Introduction

JOHN J. OHALA
JERI J. JAEGER

WHY EXPERIMENTS?

The evolution of a scientific discipline is probably similar to that of human economic evolution. Originally the food and clothing needed for survival were obtained by hunting and gathering; later it was discovered that these necessities could be obtained in greater abundance—and virtually on demand—by raising livestock and cultivating plants. That is, rather than rely on chance opportunities for obtaining sustenance, humans found that they could create and control these opportunities themselves. Likewise, one can generally identify a "hunting and gathering" stage in the history of most scientific disciplines, where the evidence needed to sustain a hypothesis is obtained by chance observations. Gradually the discipline's practitioners discover how to create and control the opportunities they need in order to make those crucial observations, that is, they learn how to conduct experiments.¹

We would like to believe that phonology is on the verge of developing into an experimental discipline. We could be wrong. There have been many prior calls for phonologists to turn to experimental studies of speech behavior, and in some isolated cases this call has been answered by practice (e.g., Osthoff & Brugmann 1878/1967; Rousselot 1891; Thumb & Marbe 1901; Ohala (1974) differentiates between 'nature-made experiments' and 'man-made experiments.' For the purposes of this introduction and the sense intended by the title of this book, the term 'experiment' used here refers to the latter.

¹Ohala (1974) differentiates between 'nature-made experiments' and 'man-made experiments.' For the purposes of this introduction and the sense intended by the title of this book, the term 'experiment' used here refers to the latter.
Verner 1913; Esper 1925; Grammont 1933; Haden 1938; Fry 1960; Lisker, Cooper, & Liberman 1962), but these have not triggered a self-sustaining tradition of experimental phonological studies.

Physiology has a similar history. Although there are isolated examples of physiological questions being answered or explored experimentally before the nineteenth century (e.g., by Harvey, Haller, Priestley), it was not until the mid-nineteenth century that the experimental approach was firmly established in the field through the efforts of Claude Bernard and those of like mind (e.g., Helmholtz, Pasteur, Purkinje). To a large extent their struggle was to demonstrate the unity of all scientific endeavor, that is, that there is fundamentally only one method in science. This method does not belong exclusively to physics, which was the first discipline in history to firmly establish what is known as the scientific method; rather, there is no bar to the application of knowledge derived from other disciplines to questions in one's own field. (This view should not be confused with naïve reductionism, the notion that the ultimate understanding of all aspects of the universe requires that they be interpreted or reduced in terms of some common set of fundamental physical entities, for example, subatomic particles. All that can be asked is to allow such interpretation where it enhances our understanding.)

The experimental method is based on the recognition that our knowledge of the world is subject to many distortions, in other words, that the world is not as it may seem. We must take special pains, then, to make as carefully as we can the observations that our knowledge is based on; we must refine these observations and structure them in such a way that we can eliminate or attenuate the factors which might distort them or render them ambiguous (with respect to the light they shed on the object of our study). An experiment, then, is simply the creation—contrivance, if one prefers—of a situation in which crucial observations, those relevant to a given question, may be made in such a way that they will be free from as many anticipated distorting influences as possible.

The primary purpose of experimentation is not to create knowledge, although by chance it often happens that completely new, unexpected things are observed during experimentation. It is, rather, a way of refining our knowledge. Following Popper (1959), one might even say that in a sense experiments actually destroy knowledge; at least they help to show which of our beliefs about the workings of the world do not agree with observations and hence should be discarded.

