

The role of experiments in phonological investigation

David Eddington

Abstract

The psychological status of many phonological analyses has been questioned on several grounds. Experimental evidence is presented as a crucial part of an analysis that professes psychological significance. A strong and weak sense of psychological reality are defined and discussed in relationship to the traditional and experimental approaches to phonology. Several criticisms of experimental evidence are presented and contested. Examples of experimentally acquired evidence are given, which address the issue of the psychological reality of certain phonological analyses.

Introduction

The goal of phonological investigation has always been to discover and systematize the patterns and generalizations which are found in the phonological systems of natural languages. This could be called the search for linguistic realities. Since the advent of generative linguistics, the search for linguistic realities has been extended, and the search for psychological realities rekindled following the behaviorist era.

In this paper, I discuss the role of experiments in the search for phonological analyses which are relevant to linguistic cognition. In the first section, I argue that most phonological analyses adequately describe linguistic realities. However, they are not necessarily adequate descriptions of psychological realities as well. I claim that a stronger case may be made for the psychological relevance of an analysis which includes experimental evidence.

Of course, the term 'psychological reality' needs to be defined. There are two different meanings which are generally associated with it. I discuss how these meanings relate to the traditional and experimental approaches to phonology. In addition, I present and contest several objections to the experimental approach which have been voiced over the years. In the final section, I review some examples

of experimentally acquired evidence which bear directly on the issue of the psychological significance of certain phonological analyses.

Linguistic realities are not necessarily psychological realities

Over the years, several reasons have been given for questioning the psychological relevance of phonological analyses. The first has to do with the assumption that phonological phenomena are necessarily psychological phenomena as well. In the course of performing a phonological analysis, it is entirely conceivable to discover 'linguistic realities' which have no counterpart in psychological processes such as storage, production or comprehension. A number of linguists have discussed this topic (Botha 1971; Chomsky 1980b:45; Derwing, Prideaux, and Baker 1980; Goyvaerts 1978; Lass 1976; Morin 1988; Skousen 1989). It follows then, that the search for linguistic realities and the search for psychological realities, belong to different but related fields of inquiry (see Ingve 1986, 1996).

Another reason for questioning the psychological content of phonological analyses is their nonempirical status. It has been widely argued that linguistics has more in common with nonempirical sciences such as formal logic, pure

mathematics, and philosophy, than with the empirical sciences (Botha 1971, 1973; Derwing 1973; Hall 1987; Itkonen 1976, 1978a, 1978b, 1983; Katz 1981, 1985; Katz and Postal 1991; Lass 1976; Ringen 1975; Steinberg 1975; Yngve 1986, 1996). Empirical sciences deal with theories that are testable (Bunge 1980: 31-33). This means that they are subject to possible refutation based on events which take place across time or through space (Itkonen 1978a:80; Popper 1968:27-48, 102-3). Since most phonological analyses do not deal with spatiotemporal events, they are nonempirical by nature.

An analysis which can be proven or disproved on the basis of spatiotemporal events possesses a sense of tangibility and concreteness. This is because events which are considered real (as opposed to abstract) take place in time and space. This same sense of tangibility is missing in a theory for which there are no spatiotemporal manifestations. Therefore, a phonological analysis which does not relate to spatiotemporal evidence is a nonempirical analysis of a linguistic reality, and not necessarily one which gives insight into human language capabilities.

The concept of phoneme, for example, is fairly abstract. Unlike allophones, one cannot measure the frequency or duration of a phoneme. Initially, then, it appears that phonemes have no spatiotemporal manifestations apart from their respective allophones. However, the reality of the phoneme has been demonstrated

in studies of perception (e.g. Jaeger 1980), and perception is a process that occurs in time.

Nevertheless, many contemporary phonological analyses are regarded as analyses of linguistic instead of psychological realities because of the evidential base on which they are founded. Included in this category are those analyses that are founded almost exclusively on data from internal reconstruction, which is interpreted in accordance with theory-internal assumptions. The objection which is commonly raised against this sort of evidence is this: How can a theory profess to be pertinent to a person's ability to speak and understand a language, if it is arrived at with little or no recourse to actual speakers, and is carried out as if human language were an entity separate from humans (Derwing 1980)? In order for an analysis to be able to profess, not only linguistic, but psychological relevance, it should include evidence which involves the speakers' actual use and manipulation of linguistic elements.

The role of experiments in the search for psychological realities

In summary, there are several reasons to be skeptical of the psychological significance of many phonological analyses. There are three ways in which

experimentally adduced evidence can aid in the search for psychological realities in phonology: 1) experiments provide empirical evidence; 2) experiments involve attempts to gain insight into the psyches of language speakers by more direct means; 3) experiments help determine which linguistic realities are psychologically pertinent and which are not.

Phonologists have had a great deal of success in discovering phonological structures, patterns, and generalizations which may be used to systematize human languages. These could be termed linguistic realities. Since these linguistic realities are based on human language, it is possible that they also have some relevance to the way people processes language. However, the fact that a phonological analysis yields a certain structure or pattern only demonstrates that it is available to be potentially known or used. In order to determine what is actually known or utilized by the speakers, the focus of the research must turn back towards the speakers themselves.

It is highly possible that a phonological structure proves useful in systematizing a given language, but that has no correlation in the minds of language speakers. A phonological structure may have arisen by chance, or may be the result of a diachronic process which has long since died. Certain alternations may be due to purely articulatory or aerodynamic influences, and in that case, would neither be

knowable nor psychological at all. Therefore, evidence from psycholinguistic experimentation helps to differentiate between those linguistic realities which are psychologically relevant, and those which are not.

The second advantage of experimental evidence is that it deals with events which take place through space and in time. A hypothesis which is stated in such a way that it is subject to experimental refutation is empirical. As mentioned before, there is a willingness to accept the reality of hypotheses that are supported by spatiotemporal evidence. By the same token, there is skepticism as to the reality of hypotheses for which there is no spatiotemporal evidence. It is for this reason that well designed experiments are thought to offer insight into human linguistic capacities. If one's goal is to study the variety of structures and patterns that can be found in languages, psychological experimentation is not essential. However, if one's goal is to study how humans produce and comprehend language, direct access to the speakers of the language is imperative.

