence), then those interpretations that satisfy the theory are the models of the theory. Since there is a one-many relationship between the theory specification (which typically is a set of statements) and its models, one might think that the theory is general while the models are particular. There are, however, two problems with this approach. In the first place, according to the semantic conception, a scientific theory is not to be identified with any particular theory specification. The theory itself _is_ the class of models specified by the theory specification. Hence, the relationship between a theory and its models is not one of general to particular, but of class to member. Second, on the conception of generality relevant to the distinction given by SM' and ST', the generality of theories would be expressed by the fact that a linguistic specification of a scientific theory will typically consist only of general statements not containing constant terms. Similarly, the particularity of models will be expressed by the fact that a linguistic specification for a model will include statements containing constant terms. The particularity of the model thus derives from the presence of constant terms in the model's specification, not from the fact that it is a particular model in the semantic sense.[3]

B. Few if any theories are really general, so the distinction between models and theory cannot be in terms of generality.

The view that scientific theories are characterized by their generality has a long history, and is one part of the classical positivist conception of scientific theories. The scientific theories that most obviously have this character are fundamental physical theories, such as Newton's theory of gravitation or Maxwell's theory of electricity and magnetism. The laws of these theories are characterized by their universal form and by the absence of any reference to particular individuals. Many advocates of the semantic conception have called into question the claim that all theories are general in this sense. In particular, biological theory is thought to have a fundamentally different structure than physical theory (Beatty 1980). The difference is thought to derive from the lack of biological laws (Beatty 1995), which in turn is derived from the historical character of biological theory. Biologists describe the properties of organisms, as well as the evolutionary processes by which organisms have come to have these properties. But the fact that organisms have the properties they do depends entirely upon contingent facts of evolution. Mendel's laws, for instance, are not universal generalizations of unrestricted scope, but ultimately only exception-ridden generalizations about the changes in gene frequencies in populations of organisms that utilize a particular type of historically contingent mechanism for recording and transmitting genetic information. Had evolution not produced the molecular and cytological structures that it did, then Mendel's laws would not apply. And in fact there are actual processes of reproduction (e.g., asexual reproduction) that use entirely different mechanisms. To the extent that there are universals in biology, it is only because there are universal mechanisms. But even if such universals do arise, these universals are ultimately contingent. They are at best universally true of the organisms we actually have on this planet, and their universality derives from the historical particulars of this planet. If there are truly universal laws in biology at all, they are probably not empirical at all, but are rather theorems that characterize the properties of certain mathematical models (Sober 1997).

While Beatty and others are certainly correct about the historical and contingent character of biological theory, this fact does not undermine the distinction between theories and models as it is characterized by SM' and ST'. According to ST', a theory is a set of generalizations that hold universally for some domain of entities. Nothing in this definition excludes domains consisting of entities whose properties are historically contingent. Even though Mendel's laws apply only to a certain set of individuals whose existence is historically contingent, those laws are true counterfactual-supporting generalizations about the behavior of that class. And while this contingency gives a kind of particularity to Mendelian genetic theory, it is not the kind of particularity we find in an actual model. The model is meant to represent a particular population with a particular structure—for instance, a 2 allele 2 locus system.

C. Models can be general as well, so again, the distinction between models and theories cannot be couched in terms of generality.

The putative distinction here offered between models and theories can be attacked from the other direction as well. That is, instead of objecting that theories really depend upon particulars, one might claim that models are in fact general. It is not usually the case that a model represents only to a single system. Models, instead represent a system of a certain type. For instance, if we formulate a model of structure of a neuron, this model will be general because there exist many neurons.

This objection can be met by appealing to a distinction between abstract models and general theories[4] A model is abstract to the extent that it leaves out certain details in its representation of a modeled system. Different models may provide more or less abstract characterizations of a modeled system, but these models are still models of a particular system, for what is asserted in claiming that a model is a model of a modeled system is that the model is, to some degree, similar to that system. This assertion, which Giere (1988) calls the "theoretical hypothesis", connects a particular model to a particular system.

The confusion of thinking that models are general in the sense that theories are general arises because abstraction sometimes makes models generally applicable, in the sense that they can be said to resemble (via theoretical hypotheses) each member of a large class of particulars. But whether a model is generally applicable, or whether increasing its abstractness will increase its applicability is simply a matter of contingent fact. For instance, a model of a neuron has wide application, because, as a contingent matter of fact, there are a large number of organisms which have neurons, and these organisms each have a large number of neurons. Moreover, it is also true, because of contingent facts of evolution, that neurons are a rather specialized type of cell, and that there are many cells of other types. Consequently, a more abstract model that applies to cells generally will have wider scope.

While there is a logical difference between a generalization (or class of generalizations) which are true of a set of individuals in a domain and a particular model which can be applied to a number of individuals via different theoretical hypotheses, it is probably the generality of models in this sense that accounts for the interchangeable use of the terms 'model' and 'theory' among many scientists. It is not uncommon, for instance, to speak either of the Bohr model of the atom or the Bohr theory of the atom. According to the usage I propose here, Bohr gives us a model of the atom, albeit a very abstract one which can be applied to lots of cases.

III. Mechanisms and Mechanical Models

SM' provides us with a very abstract and, consequently, general model of scientific models. It is abstract enough to be compatible with SM, for semantic models can in fact do the job of representing complex systems in the world. But while SM' may have been helpful in characterizing the difference between scientific models and scientific theories, as a model of models its high degree of abstraction makes it uninformative about many aspects of modeling practice. In the remainder of the paper I will try to remedy this defect by looking at a more detailed model of models. The trick in constructing this model (as in constructing any model) is to minimize abstractness while maximizing generality. The type of model that I shall discuss I call a mechanical model. Many, but not all, scientific models are of this kind.

A mechanical model is (not surprisingly) a model of a mechanism, so to present an account of mechanical models, a first step is to offer some analysis of the concept of mechanism. In an earlier paper (Glennan 1996) I have offered the following definition of a mechanism:

(M) A mechanism underlying a behavior is a complex system which produces that behavior by the interaction of a number of parts according to direct causal laws.

While I shall not defend this definition here, a few clarifications are in order. In the first place, mechanisms underlie behaviors. The behavior which the mechanism underlies, or, more simply, the behavior of the mechanism, is what the mechanism does. A heart is a mechanism for pumping blood, a Coke machine is a mechanism for dispensing Cokes in return for coins, and so on. Darden and Craver (forthcoming) emphasize the same point about mechanisms, referring to what the mechanism does as the "phenomenon" that the mechanism produces.

Many mechanisms are, like the heart and the Coke machine, actually designed to behave as they do, but one can equally well characterize the mechanisms underlying behaviors that are not the product of design. For instance, one can ask what is the mechanism responsible for the mass extinction of the dinosaurs at the end of the Cretaceous period. Supposing that the mechanism did involve the impact of a large meteorite on the earth's surface, few would suppose that the meteorite was designed to have this effect; yet, mass extinction can be treated as the behavior of this astronomical/geological mechanism.

The second part of the definition that requires explanation is the notion of 'direct causal laws'. These are counterfactual-supporting generalizations that describe how changes in the state of one part of the mechanism directly produce changes in the state of other parts. For instance, in the Coke machine case, it might be that depressing the button completes an electrical circuit that is another part of the machine. If models are represented in terms of laws describing transitions in a state space, these laws would be the laws of succession. The stipulation that the laws be direct is meant to rule out generalizations that describe more remote effects. For instance, depressing the button on the Coke machine causes a Coke to be dispensed, but only via the complex interaction of a number of parts. Hence, there is no direct law connecting button presses to the dispensing of Cokes.

The notion of law employed here is weaker than many philosophers would like to have. While, logically speaking, such laws have universal form, as in "Whenever the button is pushed, the circuit is closed", the laws refer to particular kinds of entities, and are true of these entities because of contingent facts about how these entities are constituted and interconnected. A major advantage of this construal of 'law' which has already been suggested in the previous section is that it treats as laws generalizations such as Mendel's laws, even though such laws are only locally applicable and are true in virtue of contingent facts. And while laws characterizing interactions between parts of Coke machines may be considerably more homely than Mendel's laws, the differences between them are differences of degree rather than of kind. A second advantage of this conception of law is that it shows how laws can actually be explained by mechanisms. What I have called the behavior of the mechanism can often be described by a counterfactual-supporting generalization. We might describe the behavior of a simple Coke machine in the following way: "Whenever $1 is inserted and the button marked 'Coke' is pressed, a Coke appears in the slot at the bottom of the machine". This statement is, on the account I give here, a law statement, and the fact that the statement is true is explained by facts about how the Coke machine (the mechanism) works. Laws of this kind I call mechanically explicable (cf. Glennan 1996, 1997). As a less homely example, Mendel's first law, which we might state as "Whenever a parent is heterozygous at a locus, the proportion of gametes produced by the parent carrying each allele will be .5", is also mechanically explicable. It follows from facts about the mechanisms used in organisms to produce gametes.

There is thus a two-way relationship between laws and mechanisms. First, reliable behavior of mechanisms depends upon the existence of lawful relations between their parts, and direct causal laws characterize these relations. Second, many laws are mechanically explicable, in the sense that they are just generalizations about the behavior of mechanisms. A single law can both be explained by a mechanism and characterize the interaction between parts of a larger mechanism. For instance, Mendel's first law is explained by the standard mechanisms of gamete formation, while evenly segregated gametes, produced in accordance with Mendel's first law, play a part in larger mechanisms, for instance the mechanism by which deleterious alleles will be (generally) driven to extinction.

Given this understanding of what a mechanism is, we now can define a mechanical model:

(MM) A mechanical model is a description of a mechanism, including (i) a description of the mechanism's behavior; and (ii) a description of the mechanism which accounts for that behavior.

The two-part characterization of a mechanism, in terms of a behavior and the mechanism that produces it, leads naturally to a two-part characterization of a mechanical model. The behavioral description is a description of the external behavior of a mechanism. The mechanical description is a description of the internal structure — the guts of the mechanism.[5] The distinction between behavioral and mechanical descriptions is roughly the distinction between what a system is doing and how it is doing it. It is analogous to a number of other distinctions made in the discussion of complex systems, including Chomsky's (1965) distinction between competence and performance theories, Simon's (1978) distinction between substantive and procedural rationality, and Marr's (1982) distinction between computational and algorithmic theories.

The sense in which the two parts of a mechanical model are descriptions requires clarification. They are not descriptions in the sense of 'descriptive phrases' familiar from Russell's theory of descriptions, nor indeed can they be identified as sets of sentences or as syntactically defined entities of any kind. Rather, they are semantic entities, in the sense that there can exist many syntactically different formulations of the same description. For instance, a model of a pendulum would include a behavioral description which characterized the motion of the pendulum. The same behavioral description could be expressed in a number of ways, using, e.g., different kinds of equations, coordinate systems or diagrams. For one to have different behavioral descriptions, the behavior described must be different. Thus, if we have one model of the pendulum which specifies that the range of motion will dampen over time, and another in which no damping occurs, we have different behavioral descriptions, and hence different models.

Following Giere (1988), I take the relation that obtains between a model and a mechanism is one of approximate similarity. The behavior of the system in nature is described (to varying degrees of approximation) by the model's behavioral description and the internal structure of the system is described (again to varying degrees of approximation) by the model's mechanical description. To make claims about the nature of a mechanism, one constructs a model and asserts that it is similar to a system in nature.

The most important difference between mechanical models and the models of the semantic conception is that mechanical models have two parts. Let us consider this difference with respect to the state space version of the semantic conception. According to this conception, the possible states of the system modeled are represented by the set of vectors in a state space, and a particular model can be understood as a curve through this space. As a very simple example, consider how we might construct a state-space model of an analog watch. The state of the watch at a given time can be characterized by two variables, one representing the position of the hour hand and the other the position of the minute hand. One can then plot the evolution of the state of the watch through time as a periodic curve in this space. From the mechanistic point of view, this description of the watch is only half a model of the watch. It is, in particular, a description of the behavior of the watch. What is lacking is a description of the mechanism that produces this behavior. A mechanical description would characterize how the various parts of the watch (the battery, the quartz crystal, the hands, the internal gears, etc.) cause the hands to move in the way characterized by the behavioral description.

The internal workings of the mechanism can be represented using the state space approach as well. The relevant properties of the interacting parts of the mechanism could be coded as state variables. Laws of succession could be specified that express how changes in values of these state variables lead to changes in the values of state variables representing properties of directly causally linked parts of the mechanism. In this case, the laws of succession are the direct causal laws referred to in (M). What this example shows is that the notion of a mechanical model is more restricted than that of a state-space model. Whether a state space model is a mechanical model depends upon what state variables are chosen, and whether the laws of succession used to characterize the state changes represent direct causal interactions between the parts of the mechanism.

This example shows that it is possible to formulate a mechanical model using a state space representation but that not all state space models are mechanical models. This is because the requirements for a model being a description of a mechanism place substantive constraints on the choice of state variables, parameters, and laws of succession and coexistence. It is meeting these additional constraints which accounts for the explanatory power of mechanical models. The division between the behavioral description and the mechanical description is in essence a division between explanandum and explanans. The mechanism characterized by the mechanical description brings about, and hence explains, the behavior characterized by the behavioral description.[6] If the behavioral description is a statement of a law (in the weak sense described above), then that law is mechanically explicable.

Nancy Cartwright advocates a similar view of the relationship between laws and mechanisms to the one I have suggested here. She does not use the term ‘mechanism’, but instead speaks of ‘nomological machines’:

> It takes very special arrangements, properly shielded , repeatedly started up, and running without a hitch, to give rise to a law; it takes what I call a ‘nomological machine’. … [M]odels serve as blueprints for nomological machines (Cartwright 1997, S292).

What Cartwright apparently has in mind here is the kind of elaborate arrangements often required to produce direct empirical confirmation of theoretical laws. As Cartwright often has emphasized, it takes quite elaborate experimental set ups to isolate and amplify microphysical effects which may be seen to represent the isolated effects of basic theoretical laws. These setups are examples of Cartwright’s nomological machines. In her view, there aren’t many nomological machines, because “laws are scarce”. I would agree with Cartwright in one sense. It takes special conditions to see the effects of the kinds of laws we find in physics textbooks, and most ordinary phenomena are not covered under them. However, if one takes laws to be of the more homely sort I commend, there are actually lots of laws — as many as there are mechanisms — for these laws just describe the behavior of these mechanisms. There is, for instance, a mechanism that produces the spiraling motion of the leaf as it falls from the tree outside my window, and there is a law, in the sense of a counterfactual supporting generalization, about how the leaf behaves as it falls, and how it would have behaved if it were dropped under those conditions. The generalization has very narrow scope, since probably only one leaf in the history of the world will satisfy the conditions requisite to produce the particular pattern of behavior which the actual leaf did. For this reason, we won't ever build a model of that specific mechanism, but there is a law and a mechanism nonetheless.

In concluding this discussion of mechanisms and models, two points should be made which will have significant implications for the subsequent discussion of testing mechanical models. First, notice that the concept of a mechanism's behavior presupposes a concept of normal functioning. When we describe the behavior of a mechanism, we are describing how it will behave if it is not broken. For instance, in describing the behavior of the watch in terms of the periodic rotation of hands, we are presupposing that the watch's battery has not worn out, and many other things of that sort. The "if it not broken" clause is a kind of ceteris paribus clause for behavioral descriptions. This idea of "normal function" is required even for mechanisms that are not the product of design or selection. If, for instance, we describe the behavior of the El Ninõ mechanism, our description presupposes that the normal mechanism by which El Ninõ produces its effects is not disrupted by exogenous factors (like the earth being hit by a large asteroid).

The second point is that there is a one-many relationship between behavioral and mechanical descriptions. This is simply because the same behavior can be produced by different mechanisms. A spring wound watch and a Quartz crystal watch will (relative to most descriptions) behave the same way, even though the mechanism that produces the behavior will be quite different. This fact gives rise to a mechanistic version of the underdetermination thesis. This raises the following question: If one has two competing models of a mechanism which both predict the same behavior, how does one choose? The models described in the remainder of this paper have exactly this structure, and examination of the techniques used to resolve this question will provide us with some more general insights into how models are built, tested and revised.