Strong and weak reality

To this point, I have discussed the issue of psychological reality in phonological investigation without giving an explicit definition of what it means

for an analysis to have psychological significance. There are two senses in which an analysis may be considered pertinent to linguistic cognition. Cutler (1979:79) defines them in this way:

In the strong sense, the claim that a particular level of linguistic analysis X, or postulated process Y, is psychologically real implies that the ultimately correct psychological model of human language processing will include stages corresponding to X or mental operations corresponding to Y. The weak sense of the term implies only that language users can draw on knowledge of their language which is accurately captured by the linguistic generalization in question. (see also Steinberg 1975:218-20)

While science can never determine absolute truth or unearth "ultimately correct" models, Culter's observation is still extremely useful. The strong sense of reality implies a close relationship between the way an analysis works out on paper, and the internalized representations, and mental processes speakers possess. The weak sense involves little correspondence between formal constructs and psychological mechanisms.

In many respects, human language processing can be considered a virtual black box. Since the box cannot easily be opened for inspection, the major clues

to its contents come from the output it produces. The weak reality approach to the black box involves determining *in principle* what the box might contain. If an analysis produces the same output as the box does, then in a limited way, it may claim to have achieved psychological validity (Rischel 1978:442). The strong claim to reality, on the other hand, involves establishing *in fact* what the box contains.¹

It is often uncertain what level of psychological reality phonological analyses strive to attain. At times, they are spoken of as if the rules and representations they are comprised of were step-by-step formulae for assembling or producing forms. In this case, they would be candidates for attaining reality in the strong sense. However, rules are generally not defined in these terms. Bradley's statement (1980:38) is representative of this sentiment. She observes that 'grammars do not (and, moreover, are not intended to) dictate the ways in which the computations of speaking and listening proceed.' Chomsky and Halle (1968:117) made the same point three decades ago:

Although we may describe the grammar *G* as a system of processes and rules that apply in a certain order to relate sound and meaning, we are not entitled to take this as a description of the successive acts of a performance model.

If formal rules do not reflect the mental algorithms which are used in speech

and comprehension, the natural question is in what sense they are relevant to linguistic cognition. The claim is that rules somehow represent a speaker's tacit knowledge of the language--knowledge in the sense of being able to speak and comprehend the language. For example, in response to charges that phonological rules have no psychological validity Kiparsky (1975:198) states:

To deny that grammatical rules are utilized in speech behavior is not necessarily to deny their psychological reality. . . . In phonology, the system of rules and underlying forms might be a representation of the speaker's knowledge of the systematic relationships among words in the language; not in any sense a mechanism which is applied whenever words are spoken and heard.

If rules and representations are defined in these terms, they cannot ever attain psychological reality in the strong sense. This is because reality in the strong sense implies that rules correspond to mental operations and mechanisms. If phonological analyses are merely abstract representations of a speaker's ability to speak and understand the language, instead of mirroring actual processes, they may only attain reality in the weak sense. This means that what is potentially real in an analysis is not the rules, constraints, intermediate derivations, or rule orderings, but merely 'the function that these constructs serve to specify' (Matthews 1991:197), or in other

words, the 'content' of the rules (Rischel 1978:442). Any analysis is real in the weak sense, so long as it produces the same outcome as is produced by the speakers of the language, regardless of the way it goes about producing the output.

Misunderstandings about what level of reality linguistic analyses represent sometimes lead to confusion. For example, Derwing (1979:114) charges that

a grammar which describes utterance forms can no more 'explain' them than can a description of a painting tell how the painting came about--or a 'grammar of a cake' tell how to make a cake.

The major contention here is that a description of some phenomenon is by no means an explanation as well. Derwing feels that rules should explain language production and comprehension. Fromkin (1980:200) charges that this assumption is responsible for Derwing's misunderstanding of what rules represent; they are not thought to explain language behavior, but are merely an abstract representations of the ability to use language. Most phonological analyses do not claim to explain the actual processes involved in language perception and production, only to describe phonological structure, with the assumption that the structure is somehow relevant to actual processing.

In regards to Derwing's analogy between a linguistic analysis and painting a picture, it could be countered that a description of the structure, composition, and

use of color and light in a painting, does not explain how it was painted by the artist. However it does say something about what factors the artist had to have in mind while producing the painting. It is in this abstract way that analyses correspond to speakers' linguistic abilities. Of course, some linguists (e.g. Bresnan and Kaplan 1982:xxii) perceive the search for the actual cognitive processes which underlie language use as a much worthier goal.

In summary, many researchers are interested in analyses of language which possess some relevance to the human psyche. Confusion often arises because rules are often spoken of as if they represented actual steps in language processing. However, they are defined in other terms. They are defined, not as linguistic algorithms, but as a more generic type of knowledge which underlies speakers' ability to speak and comprehend their language. An analysis which posits algorithms may justifiably profess to achieve psychological reality in the strong sense of the word. However, an analysis which deals in abstract representations may only hope to be proven real in the weaker sense of the word. I will return to the notion of strong and weak reality as it pertains to experimental evidence in a later section.

Criticisms of psycholinguistic experiments

In the preceding pages, I have suggested that experimental evidence has advantages over the more commonly utilized types of evidence in resolving issues of psychological reality. The experimental approach, however, has not been free from criticism. I review several of these criticisms in this section.

Lack of adequate knowledge

One misconception about experiments is the idea that they are valuable only in a field which is well developed, and about which much is known. For example, Matthews (1991:190-1) is of the opinion that,

We know very little about the computational machinery involved in language processing. We are therefore not in a position to use experimental evidence regarding language processing.

Kac expresses similar sentiments. He concludes that since an adequate theory of linguistic structure has not been achieved, mingling psychology and linguistics is a wasteful and unfruitful endeavor (1974:45-6; 1980:243).

It is difficult to accept this line of reasoning. It is tantamount to refusing to perform experiments relating to subatomic particles, on the grounds that so little is known about them. In the same way, it would be absurd for a psychologist to

denounce the utility of psychological experiments simply because s/he felt that too little is known about perception, learning, or cognition. The truth is that much of what is known about physics and psychology is a direct result of theory-based experimentation, and could not have been established in any other way.