IV. A Case Study -- Models of Vowel Normalization

In this section I turn to a detailed application of the theory of mechanical models that I have developed in the previous section. I shall present two alternative models that seek to explain a phenomenon in speech perception called vowel normalization. This case study serves two purposes. First, it illustrates how my model of models applies to a real and complex case. Second, it provides insight into how to evaluate competing models of the same phenomena.

The fundamental problem of speech perception is to understand the mechanism by which a listener transforms an acoustic signal into a sequence of phonemes. Acoustically speaking, a speech signal is just an aperiodic sound wave. One way to characterize a speech signal is with an oscillogram, which is simply a plot of the wave's amplitude over time. An alternative way to characterize a wave is with a continuous spectogram. In a continuous spectogram, the complex wave is decomposed into simple sinusoidal waves over a continuous range of frequencies. The result is a three-dimensional plot representing intensity (wave amplitude) at each harmonic frequency over time. Research suggests that the majority of acoustical information required for phonological recognition is coded by the wave spectrum, and in particular by peaks (local maxima) of intensity in the spectrum.

Speech is produced by passing vibrating air from the vocal cords, which function as a forced harmonic oscillator, through the articulatory tract, which functions as resonator. The configuration of the articulatory tract at any given time can be characterized by a frequency response curve. The local maxima in this curve (i.e., the frequencies at which the resonator resonates) are called <u>formants</u>. By altering physiological features of the articulatory tract (the openness of the mouth, the position of the tongue, etc.), a talker can change the position of these formants. Acoustically, formants are detectable as peaks of frequency intensity in a spectogram. Formants are numbered from lowest to highest (F1, F2, etc.) according to the frequency of their peak resonance. The frequency of the formants (especially F1, F2 and F3) turns out to be a major determinant in the perception of different phonemes. Besides formant frequencies, the fundamental frequency of the wave (which is the frequency of oscillation of the vocal cords), called F0, also influences phoneme perception.

The most significant information used by listeners to recognize vowel sounds are the frequencies of the first and second formants. This conclusion is supported by two kinds of evidence. First, artificial vowels that are constructed by masking all but F1 and F2 cues from naturally produced vowels are recognized with a high degree of accuracy by most listeners. Second, differences in F1 and F2 intensities correspond fairly closely to different places of articulation of vowels in the articulatory tract. Front vowels have relatively higher F2 frequencies than back vowels. Low vowels have relatively higher F1 frequencies than high vowels. Using F1 and F2 frequencies it is possible to construct for a map of the speaker's vowel space. An F1/F2 map for an average talker is pictured in figure 2:

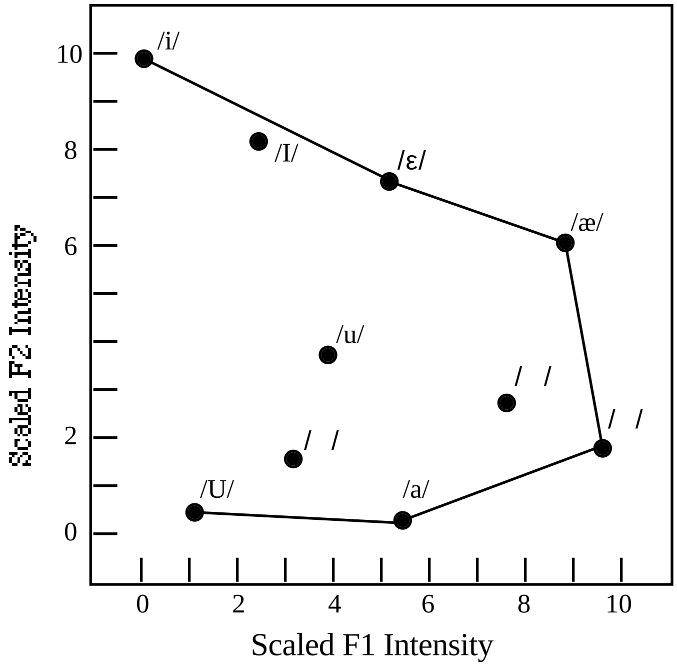

Fig. 2. An Average Talker's Vowel Space[7]

The difficulty with this simple characterization of vowel perception is that it does not take into account variations in the vowel spaces across different talkers. Because different talkers' F1/F2 maps are different, signals that are similar with respect to F1 and F2 frequencies can represent different vowels in different talkers. The process by which listeners adjust to different talkers' vowel spaces is called vowel normalization.

In the remainder of this section, I describe two competing models of this mechanism, one proposed by L. J. Gerstman (1968) and the other proposed by A. K. Syrdal and H. S. Gopal (1986). While my primary concern will be with vowel normalization itself, it is not possible to examine this component of the auditory system in isolation. To evaluate different models of vowel normalization, one must first integrate them into models of the larger recognition process. Models of normalization can be compared indirectly by comparing models of vowel recognition which incorporate different models of normalization. The two models that I will describe are thus not models of vowel normalization alone, but models of the vowel recognition mechanism that incorporate a normalization component.

These models are competing because they represent alternative hypotheses about what mechanism underlies a particular behavior. The two models have a common behavioral description, but differ in their mechanical description. To describe, and ultimately evaluate, these models, we must first examine this common behavioral description.

In the broadest sense, the behavior of a vowel recognition mechanism can be described by a function from acoustic signals to vowel tokens. The acoustic signals of interest are any segments of human-produced speech that contain vowels.[8] A perfect vowel recognition mechanism would be able to take any such signal and identify what vowel(s) the talker intended to produce. This description of vowel recognition behavior is, however, unmanageable. It is just not possible to evaluate predictions of the models versus actual human listener behavior for such a wide range of stimuli.

To make the study of vowel normalization more manageable, researchers have restricted themselves to a small standard data set. The data set used both by Gerstman and by Syrdal and Gopal was produced by Peterson and Barney (1952). Peterson and Barney recorded vowels produced by 76 talkers (a mixture of men, women and children). Each talker read from a randomly ordered list containing two repetitions of ten words ('heed', 'hid', 'head', 'had', 'hud', 'hod', 'hawed', 'hood', 'who'd' and 'heard'). Peterson and Barney then made spectrographic measurements of each vowel, measuring values of F0, F1, F2 and F3. Finally, they tested the intelligibility of these vowels by playing them to twenty-six listeners. 1199 of the 1520 were unanimously recognized.

With respect to this data set, human listeners display good but not complete accuracy in recognizing vowels. This level of performance with respect to the Peterson and Barney data set serves as the (partial) behavioral description of the "average" human vowel normalization mechanism. A model of vowel normalization should identify vowels with approximately the same degree of accuracy as human listeners do.[9] Both the Gerstman and Syrdal/Gopal models satisfy this requirement. The two models therefore have the same behavioral description. Now consider their differing mechanical descriptions.

A. The Gerstman Model

The basic idea behind the model proposed by Gerstman (1968) is that as a talker begins to speak the auditory system acquires information about the F1/F2 position of a small number of vowels, and that this information is sufficient to tune the recognition mechanism to that talker's particular vowel space. The model suggested by Gerstman is composed of four parts, which I shall call the formant frequency analyzer, the calibration mechanism, the normalization mechanism and the identification mechanism. The frequency analyzer is the part of the system that analyses a signal into a frequency spectrum and identifies formants. The calibration mechanism analyzes a calibration signal consisting of an initial segment of a particular talker's speech in order to determine the dimensions of that speaker's vowel space. The normalization mechanism transforms absolute formant frequencies F1 and F2 into frequencies F1' and F2' that are normalized to eliminate variations between different speakers' F1/F2 spaces. The identification mechanism uses F1' and F2' to identify the vowel token. The structure of the overall mechanism is indicated in figure 3 :