In the empirical sciences, theory building and experimentation are inseparably connected. Extensive knowledge, well-developed, and completely adequate theories are not prerequisites for experimental research. The only prerequisite is a hypothesis which is consistent with the bulk of the existing scientific knowledge, and which is stated in such a way as to specify what outcome would support the hypothesis and what outcome would disprove it (Bunge 1980:33). Once a hypothesis has been confirmed or refuted, the theory is then modified, which in turn leads to better hypotheses and more experimentation. Analysis which never goes beyond the theoretical stage are just as un insightful as data gathering and experimentation carried out with complete disregard to theoretical underpinnings. Clearly stated theory should naturally lead to, and precede empirical research. The role of descriptive analyses is that they provide insight into the structure of language upon which theories are built. In Kac's (1978:155) words:

The question how a language is organized (which is the same as the question of what its structure is) is a different one from that of how

a speaker comes to be able to use it--though in answering the first we contribute to some extent to the answer to the second since it is precisely the fact that languages have structure that renders them knowable and learnable in the first place.

In the case of linguistics, descriptive analyses are the forerunner of empirical investigation (Baker 1979:141; Derwing 1979:125; Kac 1980:243).

Experiments and competence

A common charge against experimental evidence is that it is not pertinent to the generative domain of enquiry. Kiparsky (1968:174) phrases it in these terms:

The fact that grammars are not performance models presumably means that the answer to the question of whether they are correct competence models is not likely to be forthcoming by any currently known experimental techniques until the contributions of competence can be separated out from the facts about performance.

In one respect, Kiparsky is absolutely correct. Competence is an idealized concept which comprises the system of rules which are thought to underlie a speaker's ability to produce and understand language; performance is the actual realization of the

speaker's linguistic ability (Chomsky 1980a:205). As long as this view is maintained, competence is effectively shielded from experimental probes and possible refutation. This is so because, according to this dichotomy, all spatiotemporal manifestations of language fall into the domain of performance, and as a result, can never be directly relevant to the study of abstract representations of linguistic ability which competence is thought to embody.

Suppose, for example, that two subjects react in a different manner to a question in a linguistic experiment. One could say that the subjects have the same underlying linguistic ability (competence), but that the experimental design was responsible for the differences. That is, something in the experiment kept the subjects' reactions from reflecting their competence (Wheeler 1980:78-90). Therefore, it could be argued that the experiment must have measured performance instead. Since generative theories are theories of competence, and by definition, the results of the any experiment are indicative only of performance, the results are irrelevant to a theory of competence.

Derwing (1983:66) demonstrates how a similar argument can effectively insulate a theory from any sort of counter evidence:

Suppose we find some child who is quite adept at basic arithmetic.

One possible hypothesis about the "competence" thought to underlie

this skill might be to attribute the child, not with something so mundane as a learned, laborious, step-by-step procedure for carrying out simple arithmetic operations, but rather with knowledge of number theory. And what if experimental results are found that seem to fly in the face of this hypothesis? Just chalk them up as "performance errors" and the well-formed theory remains inviolate.

The question of where competence stops and performance begins is also difficult, if not impossible to determine. Any time an utterance is made, whether in the laboratory or in spontaneous speech, it necessarily involves both competence as well as performance. That is, there is no such thing as 'pure competence' unfettered by performance factors (Stemberger 1994; Wheeler 1980:67; Zimmer 1969:320). Competence cannot be produced alone and uninfluenced by performance in time or space.

Until competence is defined in such a way that it is subject to possible falsification, any empirically testable hypothesis will ultimately be a hypothesis about an aspect of performance. Therefore, researchers whose principle concern is psychological reality should be content to relegate competence to the domain of nonempirical science along with logic, virtue, number theory, and ethics, and to

focus on the reality of entities that emerge through spatiotemporally observable performance.

Negative experimental results

Kiparsky (1975) voices another objection to the use of experiments. In this case, it has to do with the validity of negative experimental results. He accepts the results of production and perception tests, but only insofar as they produce evidence in favor of the psychological significance of an analysis. According to him, negative evidence is always inconclusive. His justification for this reasoning is this: If an archeologist finds bones at an archeological dig, the existence of those bones provides positive evidence that a certain animal inhabited that area. However, if no bones are found, that is not evidence that the animal in question never lived there.

This illustration is objectionable on two counts. In the case of bones found at an archeological dig, Kiparsky's logic is sound. However, this does not imply that negative evidence is irrelevant to the testing of a hypothesis in all fields of research. For example, the claim that cold nuclear fusion had been produced in the laboratory was corroborated by some experiments, and refuted by others. In the end, the claim

was rejected on the basis of the negative evidence presented.

Second, Kiparsky's argument cannot be logically extended to physics or to phonology. This is because archaeologists look for evidence for things that used to exist in a given place. On the other hand, physicists and mentalistic phonologists seek evidence about things which theoretically exist in the present. Fortunately, not all phonologists are willing to dismiss negative evidence so quickly. For example, Mohanan (1986:58-9) suggests that if there is abundant evidence that a rule does not play a part in storage, recognition, or production, it should not enter into a description which professes psychological significance.

Experiments and external validity

Another criticism of psycholinguistic experiments is what Kiparsky and Menn term the 'strangeness effect' (1977:63-4). This suggests that the unusual circumstances which are involved in obtaining experimental evidence, influence the subjects to give unusual responses. In other words, the experimental situation causes the subjects to answer in ways they would not under normal circumstances. The strangeness effect is known in science as threats to the external validity of an experiment.

External validity involves the extent to which experimental results can be considered valid outside of the experimental setting. Experimental results are externally valid if they are applicable to objects (or people) other than those which are tested in the experiment. Externally valid results are thought to be indicative of how objects (or people) react outside of the experimental situation.

Any experiment must try to control for external validity. One way to do this is to compare the results of experiments performed in the laboratory with results obtained under more natural circumstances. External validity may be checked in another way also. Ideally, a hypothesis about psycholinguistic functions should be tested with a number of distinct experimental paradigms. In this way, influences which are attributable to the experimental situation may be factored out. If several different experiments yield similar results, the results are more likely to be due to actual mental processes, and not to something inherent in the experiment (Derwing 1979:126-7).

Conflicting experimental results

The fact that experiments often produce conflicting results could also be adduced by a skeptic to argue that experiments are ineffective in deciding questions

of psychological reality. This position is clearly untenable. Conflicting experimental evidence is a reality for all fields of science which incorporate experiments. The mere existence of conflicting evidence is not grounds for abandoning the experimental approach. On the contrary, it should compel researchers to refine their experimental methods, as well as to explore others. It should lead to closer inspection of the phenomenon under investigation, as well as the experimental means used to probe it.

Experiments and strong reality

Ultimately, the claim that an analysis is real in the strong sense means that it relates to actual mental mechanisms. Since little is currently known about the workings of the brain at this level, it may seem impossible for any analysis to achieve strong reality. However, this is not the case. Consider the work on the processing of regular and irregular past tense verbs in English.

Two theories exist. One suggests that different mechanisms are responsible for regular and irregular inflection (Marcus et al. 1993; Pinker 1991; Pinker and Prince 1988, 1994; Prasada and Pinker 1993). According to competing hypotheses, regular and irregular past tense forms are produced by the same mechanism (Bybee

1985, 1988, 1995; Daugherty and Seidenberg 1992, 1994; Seidenberg 1992; Stemberger 1994). These hypotheses lend themselves to possible empirical refutation, which makes them viable candidates for strong reality. If one area of the brain is activated when processing regular forms, and another area is activated when processing irregular forms, that would constitute evidence that regular and irregular forms utilize different mental mechanisms. This is actually the goal of several recent studies (Jaeger et al. 1996; Ullman in press).

Experiments, strong reality, and competing analyses

As already discussed, most phonological analyses do not purport to represent mental mechanisms, which prevents them from attaining strong reality. If one were to choose a given phonological phenomenon, it would be a simple task to find a number of analyses, carried out within different frameworks, that claim to account for the phenomenon in question. In many cases, each of the analyses would produce the same outcome, but they each assume differing mechanisms in order to yield the output. Often, each successive analysis is regarded as superior to its predecessors. The question that is raised, is whether experimentation can help decide between

such competing analyses.

Most psycholinguistic evidence involves this assumption: A theory postulates certain mental mechanisms. Therefore, if the results of an experiment correspond in a statistically significant way to the hypothesized mechanisms, that constitutes evidence in favor of the reality of those mechanisms. Botha (1971:128-30) challenges the validity of this assumption:

The fact that a theory correctly PREDICTS some events does not necessarily imply that it also correctly DESCRIBES or REPRESENTS the structure of the mechanism from the operation of which the predicted events result. The predictive and representative or descriptive functions of a theory are distinct, and achievement of success in one of them does not necessarily imply that success has been achieved in the other as well.

In other words, if the human mind is viewed as a black box, any number of different mechanisms may be responsible for producing the output of the box. The fact that experimental evidence supports one hypothetical mechanism is no guarantee that a distinct mechanism is not actually responsible for the output (Botha 1971:131-5). Therefore, if two distinct mechanisms are postulated which would produce the same output, psycholinguistic evidence would be incapable of determining which

mechanism is real (Fromkin 1975:56).

Botha and Fromkin are correct in this regard. As already discussed, a phonological analysis that does not claim to mirror actual mental processes, has the potential of relating to psychological mechanisms only in the weak sense. In this sort of analysis, only its substance may be proven to have some significance for speakers. Consequently, any and all analyses which make the same predictions are significant in the weak sense, if those predictions are borne out.

For many phonologists, it is frustrating that psycholinguistic experimentation is unable to decide between competing analyses of the same phenomenon. Much of the contemporary work in phonology centers on demonstrating the superiority of one analysis of a given process over another. Therefore, it appears that if experimentation is incapable of settling the issue of which analysis is most correct, it is of little value. An example from Spanish will clarify this point.

Spanish diphthongization and strong reality

The alternation between the unstressed mid-vowels [o, e] and the stressed diphthongs [jé, wé], has been accounted for in several different analyses. A crucial part of these analyses is that they distinguish between unstressed mid-vowels which

alternate with stressed diphthongs (e.g. *c[o]ntámos* 'we count,' *c[wé]ntan* 'they count') and those which do not (e.g. *t[o]sémos* 'we cough,' *t[ó]sen* 'they cough').

A number of different ways have of distinguishing diphthongizing from non-diphthongizing mid-vowels have been proposed. Harris (1969:74-75; 1977) claims that the mid-vowels which do not undergo diphthongization are represented underlyingly as /o/ and /e/. Those which are transformed into diphthongs appear in the deep structure with a diacritic feature [D]. St. Clair (1971:421) utilizes a tense-lax distinction in his analysis. The lax vowels which he transcribes as /O, E/, become diphthongs, while the tense vowels, /o/ and /e/, do not. Hooper's analysis (1976:157-60) is different in that it does not transform vowels into diphthongs. Instead, in stem morphemes which demonstrate the alternation, both the mid-vowel and the diphthong are listed disjunctively. For example the verb *contar* 'to count' is represented as in (1):

(1) o
 /k^owe At-/

As a result, only mid-vowels which are disjunctively listed with a diphthong alternate. In a later analysis, Harris (1985:31) suggests that the diacritic that marks mid-vowels which undergo diphthongization is part of the syllable structure.² According to this view, only mid-vowels which are adjacent to an empty prosodic slot trigger the formation of diphthongs.

The four different formal representations of the stems *cont-* 'to count' and *tos-* 'to cough' are listed in (2) for comparison:

(2)	(i)	Harris 1969, 1977	/k ^o nt /	/t ^o s /
	(ii)	St. Clair	/k ^o nt-/	/tos-/
	(iii)	Hooper	/k ^o nt /	/tos-/
	(iv)	Harris 1985	/k o n t-/	/t o s-/
			* * * * *	* * *
			x x x x x	x x x

Each of these analyses express the generalization that some, but not all mid-vowels, alternate with diphthongs, and each effectively distinguishes diphthongizing from non-diphthongizing mid-vowels. In other words, they each embody the substance of the alternation. At this level, the claim that diphthongization is psychologically valid means that it is in some way relevant to the language faculties of Spanish speakers. This corresponds to psychological reality in the weak sense.

Proving that one of these analyses is more psychologically valid than the others would involve more than confirming the reality of the substance of the alternation. It would include confirming the reality of the actual form of the representation used in the analysis. Although the substance of each analysis is the same, (i.e. some mid-vowels alternate with diphthongs), the form of each one is

distinct. For instance, in Harris' earlier analyses the diacritic [D] is crucial, while in St. Clair's the abstract feature [lax] is. At this level, any strong claim to the reality of an analysis is necessarily a claim that the actual form of the representation has a correlate in speakers' minds. Therefore, an attempt to prove that one analysis is true and another false would necessarily entail demonstrating the psychological reality of notational elements such as empty prosodic slots, or features such as [D] or [lax].