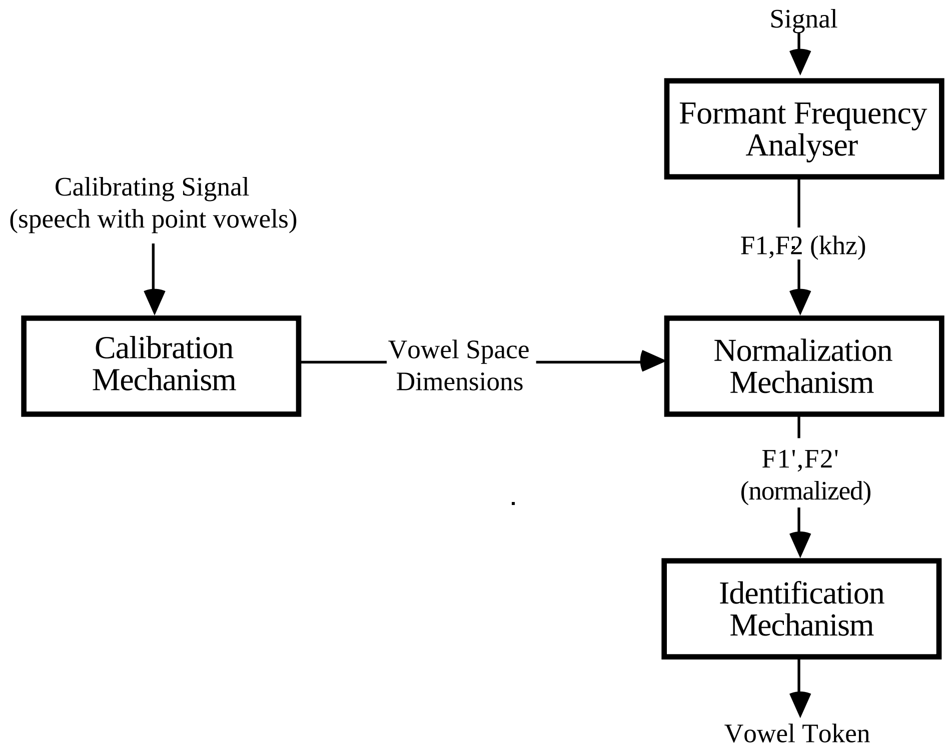

Fig. 3: The Gerstman Vowel Normalization Model

The distinguishing feature of the Gerstman model is that the normalization mechanism requires certain parameters that are not determined by the vowel signal being processed. These tuning parameters must be determined from an initial sample of the speaker's speech. The calibration mechanism is the part of the system responsible for the acquisition of this information. The tuning parameters required by the normalization mechanism represent invariant information about the vowel space of a particular talker. The parameter values will vary from talker to talker but should remain constant for a given talker. Once a listener has identified these parameters for a talker, further calibration is unnecessary.

According to Gerstman's model, the parameters that must be determined by the calibration mechanism are the maximal and minimal F1 and F2 frequencies among the set of vowels used by the talker. In other words, the listener must determine the F1 frequencies of the two vowels with the lowest and highest first formants and the F2 frequencies of the two vowels with the lowest and highest second formants. These vowels are known as the point vowels. I shall denote the minimal and maximal formants $F1_{min}$, $F1_{max}$, $F2_{min}$, and $F2_{max}$. Referring to figure 2, note that $F1_{min}$ is the first formant of /i/, $F1_{max}$ is the first formant of / /, and $F2_{min}$ is the second formant of /a/ and $F2_{max}$ is the second formant of /i/. Reasonable approximations of these values could be achieved using fewer or different vowels (e.g., taking $F2_{min}$ to be F2 for /U/).

Once the tuning parameters have been acquired, the raw F1 and F2 frequencies of vowel signal tokens can be processed by the normalization mechanism. The behavior of the normalization mechanism is described by the following equations:

$$F1' = \frac{F1}{(F1_{max} - F1_{min})} \qquad F2' = \frac{F2}{(F2_{max} - F2_{min})}$$

The output of the normalization mechanism for each vowel processed is a point (F1', F2') with $0 \le F1', F2' \le 1$. This point is then used by the identification mechanism to identify the vowel in question. The identification mechanism works by comparing the position of the point (F1', F2') in the scaled vowel space (see figure 2) with the position of average vowels within that space.[10]

I have already argued that a mechanical model is more than just a set of generalizations about the external behavior of a system; a model purports to describe the structure of the mechanism that accounts for the behavior. The Gerstman model appears to pass the first test of accounting for the behavior of the system, at least insofar as identifying vowels in the Peterson and Barney data set with accuracy comparable to humans counts as predicting the behavior of the actual auditory mechanism. But beyond this, what claims has Gerstman made about the internal structure of the auditory mechanism?

The claims about internal structure are summarized by the diagram of the mechanism in figure 3. Acceptance of the model involves commitment to the view that there are four distinct components of the auditory mechanism which exhibit the behaviors described in the above paragraphs. The internal structure of these components has been left largely unexplained. For instance, the calibration mechanism must in some way acquire F1 and F2 frequencies for point vowels, but Gerstman has not specified how these vowels are identified. They would have to be identified by a mechanism that either normalizes vowels in a different way or does not normalize at all. There are a number of ways in which the calibration mechanism might work: it might rely on other features of the vowel token besides F1 and F2 to identify unnormalized vowels; alternatively, it might rely on lexical or pragmatic information. One might for instance suppose that a listener might rely on the fact talkers say fairly standard things when they first begin speaking ("Hello!", etc.) and use this information to help disambiguate any vowels that are initially ambiguous. A major consideration in evaluating the Gerstman model is whether or not one can give plausible accounts of how this and the other parts of the mechanism hypothesized by that model might work.

B. The Syrdal/Gopal Model

Syrdal and Gopal (1986) have developed a model of vowel recognition that uses a different kind of vowel normalization model. Unlike the Gerstman model, the Syrdal/Gopal model suggests that each vowel is self-normalizing. In other words, the model suggests that the F1 and F2 values of each vowel token can be normalized by using other parts of the acoustic signal. and do not need tuning parameters extracted from the context of the utterance.

Syrdal and Gopal's model is represented schematically in figure 4:

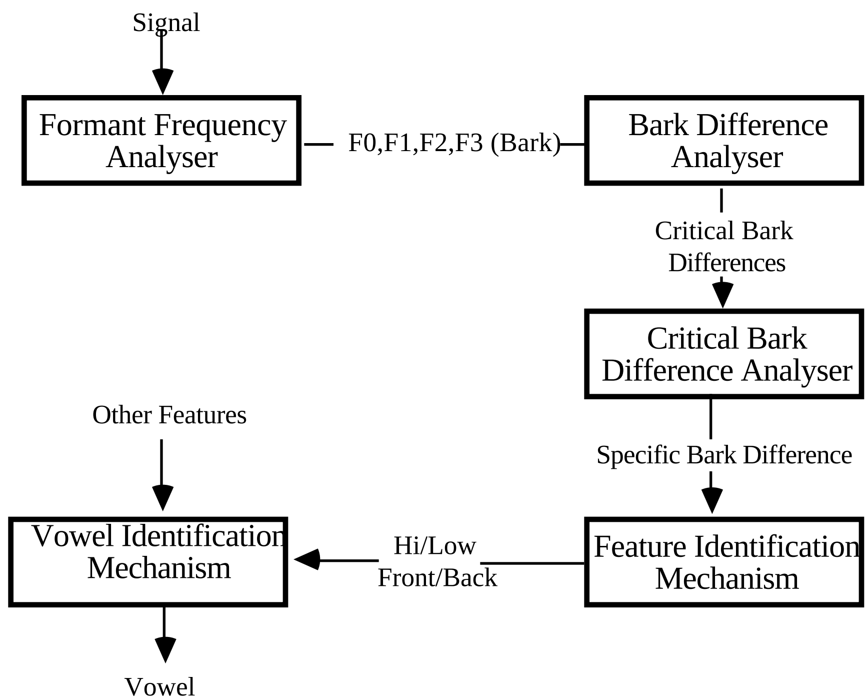

Fig. 4: The Syrdal/Gopal Vowel Normalization Model

Processing of each vowel, according to this model, occurs in five stages. First, the acoustic signal is analyzed into formant pitches measured using the Bark scale.[11] Second, differences between the formants F3-F2 and F1-F0 are calculated. Third, these formant differences are classified into one of two categories depending upon whether or not their difference exceeds a critical distance of 3 Bark. Syrdal and Gopal show that this twofold binary classification is related to the vowel's phonetic features. The relation is given by the following map:

	< 3 Bark	> 3 Bark
F3-F2	front	back
F1-F0	high	mid, low

Since this classificatory schema only distinguishes four possible categories, it is not by itself sufficient to uniquely identify the ten different American vowels being considered. Additional information is required to identify the vowels. Syrdal and Gopal have analyzed critical bark differences in several other dimensions (F2-F1, F4-F2 and F4-F3) and determined that these measures could be used reliably to further discriminate between all but three pairs of vowels (/ɛ/ and /æ/, /u/ and /U/, /a/ and / /). These other critical bark difference dimensions do not correspond to phonetic feature classifications. Other kinds of information (like vowel duration) can also be used to discriminate vowels.