Concepts and strong and weak reality

The notational elements of phonological investigation are conceptual elements on a par with the notational elements of logic and mathematics. Concepts are created fictions that are defined and granted existence by those who use them. This means that they may be useful within the domain they have been created, but unlike physical elements, they do not exist outside of the conceptual domain (Bunge 1980).

A number of linguists have related linguistic representations to maps (Harmon 1980:21-2; Matthews 1991:196). This is a fruitful analogy. Suppose that, in addition to latitude and longitude, other notational devices have been invented which also locate topical features on the Earth. It could be argued that one device

does the job better, or is more elegant or precise than another. However, it would be absurd to assert that one is real while the others are not. What is real are not the devices used to locate islands, oceans and rivers. The islands, oceans and rivers themselves are real. The notations are simply extremely useful tools. So it is with the theoretical notations of phonological analyses. They are convenient fictions.

The fact that notational elements belong to the conceptual realm means that they are not subject to empirical test. In other words, it would be impossible to discover empirical evidence which would support the psychological reality of one of the analyses of Spanish diphthongization in (2) over another. What spatiotemporal event would argue in favor of the feature [D] over the disjunctive ordering of vowels and diphthongs? Would it be possible to formulate a hypothesis about the existence of empty prosodic slots in such a way that it would be clear what evidence would disprove or support them? (e.g. If we find X, we know that diphthongizing vowels appear next to an empty prosodic slot.) Any attempt along these lines would be futile.

Botha and Fromkin are correct in regarding experiments as an unsatisfactory means of distinguishing between competing analyses which are notational variants. This is true because empirical data is relevant only to empirically confirmable entities. Since the notational elements of phonological analyses belong to the

conceptual realm, they simply cannot be proven by empirical means such as experimentation. Therefore, the charge that theoretical problems cannot be solved via experimentation are essentially complaints that experiments cannot resolve nonempirical questions.

In phonology, the majority of the evidence presented to demonstrate the superiority of one analysis over another is nonempirical, (appeals to universal grammar, simplicity, elegance, generality, etc). Therefore, one analysis may indeed be proven superior to another in the conceptual realm. However, it does not follow that the superior analysis more closely models the mental processing of language which exists in the empirical-physical world. Unfortunately, phonologists often make the mistake of confusing the empirical and conceptual domains. Katz (1985:193) expresses it in this manner:

No one confuses psychological theories of how people make inferences with the logical theories of implication, or psychological theories of how people perform arithmetical calculations with mathematical theories of numbers. Yet, in the exact parallel case of linguistics, conceptualists do not make the distinction, conflating a psychological theory of how people speak and understand speech with a theory of the language itself.

Can experimental evidence decide between competing analyses of the same phenomenon? Can it be used to answer questions about the mental processing of phonology? Yes, but what is required is that the analysis be based on an empirically testable theory of actual language processing, and not merely a conceptual theory of language structure.

The reality of an analysis in the weak sense may be verified or refuted experimentally. This entails determining whether the substance which the analysis captures is psychologically significant. In this case, if experimental results yield an output similar to those predicted by a hypothesized mechanism, then the existence of some mechanism is supported. The results do not indicate what the form of the mechanism is, only that speakers appear to have, or not to have, internalized some mechanism which generates the output (Cutler 1979:79; Steinberg 1975:218-20). In other words, experiments are effective in determining which analyses have mental significance, and which do not. Although it is more difficult, it is also possible to verify the strong reality of an analysis. The studies on the processing of the English past tense which were cited earlier demonstrate this possibility.

Examples of Experimentally Acquired Evidence

English *ng* and Lexical Phonology

Nonce word experiments have the ability to answer some questions about the psychological validity of grammars. For example, Schlenck (1988) carried out a nonce word experiment in English to test a hypothesis of Lexical Phonology. In Lexical Phonology, it is hypothesized that certain phonological processes are related to certain groups of affixes. In English, the letters *ng* are pronounced either [ŋ] or [ŋ̥]. With a few exceptions, [ŋ] appears before the suffixes *-ation*, *-ize*, and *-er*. These are called Class I affixes. On the other hand, before Class II affixes such as *-ing*, *-ly*, and *-less*, *ng* is pronounced [ŋ̥]. This observation is founded on data such as in (5):

(5)	(i) Class I aff.	<i>-ation</i>	<i>prolongation</i>	[ŋ]
		<i>-er</i>	<i>stronger</i>	[ŋ]
	(ii) Class II aff.	<i>-ing</i>	<i>prolonging</i>	[ŋ̥]
		<i>-ly</i>	<i>strongly</i>	[ŋ̥]

The empirical question is whether this analysis merely describes a linguistic reality in English, or whether it also represents a psychological reality as well.

In Schlenck's study, subjects were recorded as they read a fairy tale which included a number of nonce words containing *ng* followed by different affixes.

Afterwards, the nonce words were transcribed phonetically. The predictions inherent in Lexical Phonology were borne out. Before Class I affixes [ɔ̃] appeared significantly more often than [ɔ̃]. Before Class II affixes, [ɔ̃] appeared significantly more often than [ɔ̃]. The subjects' ability to use [ɔ̃] or [ɔ̃] in the appropriate context, implies that they have some sort of knowledge of the distributional pattern of these allophones in English and put that knowledge to use.

The English vowel shift

Perhaps one of the most experimentally tested rules in English is the English vowel shift. The vowel shift rule is designed to account for the following vocalic alternations (Chomsky and Halle 1968:50-5):

- (6) (i) [aj]- [I] *divine-divinity*
 (ii) [ij]- [ɛ] *serene-serenity*
 (iii) [ej]- [æ] *sane-sanity*
 (iv) [aw]- [ʌ] *profound-profundity*
 (v) [uw]- [a] *lose-lost*
 (vi) [ow]- [a] *verbose-verbosity*

In 1977, Halle reformulated the rule so that it incorporated another vocalic

alternation (614-18):

(vii) [uw]- [ʌ] *reduce-reduction*

However, this reformulation of the rule excluded the alternation between [uw] and [a]. In 1985, Halle and Mohanan added another alternation to the vowel shift rule (72-9):

(viii) [ej]- [ʌ] *destroy-destruction*

Each pair of vowels is thought to be derived from a single underlying long vowel which never surfaces. For example, the alternation between [ij] and [ɛ] is thought to stem from an underlying /e:/. Although these alternations arose from historical processes, they are asserted to constitute part of the internalized mental grammar of Modern English speakers.