Although Syrdal and Gopal think there is good evidence that the human auditory system uses binary feature classifications to minimize inter- and intraspeaker variability with respect to the production of vowels, they do not commit themselves to one particular computational model for vowel recognition. Instead they categorize vowels using a number of different features including critical bark differences, absolute bark values and vowel duration. They do not try to argue which of these techniques (if any) the human auditory system uses for feature discrimination. For this reason I have left some of the inputs in the vowel identification mechanism in figure 3 unspecified. These details turn out not to be significant for the kinds of tests described in the next section. What is critical is the general approach to normalization and identification. Syrdal and Gopal believe that these discriminations can be made based upon the structure of the signal being identified, regardless of what particular acoustic features are actually used to make binary discriminations. This is the sense in which the Syrdal/Gopal model is self-normalizing,.

The Gerstman and Syrdal/Gopal models illustrate several noteworthy features of mechanisms and mechanical models. The most important characteristic of mechanisms is that they are complex structures consisting of a number of parts. Thus, a mechanical model must specify the hypothesized parts of the mechanism. The parts of the Gerstman and Syrdal/Gopal mechanisms are given respectively in the boxes found in figures 3 and 4. These diagrams do not simply specify a list of parts, but also their causal arrangement. In particular, the arrows represent the causal interactions that occur between the different parts. The flow chart representation I have used to summarize these mechanisms is certainly not the only possible way to present a mechanical model, but it is an especially natural one. And while these diagrams are not taken directly from the scientific literature on vowel normalization, they represent the models in a way that would be intelligible to scientists working in this domain. Diagrams of this general type, with graphical representations of a mechanism's parts, together with indications of the types of interactions between these parts, can be found in the scientific literature for many fields.

If one compares the diagrams in figures 3 and 4 with visual representations of more paradigmatic mechanisms, like schematics for electronic or mechanical devices, what is striking is that there is no indication of the size, location or arrangement of the parts. Parts are instead specified functionally, that is, in terms of the causal role of the part within the overall mechanism. Note that the parts are themselves complex mechanisms. Their behavior is described by the functional characterization, with the mechanism underlying the behavior left unspecified. This strategy has an important advantage. By leaving out the details by which these functions are implemented, the model highlights the explanatorily relevant features of the mechanism.[12] The correctness of these explanations will not turn on how it is, for instance, that the formant frequency analyzer works, so long as there exists some mechanism that performs this function.

V. Testing and Refining Mechanical Models

The final issue I shall consider in this model of models is what procedures can be used to test and refine mechanical models. I begin by discussing in general some of the available strategies, and then consider how these strategies have been applied in the case of the Gerstman and Syrdal/Gopal models.

To assess the adequacy of a mechanical model, one must assess the adequacy of both the behavioral and mechanical descriptions of the model. In other words, one must assess both the extent to which the behavior predicted by the model conforms to the behavior of the mechanism and the extent to which the model's description of the mechanism's parts and their interactions conform to the actual workings of the mechanism.

In cases where mechanical models make substantially different claims about the behavior of a mechanism, at least part of the task of model testing is straightforward. Experiments may be performed to see if the mechanism in question behaves as is predicted by each of the models, and any model that fails to adequately describe this behavior may be rejected.[13] However, saving the phenomena is not a sufficient condition for a model being a good model of the mechanism, for the mechanical model claims to do more than merely predict the mechanism's behavior. This point is illustrated by the two models under consideration. Though both models save the phenomena, they postulate very different mechanisms underlying the phenomena. The actual mechanism may be dissimilar to one or both of them. This is a very practical version of the underdetermination problem. But unlike most philosophical versions of the underdetermination problem, where we imagine theories that are underdetermined by an infinite set of actual and possible observation statements, there is a well-understood strategies for choosing between alternative models.

Probably the most obvious strategy is to try to take the mechanism apart and study the behavior of its parts. The functional localization strategies explored by Bechtel and Richardson (1993) and the strategies discussed by Craver and Darden (forthcoming) for discovering neurobiological mechanisms are principally strategies of this kind. These strategies are important, and indeed a good deal of work on mechanisms of speech perception relies upon them. However, there are sometimes circumstances in which it is impractical or impossible to dissect a mechanism. In the case of high level cognitive mechanisms like those we are considering, the parts themselves may be complex and highly distributed and may defy our strategies for localization. Moreover, in the particular case of speech research where we have no model organisms other than human beings, ethical considerations place considerable constraints on the kind of physical dissection of mechanisms that we might otherwise undertake. Thus in cases such as this one must utilize testing techniques that allow one to make inferences about the internal structure of a mechanism simply by examining the mechanism's external behavior. These strategies typically involve examining the behavior of the mechanism under non-standard conditions. Placing the mechanism in these non-standard conditions will in some cases cause it to behave in unusual ways—that is, it may break the mechanism. Observing when and how a mechanism breaks can tell one a lot about how a mechanism works. Let us consider how such a strategy can be used in comparing the two speech perception models under review.

The Gerstman and Syrdal/Gopal models save the same class of phenomena, because both perform comparably to human listeners in identifying vowels in the Peterson-Barney data set. Furthermore, it is hypothesized (plausibly, but without much data) that similar behavior could be maintained under a variety of more natural conditions. But in each case, the generalizations describing the behavior of the normalization models must be understood as having a ceteris paribus clause attached to them. For generalizations about the behavior of mechanisms, ceteris paribus clauses will not be violated so long as nothing has occurred to disrupt the functioning of the mechanism. The crucial point here is that alternate mechanisms which normally would produce similar or identical behaviors have different conditions under which they break down. Investigating when and how mechanisms break thus can provide information about what is going on inside the mechanism.

Though, in response to ordinary human speech, both models mimic the normalization capacities of the human auditory system, the models rely upon different features of that speech to perform the normalization. To evaluate the models one creates contexts in which some of ordinarily available features are absent and in which, consequently, one of the models predicts a breakdown (i.e., a misidentification or ambiguous identification of a vowel). If the actual mechanism breaks down in the way in which one model predicts but the other does not, then the breakdown constitutes evidence for the first model.

Nusbaum and Morin (1991) have conducted experiments that use these strategies to evaluate the models in question. The results of these experiments neither falsify nor verify either of the models in their entirety. Instead, Nusbaum and Morin identify certain features of the models and perform experiments that seek to determine whether or not the actual mechanisms used in vowel normalization and recognition have these features. I shall structure my discussion of the experimental results in terms of five features: contextual versus structural tuning, use of F0 and F3 information, use of Bark Scale normalization, sensitivity to intraspeaker vowel variations, and correlations with phonetic features. Nusbaum and Morin's investigation of these features suggest that actual speech perception mechanisms combine features from both the Gerstman and Syrdal/Gopal models.

Contextual versus Structural Tuning: The most significant structural difference between the two mechanisms has to do with the stage at which they acquire the information necessary to normalize a particular vowel token. In the Gerstman model, the tuning mechanism must acquire parameters that describe the talker's vowel space prior to processing a vowel token. In Syrdal and Gopal's model the parameters for normalization are carried by the vowel token itself. Nusbaum and Morin (1991) have characterized this difference as a difference between "contextual tuning" theories and "structural estimation" theories:

> Contextual tuning theories state that listeners use a sample of speech to either calibrate or learn vocal characteristics of the talker that are necessary to interpret the phonetic structure of subsequent utterances. . . . By contrast structural estimation theories state that each sample of speech is self-normalizing. Information from within the structure of an utterance is used to estimate the vocal characteristics of the talker (Nusbaum and Morin 1991, 114).

How does one experimentally investigate whether or not the normalization mechanism uses contextual tuning (as Gerstman suggests) or structural estimation (as Syrdal and Gopal suggest)? Nusbaum and Morin have devised and performed an experiment that provides some clues. Their experimental design relies on the following considerations: A contextual tuning mechanism relies upon cues from some sample of speech to calibrate the normalization mechanism. If these cues are removed by isolating a vowel to be normalized from other samples of the talker's speech, then on the basis of a contextual tuning model, one would expect a degradation in the recognition capacity of the speaker. On the other hand, if vowel tokens are self-normalizing, no degradation would be expected.

Nusbaum and Morin's experiment compares the performance of human subjects in recognizing vowels under two different conditions. In both conditions listeners are asked to repeatedly listen to sequences of sixteen syllables and listen for occurrences of a particular vowel. Each stimulus sequence contains four instances of syllables containing the target vowel, randomly located among twelve syllables containing other vowels. Subjects are asked to press a button whenever they hear a syllable containing the target vowel. In the first (mixed-talker) condition, the target and distracter syllables are chosen from a pool of syllables spoken by a variety of talkers. In the second (blocked-talker) condition, all sixteen stimuli for a given sequence are spoken by a single talker. If the normalization mechanism using a contextual tuning mechanism then one would expect there to be a significant degradation in subject performance in the mixed-talker condition compared to the blocked-talker condition. If a structural estimation mechanism is being used then no such degradation should occur.

Nusbaum and Morin found that there was no significant degradation in the accuracy of vowel identification between the blocked and mixed-talker conditions. However, they did find that subjects' response times were significantly faster in the blocked than mixed talker conditions. These results led them to hypothesize that there are multiple mechanisms available for vowel normalization. They hypothesize that in the blocked-talker condition a contextual tuning mechanism is used while in the mixed-talker condition, where no contextual cues are available, a structural estimation mechanism is used. Furthermore, the increased processing demands of the structural estimation mechanism account for the increase in reaction time.

To provide further evidence, Nusbaum and Morin investigated the effects of increasing cognitive load on performance in both the blocked and mixed-talker condition. They hypothesized that demands on attention might decrease speed and accuracy of responses. They discovered that increasing cognitive load by requiring subjects to perform distracting tasks during a trial significantly degraded response time in the mixed-talker condition but not in the blocked-talker condition. From this they concluded that structural estimation mechanisms require significantly more attention than contextual tuning mechanisms.

F0 and F3 information: A second feature that differentiates the Gerstman and Syrdal/Gopal models is the role of information other than F1 and F2. The Syrdal/Gopal model uses F0 (pitch) and F3 frequencies to provide the additional information that normalizes the vowel space whereas Gerstman's model relies on F1 and F2 alone. Since in most conditions all formant and pitch information will be available, this difference in the models would not typically lead to differing behavioral predictions.[14] Nevertheless it is possible to construct auditory stimuli in which F0 or F3 information is not present. If F0 and F3 information is required for normalization, as is suggested by Syrdal/Gopal model, then one would expect a significant degradation in accuracy of vowel identification. To test this hypothesis, Nusbaum and Morin evaluated listener performance under four conditions. Starting with a set of stimuli containing all formant and pitch information, they produced additional sets of stimuli in which some of that information was removed. In one stimulus set, all formant and pitch information was maintained; in a second, formants above F2 were removed; in a third, pitch (F0) was removed; and in a fourth both formants above F2 and pitch were removed. These conditions can be referred to by a natural notation referring to which frequencies are present as (0-3), (0-2), (1-3) and (1-2). In trials where all stimuli came from a single talker (and thus normalization was not a factor), subjects were able to identify vowels correctly in approximately 95% of cases in each of the four conditions. However, in trials where vowels came from a mixed group of talkers there were significant variations in performance among the different conditions. Subjects performed approximately equally well in the (0-3) and (0-2) conditions. They performed markedly worse in the (1-3) condition and worse still in the (1-2) condition. From these results, Nusbaum and Morin concluded that pitch information is used in structural estimation, and that formants higher than F2 may provide additional estimation cues.

Nusbaum and Morin's results provide some evidence for accepting the Syrdal/Gopal model. For several reasons, however, their results should not count as decisive evidence for the model. First, although Nusbaum and Morin have shown that F0 (and perhaps F3) information is significant, the Syrdal/Gopal model makes some specific claims about how that information is used; F0 is used in conjunction with F1 to evaluate vowel height, etc. Nusbaum and Morin's results cannot be used to evaluate the Syrdal/Gopal model at that level of detail. Furthermore, as Nusbaum and Morin point out, even under the worst conditions, subjects are able to identify vowels with approximately 80% accuracy. The Syrdal/Gopal model does not really explain this rate of success. It is of course possible that the model could be extended to do so, but the Syrdal/Gopal model does not include details of how the normalization mechanism would perform in the absence of F0 and F3 information. Unless the model makes predictions about the behavior of the normalization mechanism (specifically its accuracy) under such circumstances, it is difficult to evaluate the significance of the 80% success rate in the (1-2) condition.

<u>Bark Scale Normalization:</u> Another way in which the models differ is with respect to their frequency analysis mechanism. According to both models, the initial stage of processing of a vowel token involves the analysis of the acoustic signal into pitch and formant frequencies. But it is by no means evident that the perceptual frequency (i.e., pitch) is perceived in such a way that there is a simple linear relationship between it and acoustic frequency. In fact, the most widely accepted current theory suggests that (1) all sound intensity within a certain critical band of acoustic frequencies is integrated and perceived as a single pitch; and (2) the width of this critical band is not constant over the perceptual range of the auditory mechanism. The critical band model has been tested by a variety of methods (Scharf 1970). Empirical methods have, on the supposition of the critical band model, led to estimates of critical bandwidth at various frequencies (Zwicker 1961); based upon these empirical measures, Zwicker has developed a pitch scale called the Bark scale. Bark varies approximately linearly with frequency up to 500hz and approximately logarithmically thereafter (Syrdal and Gopal 1986, 1087). The Syrdal/Gopal model assumes that the frequency analysis component produces Bark-scaled pitch and formant information whereas the Gerstman model assumes that the auditory scale for pitch and formant frequencies is linear with respect to actual acoustic frequencies.

There is an abundance of perceptual evidence that pitch perception is not linear with frequency (Scharf 1970). Furthermore, there is some evidence that there are physiological correlates of critical bands. Each critical band seems to correspond to a fixed length along the basilar membrane (Syrdal and Gopal 1986). This evidence seems to argue decisively in favor of a critical band type of frequency analyzer, and thus for Syrdal and Gopal's model. We should, however, be careful not to allow this evidence to lead to the complete rejection of Gerstman's model. While it shows that Gerstman's model of the frequency analysis component of the mechanism is probably incorrect, it does not show anything about other aspects of his model. It would certainly be possible to construct a contextual tuning model like Gerstman's in which the linear scaling frequency analyzer is replaced by a non-linearly scaled frequency analyzer and appropriate adjustments are made to the other components of the normalization mechanism.

Intraspeaker Variability: So far, we have evaluated these two models in terms of their ability to normalize variations between vowels uttered by different speakers in similar lexical and prosodic contexts. The chief data set under consideration is the Peterson and Barney (1952) set of vowels uttered in identical monosyllabic contexts by different speakers. This data set minimizes the variations in acoustic properties among tokens of the same vowel. In reality, however, there are significant variations in acoustic properties of different tokens of the same vowel produced by the same speaker (Gopal and Syrdal 1984, Syrdal and Steele 1985). These variations are attributable to variations in the syllabic, lexical or prosodic context. There is paralleling the interspeaker normalization problem an intraspeaker normalization problem. Syrdal and Gopal (1986) argue that one of the virtues of their normalization model is that it accounts for reduction of both between-speaker and within-speaker variability. Though this is clearly a virtue of their model, it should again not be taken as knockdown evidence for its acceptance. In the first place, it might well be the case that other models such as Gerstman's also reduce intraspeaker variability. In the second place, it is possible that actual human vowel recognition mechanisms do inter- and intraspeaker normalization in different ways.