A great deal of experimentation has been done in order to investigate the psychological status of the vowel shift rule (see Jaeger 1986 and Wang and Derwing 1986 for summaries). The results of these experiments have led to a greater understanding of the vowel shift, as well as to better understanding of how to tap speakers' linguistic knowledge experimentally.

Experimentation into the vowel shift has helped to determine what kinds of experiments provide insight into the mind, and which do not. One of the earliest experimental techniques which was applied to the vowel shift question were

morpheme combination experiments. In this type of experiment, the subjects are given a word and a suffix and then asked to combine them to construct a new word (Myerson 1976; Ohala 1974; Steinberg and Krohn 1975). For instance, subjects are asked to combine *maze* and the suffix *-ic* to form the word *mazic*. The outcome which would be predicted by the vowel shift rule is [mejz] | [mæzlk]. However, the most common way subjects handled these questions was to leave the vowel unchanged (i.e. [mejz] | [mejlk]). These negative results were at first regarded as evidence against the psychological reality of the vowel shift rule.

Although morpheme combination experiments do not support the vowel shift rule, other experimental methods have provided some positive evidence. Evidence in favor of certain aspects of the vowel shift rule was later found through a variety of different experimental procedures. Limited positive evidence was found by means of preference experiments, learning experiments, memory experiments, and concept formation experiments (Jaeger 1986:88-90).

In other words, morpheme combination experiments produced negative results, while many other methods yielded positive ones. This finding reveals less about the vowel shift than it does about the validity of morpheme combination experiments. It seems that morpheme combination experiments are incapable of tapping into psycholinguistic abilities, while other methods are. The realization

that morpheme combination experiments are poor experiments, at least in the case of the English vowel shift, came about as a direct result of conducting further psycholinguistic research into the vowel shift question, even when the initial experiments did not prove fruitful. What this demonstrates is that experimental inquiries into a subject not only provide a better understanding of the subject itself, they also provide a better understanding of experimental techniques. It is in this way that the experimental approach is able to hone and refine its methods.

Experiments into the vowel shift have also indicated the degree to which the rule is psychologically valid. The most significant finding is that not all of the eight proposed vowel shift alternations are supported. It appears that only five of the eight alternations have psychological significance for linguistically naive English speakers. The bulk of the data supports the validity of the alternations [aj]- [I], [ij]- [ε], [ej]- [æ], [ow]- [a], and [uw]- [ʌ], while the alternations [aw]- [ʌ], [uw]- [a], and [ɨj]- [ʌ] were not found to be significant for the test subjects (Jaeger 1984, 1986; Wang and Derwing 1986, 1994). This means that none of the proposed vowel shift rules in (6) correctly groups the psychologically significant alternations together, while at the same time excluding the insignificant ones. Therefore, the psychological validity of any of the proposed vowels shift rules, as they are currently formulated, is dubious.

As is often the case, the answer to one question provokes the formulation of another one. In the case of the vowel shift, the question which arises is what the five psychologically valid alternations have in common. Jaeger (1984, 1986), Moskowitz (1973), and Wang and Derwing (1986, 1994) suggest that the five significant alternations correspond to spelling rules. When children are taught to read English they are told that each of the five vowels have a short and a long sound. The five significant vowel shift alternations correspond exactly to the short and long varieties of the five written vowels:

(7)	<u>i</u>	[aj]- [I]
	<u>e</u>	[ij]- [ε]
	<u>a</u>	[ej]- [æ]
	<u>o</u>	[ow]- [a]
	<u>u</u>	[uw]- [ʌ]

Of course, this claim constitutes a hypothesis which has yet to be tested. It would be possible to replicate the tests using illiterate English speakers as subjects. In this way, negative results to the experiment would corroborate the spelling rule hypothesis. A positive outcome, on the other hand, would suggest that the vowel shift alternations are not rooted in spelling rules.

Conclusions

The purpose for citing these experiments has been to demonstrate the utility of experiments in shedding light on theoretical issues. As a result of these experiments, much more is known about the psychological role certain phonological alternations have for linguistically naive speakers. What is essential to note is that this knowledge could not have resulted from a nonempirical study of language internal data.

Of course, experiments do have their limitations. More specifically, they may be not used to decide between competing conceptual analyses which are essentially notational variants of each other. However, they may provide clues about the mentalistic import of conceptual analyses in the weak sense discussed. Proving the reality of an analysis in the strong sense is also possible but much more difficult. It also requires a hypothesis that lends itself to confirmation on the basis of spatiotemporal evidence.

The results of psycholinguistic experiments should always be evaluated carefully. Definitive conclusions about the mental capabilities of language speakers should not depend solely on the outcome of one experiment. The results of any one study are always tentative and inconclusive. Experiments must be replicated, and

this will often yield conflicting results.³ For this reason, judgement should be suspended until a large body of evidence from varying sources has been accumulated.

1. The distinction between adequacy in fact and adequacy in principle has been borrowed from Fought (1973:157).
2. García-Bellido offers a similar analysis (1986).
3. Bertinetto gives an example of a psycholinguistic experiment which produced different results when it was replicated (1992).

References

- Baker, William J. 1979. On draining the conceptual swamp. *Perspectives in experimental linguistics*, ed. by Gary D. Prideaux, 141-149. Amsterdam: John Benjamins.
- Botha, Rudolf P. 1971. *Methodological aspects of transformational generative phonology*. The Hague: Mouton.
- _____. 1973. *The Justification of linguistic hypotheses*. The Hague: Mouton.
- Bradley, Diane. 1980. Lexical representation of derivational relation. *Juncture*, ed. by Mark Aronoff and Mary-Louis Keaton, 37-55. Saratoga, Cal.: Anma Libri.
- Bresnan, Joan, and Ronald M. Kaplan. 1982. Introduction: Grammars as mental representations of language. *The mental representation of grammatical relations*, ed. by Joan Bresnan, xvii-lix. Cambridge, Mass.: MIT Press.
- Bunge, Mario. 1980. *Epistemología*. Barcelona: Ariel.
- Bybee, Joan. 1985. *Morphology*. Amsterdam, Holland: John Benjamins.
- _____. 1988. Morphology as lexical organization. *Theoretical approaches to morphology*, ed. by Michael Hammond and Michael Noonan, 119-141. San Diego, CA: Academic Press.
- _____. 1995. Regular morphology and the lexicon. *Language and Cognitive*

Processes 10.425-455.