Feature Identification: Another desirable feature of any normalization model would be that it explicitly provide a way to translate acoustic vowel tokens into positions in a phonetic feature matrix which in turn allows for vowel identification. According to widely accepted phonetic theories, different phonemes (including vowels and consonants) can be described in terms of their position in a multi-dimensional matrix of features. American vowels can, for instance, be uniquely described by the features of vowel height, front/back location and roundedness. These feature classifications are important because they can be identified with physiological features of the articulatory tract, and the configuration of the articulatory tract, described in terms of these features, is relatively invariant across speakers and across syllabic, lexical and prosodic contexts. Quite naturally, then, both Gerstman and Syrdal and Gopal have commented on the mapping between values internal to their models and phonetic features. Gerstman has noted that low F1' values indicate high vowels while high F1' indicate low vowels. Similarly Syrdal and Gopal have noted that F1-F0 values less than three bark indicate middle or low vowels while values greater than three bark indicate high vowels. Since both of these models have this feature, it cannot be grounds to prefer one over the other. Nevertheless, features of this kind can be quite important in the evaluation of models. This kind of virtue is not a virtue that a model has isolated from the context of other models. The reason that the correspondence between phonetic features and internal variables within the models is significant is that there are models of related mechanisms — i.e., mechanisms for vowel production — in which these features are significant.

VI. Conclusion

The construction and evaluation of mechanical models can be a very complex business, as this case study of models of vowel normalization has illustrated. Let us now step back and consider what lessons can be learned from our study of these models.

First, the results of Nusbaum and Morin's experiments clearly show that observation and experiment will not generally provide evidence that will either conclusively confirm or disconfirm a model. The reason for this is not just that decisive evidence is hard to come by, but rather that the posited relationship between a model and the mechanism it models is one of similarity rather than isomorphic correspondence. Giere (1988) has argued, especially in the context of models of classical mechanics, that models are evaluated in terms of their similarity to modeled systems, and our discussion of vowel normalization models supports this view. While some authors (e.g., Hughes 1997) argue that the concept of similarity is a vague and misleading, in the context of mechanical models, the claim that models are similar to varying degrees with real systems can be spelled out in ways that make it unproblematic. Models like the two considered here can be construed as making a wide variety of claims about the properties of human vowel normalization mechanisms. They make hypotheses about what the components of this system are, specifications of the behavior of these components, and claims about the patterns of interaction between them. A model may easily be correct with respect to some of these claims and incorrect with respect to others. Thus, for instance, we saw that experimental evidence suggests that formant frequencies are grouped within critical bands, and that this evidence supports the Syrdal/Gopal model. But similarity in this respect does not imply that the Syrdal/Gopal model is correct in other respects, or that the Gerstman model is incorrect in other respects.

It is difficult to specify a principled way in which the degrees and respects of similarity and dissimilarity should be weighed against each other, and hence it may be difficult to say of two models which is the more similar to the modeled system. But this is hardly a defect in the analysis of models presented here. Indeed, it goes some way towards explaining why scientists may continue to test and articulate a number of different and incompatible models.

A second major conclusion we can draw is that there is no principled division between the discovery and testing of mechanical models (cf. Darden 1991, Craver and Darden, forthcoming). While the experimental results do indeed show that both models are not wholly adequate models of the mechanism in question, the results are interesting largely because they suggest strategies for further articulations and revisions of the model. Consider again the evidence that formant frequencies are grouped within critical bands. This evidence does not require rejection of the Gerstman model, but it does suggest that the frequency analysis component of the model should be revised. Similarly, the fact that there remains evidence for some use of contextual tuning suggests that the actual mechanism may have a variety of partially redundant sub-mechanisms that use different techniques for normalization. (In fact, the use of multiple mechanisms and information channels to increase reliability is a characteristic of many perceptual mechanisms.)

A third and final lesson we can draw from this study is that mechanistic analysis is possible and important not just in those "harder" sciences in which the component parts of mechanisms are more easily isolated and studied, but can also be used in disciplines like cognitive psychology where the complexity of the systems studied is such that scientists (at least as yet) are unable to localize and study these components.

This is one respect in which my analysis of mechanical models is more general than that offered by Machamer, Darden and Craver (2000). In their discussion of how scientists theorize about mechanisms, they use the terms 'mechanism schema' and 'mechanism sketch' instead of my 'mechanical model'. For them, a mechanism schema is a "complete" model, i.e., a model where all parts of the mechanism are identified and spatially located, and where the interactions between parts are well understood. A mechanism sketch, on the other hand, is an incomplete model—one in which there are still blanks or question marks. In their view, one's goal in studying a mechanism is to come up with an adequate mechanism schema; an adequate mechanism schema is what is required to explain the phenomenon for which the mechanism is responsible.

I am not persuaded that the distinction between a sketch and a schema is ultimately that useful in understanding how generally we theorize about mechanisms. I prefer rather to say that models come in various degrees of articulation. According to the characterization they give, the vowel normalization models I have described qualify as sketches rather than schemata. The chief reason for this is that the parts referred to in the models are not, at least as yet, spatially localized. But while it is obviously desirable for purposes of confirmation and explanation that parts should be localized, this desideratum is not so fundamental as to merit a distinction in kind between sketches and schema. The discussion of these vowel normalization models shows that they are specific enough to be empirically testable and to provide at least partial explanations of the phenomenon of vowel normalization.

In the case of vowel normalization mechanisms, it is probable that parts can be further localized with future research, but the problem in physically locating parts of mechanisms can be more than epistemic. Sometimes physical location is simply not one of the enduring properties of a mechanism's parts. Consider a technological example: the transmission of an e-mail message over the Internet. Such a phenomenon is amenable to mechanical analysis. The parts of the mechanism are the host computer, servers, switching nodes, packets, etc. A description of the mechanism will describe the manner in which the message is divided into packets, routed over nodes, and reassembled at the remote host. These parts do not, however, have distinct locations. Of course, simply as a consequence of the supervenience of the mechanism of transmission on physical hardware, at any given moment in any transmissions these parts will have physical locations, but these locations will be transient even within the transmission of a single message. Any model that assigned locations to parts (say by referring to their physical memory locations) would be defective. This problem is just another instance of the multiple realizability (Fodor 1974) or dynamical autonomy (Wimsatt 1994) of higher level phenomenon. My discussion of the mechanisms involved in speech perception however illustrates that this autonomy does not prevent mechanical analysis. It is important to see that mechanistic analysis is possible in such cases, for as one considers more complex and higher level phenomena (especially phenomena in the domain of the social sciences), it will be increasingly impossible to demand that mechanistic models characterize their parts in terms of spatial or other physical properties.

Although the term ‘mechanical model’ is a new one, much of the work of the last twenty years on how models are constructed, tested and revised is consistent with the view of models I have presented here. It is my hope, however, that explicit recognition of the fact that much scientific modeling is modeling of mechanisms will increase our understanding of the modeling process.

Notes

[1] See Suppe 1974, p. 228 for a sketch and references.

[2] It should be noted that my terminology differs from that of Suppe and other semantic theorists who follow him. Suppe (1989) uses the term 'physical system' to refer NOT to the actual physical system, but rather to the "abstract replica" of that system — that is, the model. What I call a 'physical system' is what Suppe refers to as 'the phenomena'. I confess to finding his usage strange, because the physical systems on this view are decidedly not physical.

[3] This argument is most clear when we think of scientific models and theories as both being specified by formal theories. They may of course be specified by other means.

[4] See Darden (1996) for a discussion of the differences between abstraction and generality.

[5] The distinction between internal structure and external behavior, while suggestive, can be misleading in some cases. Many mechanisms with spatially localizable boundaries may have their external behavior described in terms of features of those boundaries and their internal structure described in terms of interactions of parts that lie within those boundaries. A description of a Coke machine would have this character. Sometimes, however, the distinction between external and internal can only be made in a metaphorical, functional sense. When, for instance, one describes a software library in terms of its external interface and its internal implementation, the distinction cannot be taken to be literally a physical distinction because neither the interface nor the implementation can be said to occupy a physical position (except in the uninteresting sense that for any given computer and at any given time the software, if loaded, will occupy physical locations in memory). Although often the spatial and functional senses of the external/internal distinction will coincide, they need not; and when they do not, it is the functional sense that is of importance to my account.

[6] It may be, for deterministic mechanisms, that the behavioral and mechanical descriptions can be so stated that the former is entailed by the latter, producing a classic DN explanation. In my view,

however, the locus of explanatory insight is “in the objects”. It is the fact that the mechanism brings about the behavior (together with the satisfaction of necessary pragmatic constraints), not that the mechanical description entails the behavioral description, which yields the explanation.

[7] This graph is produced from average F1 and F2 frequencies collected by Peterson and Barney (1952) and scaled to values between 0 and 10 by Gerstman (1968). This graph is identical to one that appears in Gerstman 1968 except that I have replaced Gerstman's notation for vowels with the now standard one used by Syrdal and Gopal.

[8] It would not be realistic to isolate vowels from phonetic contexts, since such contexts have an effect upon how the vowel signal is produced and recognized.

[9] Of course there is more to saying that the model predicts actual behavior than merely that it gets the approximately same percentage of correct identifications as humans do. In fact, if one got the mechanism right, the model might be expected to do better since it does not have to cope with extraneous cognitive factors like lack of attention. The important point is that it has to get similar behavior for the right reasons — because it uses a similar mechanism. It would, for instance, be damning to a model if all of its misidentifications were of a vowel that is never misidentified by a human listener.

[10] There is some detail that I am ignoring here. Vowel tokens may have normalized first and second formant values which do not obviously identify them as one vowel or another. This happens when the point (F1’, F2’) lies a similar distance from two or more vowels. Gerstman notes that, for the Peterson and Barney data set under consideration, the tokens that have this property are the same tokens that were not unanimously categorized by a panel of 26 listeners. Gerstman then proposes an algorithm that classifies even these tokens with high accuracy. That algorithm involves examining the position of vowels within a space with dimensions (F1’, F2’, F1’+F2’, F1’-F2’).

[11] Perceived pitch does not vary linearly with actual frequency of pure tones. The Bark scale is one of a number of non-linear scales. The significance of the Bark scale is discussed in section V.

[12] The point here is analogous to the one Kitcher (1985) makes about the explanation in classical genetics. In many cases, the molecular "implementation" details are irrelevant to understanding how cytological mechanisms produce phenomena of segregation and independent assortment.

[13] It might be argued that one need not reject a model on these grounds, and it is true that empirically inaccurate models may have considerable value in characterizing mechanisms, especially at the beginning of an investigation. But one can say that no model will be wholly adequate if its predictions of behavior do not conform at least approximately to the behavior of the actual mechanism.

[14] In the case of F0 there are some "normal" contexts in which F0 information is absent. In whispered speech, the vocal cords do not vibrate; rather noise is produced by a rush of turbulent air through the narrow glottal opening.

References

Beatty, J.: 1981, 'What's Wrong with the Received View of Biological Theory', in PSA 1980: Proceedings of the Philosophy of Science Association, Vol. 2, 397–426.

______: 1995, 'The Evolutionary Contingency Thesis', in J.G. Lennox and G. Wolters (eds), Concepts, Theories, and Rationality in the Biological Sciences, University of Konstanz Press, Konstantz and University of Pittsburgh Press, Pittsburgh, 45-81.

Craver, C. and L. Darden: forthcoming, 'Discovering Mechanisms in Neurobiology: The Case of Spatial Memory', in P. Machamer, R. Grush and P. McLaughlin (eds.), Theory and Method in Neuroscience, University of Pittsburgh Press, Pittsburgh.

Cartwright, N.: 1983, How the Laws of Physics Lie Clarendon Press, Oxford.

______: 1997, 'Models: The Blueprints for Laws', Philosophy of Science 64 Supplement (Proceedings of the 1996 Biennial Meeting of the Philosophy of Science Association), S323-S336.

Chomsky, N.: 1965, Aspects of the Theory of Syntax, MIT Press, Cambridge.

Darden, L.: 1991, Theory Change in Science: Strategies from Mendelian Genetics, Oxford University Press, New York.

______: 1996, 'Generalizations in Biology', Studies in the History and Philosophy of Science 27:3, 409-419.

Fodor, J.: 1974, 'Special Sciences, or The Disunity of Sciences as a Working Hypothesis' Synthese 28, 97-115.

Gerstman, L.J.: 1968, 'Classification of Self-Normalized Vowels', IEEE Transactions on Audio Electroacoustics AU-16, 78-80.

Giere, R.: 1988, Explaining Science University of Chicago Press, Chicago.

Glennan, S.: 1992, 'Mechanisms, Models and Causation' Ph.D. Dissertation, University of Chicago.

______: 1996, 'Mechanisms and the Nature of Causation', Erkenntnis 44, 49-71.

______: 1997, 'Capacities, Universality and Singularity', Philosophy of Science 64:4, 605-626.

Hughes, R.I.G.: 1997, 'Models and Representation', Philosophy of Science 64 Supplement (Proceedings of the 1996 Biennial Meeting of the Philosophy of Science Association), S323-S336.

Lloyd, E.: 1988, The Structure and Confirmation of Evolutionary Theory, Greenwood Press, New York.

Lewontin, R.C.: 1974, The Genetic Basis of Evolutionary Change, Columbia University Press, New York.

Machamer, , L. Darden and C. Craver: 2000, 'Thinking about Mechanisms', Philosophy of Science 67, 1-25.

Marr, D.: 1982, Vision: A Computational Investigation into the Human Representation and Processing of Visual Information,W. H. Freedman, San Francisco.

Morgan, M. and M. Morrison (eds): 1999, Models as Mediators Cambridge University Press, Cambridge.

Nusbaum, H. C., and T. Morin.: 1991, 'Paying Attention to Differences Among Talkers', in E. Vatikiotis-Bateson and Y. Sagisaka (eds), Speech Perception, Production and Linguistic Structure, IOS Press, Amsterdam.

Peterson, G. E., and H. L. Barney: 1952, 'Control Methods Used in a Study of Vowels', Journal of the Acoustical Society of America 24, 175-84.

Schaffner, K.: 1993, Discovery and Explanation in Science and Medicine, University of Chicago Press, Chicago.

Scharf, B.: 1970, 'Critical Bands', in J. V. Tobias (ed), Foundations of Modern Auditory Theory, Vol. 1, Academic Press, New York.

Simon, H.A.: 1978, 'Rationality as Process and as Product of Thought', American Economic Review 68:2, 1-16.

Sober, E.: 1997, 'Two Outbreaks of Lawlessness in Recent Philosophy of Biology', Philosophy of Science 64 Supplement (Proceedings of the 1996 Biennial Meeting of the Philosophy of Science Association), S458-467

Sober, E. and R. C. Lewontin: 1982, 'Artifact, Cause and Genic Selection', Philosophy of Science 49, 157-80.

Suppe, F. (ed.): 1974, The Structure of Scientific Theories, University of Illinois Press, Chicago.

______: 1989, The Semantic Conception of Theories and Scientific Realism, University of Illinois Press, Chicago.

Suppes, : 1960, 'A Comparison of the Meaning and Uses of Models in Mathematics and the Empirical Sciences', Synthese 12: 2/3, 287-300.

______: 1962, 'Models of Data', in E. Nagel, Suppes, and A. Tarski (eds.), Logic, Methodology, and the Philosophy of Science (Proceedings of the 1960 International Congress, vol. 1, Stanford University Press, Stanford, 252-261.

______: 1967, 'What is a Scientific Theory', in S. Morganbesser (ed.) Philosophy of Science Today, Basic Books, New York, 55-67.

Syrdal, A.K., and S.A. Steele: 1985. 'Vowel F1 as a Function of Speaker Fundamental Frequency', Journal of the Acoustical Society of America 78:supplement 1, 56.

Syrdal, A.K., and H.S. Gopal: 1986. 'A Perceptual Model of Vowel Recognition based on the Auditory Representation of American English Vowels', Journal of the Acoustical Society of America 79, 1086-1100.

Wimsatt, W. C.: 1994, 'The Ontology of Complex Systems: Levels, Perspectives and Causal Thickets', Canadian Journal of Philosophy, su vol. 20, 207-274.

Zwicker, E.: 1961, 'Subdivision of the Audible Frequency Range into Critical Bands (Frequenzgruppen)', Journal of the Acoustical Society of America 68, 248.