Cena, R. M. 1978. When is a phonological generalization psychologically real?

Bloomington, Ind.: Indiana University Linguistics Club.

Chomsky, Noam. 1980a. Rules and representations. New York: Columbia

University Press.

_____. 1980b. Rules and representations. *The Behavioral and Brain Sciences*

3.1-15, 42-61.

Chomsky Noam, and Morris Halle. 1968. *The sound pattern of English*. New

York: Harper and Row.

Cutler, Anne. 1979. The psychological reality of word formation and lexical stress

rules. *Proceedings of The Ninth International Congress of Phonetic*

Sciences, vol 2, ed. by Eli Fischer-Jørgensen, Jørgen Rischel, and Nina

Thorsen, 79-85. Copenhagen: University of Copenhagen.

Daugherty, Kim, and Mark S. Seidenberg. 1992. Rules or connections? The past

tense revised. *Proceedings of the fourteenth annual conference of the*

Cognitive Science Society, 259-264. Hillsdale, N.J.: Erlbaum.

_____. 1994. Beyond rules and exceptions: A connectionist approach to

inflectional morphology, *The reality of linguistic rules*, ed. by Susan. D.

Lima, Roberta. L. Corrigan, and Gregory. K. Iverson, 353-388. Amsterdam,

Holland: John Benjamins.

Derwing, Bruce L. 1973. Transformational grammar as a theory of language acquisition. London: Cambridge University Press.

_____. 1979. Psycholinguistic evidence and linguistic theory. *Perspectives in experimental linguistics*, ed. by Gary D. Prideaux, 113-38. Amsterdam: John Benjamins.

_____. 1980. Against autonomous linguistics. *Evidence and argumentation in linguistics*, ed. by Thomas A. Perry, 163-89. Berlin: Walter de Gruyter.

_____. 1983. On drawing the right conclusions in psycholinguistics: Some critical areas of control in experimental and analytical practice. *Proceedings of the ninth annual meeting of the Berkeley Linguistics Society*, ed. by Amy Dahlstrom et al, 64-73. Berkeley: Berkeley Linguistics Society.

Derwing, B. L., G. D. Prideaux, and W. J. Baker. 1980. Experimental linguistics in historical perspective. *Experimental linguistics*, ed. by Gary D. Prideaux, Bruce L. Derwing, and William J. Baker, 1-13. Ghent, Belgium: E. Story-Scientia.

Derwing, Bruce L., and William J. Baker. 1977. The psychological basis for morphological rules. *Language learning and thought*, ed. by John Macnamara, 85-110. New York: Academic Press.

- Fought, John. 1973. On the adequacy of recent phonological theories and practices. *Annual Review of Anthropology* 2.153-164.
- Fromkin, Victoria A. 1975. When does a test test a hypothesis, or, what counts as evidence? *Testing linguistic hypotheses*, ed. by David Cohen and Jessica R. Wirth, 43-64. Washington D. C.: Hemisphere.
- _____. 1980. The psychological reality of phonological descriptions. *Proceedings of the Ninth International Congress of Phonetic Sciences*, vol 3, ed. by Eli Fischer-Jørgensen, and Nina Thorsen, 195-202. Copenhagen: University of Copenhagen.
- Goyvaerts, Didier L. 1978. From triumphant euphoria to bitter disappointment? *Communication and Cognition* 11.287-307.
- Hall, Robert A. Jr. 1987. *Linguistics and pseudo-linguistics*. Amsterdam: John Benjamins; Reprinted from 1981/82. *Review of Newmeyer, Linguistic theory in America*. *Forum Linguisticum* 6.177-88; and 1973. *Journal of English Linguistics* 7.21-42.
- Halle, Morris. 1977. Tenseness, vowel shift, and the phonology of the back vowels in Modern English. *Linguistic Inquiry* 8.611-625.
- Halle, Morris, and K. P. Mohanan. 1985. Segmental phonology of Modern English. *Linguistic Inquiry* 16.57-116.

- Harmon, Gilbert. 1980. Two quibbles about analyticity and psychological reality. *The Behavioral and Brain Sciences* 3.21-22.
- Harris, James W. 1969. *Spanish Phonology*. Cambridge, Mass.: MIT Press.
- _____. 1977. Remarks on diphthongization in Spanish. *Lingua* 41.261-305.
- _____. 1978. Two theories of nonautomatic morphophonological alternations: Evidence from Spanish. *Language* 54.41-60.
- _____. 1985. Spanish diphthongization and stress: A paradox resolved. *Phonology Yearbook* 2.31-45.
- Hooper, Joan B. 1976. *An introduction to natural generative phonology*. New York: Academic Press.
- Itkonen, Esa. 1976. *Linguistics and empiricalness: Answers to criticisms*. Helsinki: University of Helsinki.
- _____. 1978a. *Grammatical theory and metascience*. Amsterdam: John Benjamins.
- _____. 1978b. Linguistics: Nonempirical and empirical. *Proceedings of the 12th International Congress of Linguists*, ed. by Wolfgang U. Dressler and Wolfgang Meid, 157-8. Innsbruck: University Innsbruck.
- _____. 1983. *Causality in linguistic theory*. London: Croom Helm.
- Jeager, Jeri J. 1980. Testing the psychological reality of phonemes. *Language and*

Speech 23.233-253.

_____. 1984. Assessing the psychological status of the vowel shift rule. *Journal of Psycholinguistic Research* 13.13-36.

_____. 1986. On the acquisition of abstract representations for English vowels. *Phonology Yearbook* 3.71-97.

Jeager, Jeri J., Alan H. Lockwood, David L. Kemmerer, Robert D. Van Valin Jr., Brian W. Murphy, and Hanif G. Khalak. 1996. A positron emission tomographic study of regular and irregular verb morphology in English. *Language* 72.451-497.

Kac, Michael. 1974. Autonomous linguistics and psycholinguistics. *Minnesota Working Papers in Linguistics and Philosophy of Language* 2.42-7.

_____. 1978. Corepresentation of grammatical structure. Minneapolis and London: University of Minnesota Press and Croom Helm.

_____. 1980. In defense of autonomous linguistics. *Linguistics* 50.243-248.

Katz, Jerrold J. 1981. *Language and other abstract objects*. Totowa, New Jersey: Rowman and Littlefield.

_____. 1985. An outline of platonist grammar. *The Philosophy of Linguistics*, ed. by Jerrold J. Katz, 172-203. Oxford: Oxford University Press.

Katz, Jerrold J., and Paul M. Postal. 1991. Realism vs. conceptualism in linguistics.

Linguistics and Philosophy 14.515-554.

Kiparsky, Paul. 1968. Linguistic universals and linguistic change. *Universals in linguistic theory*, ed. by Emmon Bach and Richard Harms, 171-204. New York: Holt Rinehart and Winston.

_____. 1975. What are phonological theories about? *Testing linguistic hypotheses*, ed. by David Cohen and Jessica R. Wirth, 187-209. Washington D. C.: Hemisphere.

Kiparsky, Paul, and Lise Menn. 1977. On the acquisition of phonology. *Language learning and thought*, ed. by John Macnamara, 47-78. New York: Academic Press.

Lass, Roger. 1976. On generative taxonomy and whether formalisms 'explain'. *Studia Linguistica* 30.139-154.

Marcus, Gary F., Ursula Brinkmann, Harald Clahsen, Richard Wiese, Andreas Woest, and Steven Pinker. 1993. *German inflection: The exception that proves the rule*. Occasional Paper No. 47. Center for Cognitive Science, MIT.

Matthews, Robert J. 1991. Psychological reality of grammars. *The chomskyan turn*, ed. by Asa Kasher, 182-99. Cambridge, Mass.: Basil Blackwell.

Mohanan, K. P. 1986. *The theory of lexical phonology*. Dordrecht: D. Reidel.

- Morin, Yves-Charles. 1988. Morphological conditioning in phonologically transparent processes: Evidence from the evolution of vowel reduction in Vinzelles Occitan. *Canadian Journal of Linguistics* 33.431-42.
- Moskowitz, Breyne Arlene. 1973. Cognitive development and the acquisition of language, ed. by Timothy E. Moore, 223-260. New York: Academic Press.
- Myerson, Rosemarie Farkas. 1976. Children's knowledge of selected aspects of 'Sound Pattern of English.' *Recent advances in the psychology of language*, ed. by Robin N. Campbell and Philip T. Smith, 377-420. New York: Plenum Press.
- Ohala, John J. 1974. Experimental historical phonology. *Historical linguistics II*, ed. by J. M. Anderson, and C. Jones, 353-389. Amsterdam: North Holland.
- Pinker, Steven. 1991. Rules of language. *Science* 253.530.534.
- Pinker, Steven and Alan Prince. 1988. On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition* 28.73-193.
- _____. 1994. Regular and irregular morphology and the psychological status of rules of grammar. *The reality of linguistic rules*, ed. by Susan D. Lima, Roberta L. Corrigan, and Gregory K. Iverson, 321-51. Amsterdam: Benjamins.

- Popper, Karl R. 1968. *The logic of scientific discovery*. 2nd ed. New York: Harper and Row; original English edition 1959, London: Hutchinson.
- Prasada, Sandeep and Steven Pinker. 1993. Generalisation of regular and irregular morphological patterns. *Language and Cognitive Processes* 8.1-56.
- Ringen, Jon D. 1975. Linguistic facts: A study of the empirical scientific status of transformational generative grammars. *Testing linguistic hypotheses*, ed. by David Cohen, and Jessica R. Wirth, 1-42. Washington: Hemisphere.
- Rischel, Jørgen. 1978. Some general remarks in realism in current phonological work. *Papers from The Fourth Scandinavian Conference of Linguistics*, ed. by Kirsten Gregersen, Hans Basbøll, and Jacob Mey, 419-31. Odense: Odense University Press.
- Schlenck, Klaus-Jürgen. 1988. The analysis of the English velar nasal in lexical phonology: An experimental study. *Journal of English Linguistics* 21.29-38.
- Seidenberg, Mark S. 1992. Connectionism without tears. *Connectionism: Theory and practice*, ed. by Steven Davis, 84-122. New York: Oxford UP.
- Skousen, Royal. 1989. *Analogical modeling of language*. Dordrecht: Kluwer Academic Publishers.
- St. Clair, Robert. 1971. Diphthongization in Spanish. *Papers in Linguistics* 4.421-432.

- Steinberg, Danny D. 1975. Chomsky: From formalism to mentalism and psychological invalidity. *Glossa* 9.218-52.
- Steinberg, Danny D., and Robert K. Krohn. 1975. The psychological validity of Chomsky and Halle's vowel shift rule. *The transformational-generative paradigm*, ed. by E. F. K. Koerner 233-59. Amsterdam: John Benjamins.
- Stemberger, Joseph Paul. 1994. Rule-less morphology at the phonology-lexicon interface. *The reality of linguistic rules*, eds. Susan D. Lima, Roberta L. Corrigan, and Gregory K. Iverson, 147-169. Amsterdam: Benjamins.
- Ullman, Michael. In press. Evidence that lexical memory is part of the temporal lobe declarative memory, and that grammatical rules are processed by the frontal/basal-ganglia procedural system. *Brain and Language*.
- Wang, H. Samuel, and Bruce L. Derwing. 1986. More on English vowel shift: The back vowel question. *Phonology Yearbook* 3.99-116.
- _____. 1994. Some vowel schemas in three English morphological classes: Experimental evidence. In honor of William S-Y. Wang: *Interdisciplinary studies on language and language change*, 561-575. Taipei: Pyramid Press.
- Wheeler, Cathy J. 1980. On the relationship between phonology and psychology. *Papers in Linguistics* 13.51-100.
- Yngve, Victor H. 1986. *Linguistics as a Science*. Bloomington and Indianapolis:

Indiana University Press.

_____. 1996. *From Grammar to Science: New Foundations for General Linguistics*. Amsterdam/Philadelphia: John Benjamins.

Zimmer, Karl E. 1969. Psychological correlates of some Turkish morpheme structure conditions. *Language* 45.309-321.