There is a popular misconception that experimentation always involves instrumentation or complex procedures of some sort—at the very least, statistical analysis. This, of course, is not true. The research reported in this volume in the papers by Hombert and Campbell, for example, which involve word games, illustrates this. However, it is easy to identify the origin of the
misconception: A great many of the experiments in the mature disciplines do involve instruments and complex procedures. But it is important to emphasize that the complexities are simply a reflection of the advances made in those fields, that is, the accumulated wisdom about what steps are necessary to make observations of phenomena in a way that is free from distortion. Naturally, as experimentation proceeds, the discipline gains experience in recognizing previously overlooked sources of error and in finding ways to compensate for them. Several of the authors in this volume criticize their initial experimental results and conduct new experiments designed to overcome what are perceived as potential sources of error; see, for example, the chapters by Lindblom, Derwing and Nearey, Kawasaki, Ohala and Ohala, Wright, and Janson. The instrumentation or complex procedures, however, are not the essence of experimentation; the level of complexity is dictated solely by the level of sophistication attained in previous investigations of the given question. What is common to all experiments is taking as much care as possible to refine one's beliefs. It may be asked—and has been (Chomsky 1964, p. 81, 1965, p. 20)—whether such care in making observations of the universe should be the primary concern of scientists or whether the search for insight into the nature of the universe should be the principal goal. Of course, the desire to understand the workings of the universe has driven all scientific work since the time of the early Greek philosophers. But, as was demonstrated by Galileo and other fomenters of the Scientific Revolution, little progress can be made toward achieving this goal unless care is taken in observing the universe. To build a bridge it is not enough to just have the intention of getting to the other side; one must have some experience with how various materials—wood, stone, steel—behave when a load is applied to them and then how to correct any behavior which interferes with the main purpose of the structure.

Although conceptually one may regard making theories2 as distinct from making experiments, in practice the distinction is blurred. First, experiments, in the sense used here, have no meaning unless motivated by theory. Theory dictates which observations, of the infinite observations that could be made, should be made. Without theory there would be no indication of what to observe and how to interpret it once observed. Aiming data gathering, therefore, does not constitute doing experiments. On the other hand, theory construction (when this is correctly considered not as a static thing but as

2 Traditionally, a distinction was made between the terms 'hypothesis' or 'speculation' on the one hand and 'theory' on the other, in that the latter term was reserved for hypotheses that had been subjected to a substantial body of experiments and supported by their results. In current linguistic usage, however, 'theory' has been used in the sense of the original term 'hypothesis', that is, a guess or speculation that often has little or no empirical support. In the discussion here we follow, with some misgivings, the usage current in linguistics.
something that develops and evolves) that is not checked and guided by experiment is equally useless, as numerous cases in the history of science reveal to us, for example, the fantasy of Cartesian cosmology.

Second, experiments often require that a theory be refined, sometimes so extensively that the result differs greatly from the original, casually offered, theory. When trying to find an observable consequence of the theory, it is frequently necessary to make quite explicit what the theory would predict in specific situations. Especially when theories are first tested, as is the case in our field, as much or more effort may go into a careful, explicit, (re)formulation of them as into the data gathering itself (the data in some cases may already have been gathered for some other purpose). It is in this type of situation that models are important. A model is, ideally, simply a theory that is so explicit in its formulation that it eliminates all ambiguity as to what it would predict. It should be able to stand on its own without being propped up, or fudged, by its maker. The model, once wound up and made to "go," either behaves in a way that accords with observations of the thing it represents or is judged wanting. In the present book, the chapters by Lindblom, Nearey and Hogan, Wright, and Derwing and Nearay report efforts to refine a model sufficiently to bring it into alignment with empirical observations.

This discussion of models, then, also provides a clear answer to an issue that is currently debated in phonology: the role of formalism. Formalism makes theories explicit. Formally stated theories are easier to test than informally stated ones. Beyond this, formalism has no further function. The formal statement of the theory is the first step of the many needed to properly test a theory. Perhaps it is unnecessary to repeat this in a book on phonology, since it is over a century since phonologists were warned about putting too much reliance on the

method of investigation according to which people observe language only on paper and resolve everything into terminology, systems of rules, and grammatical formalism and believe they have then fathomed the essence of the phenomenon when they have devised a name for the thing . . . for on paper almost everything is possible.

(Osthoff & Brugmann 1878/1967:202, 198)

Of course, there is no sure path to the truth—with the exception of divine revelation—and the results of any experiment will always be subject to criticism of some sort and possible rejection. The same is true of the non-experimental or non-empirical methods used to evaluate theories, for example, the authority, prestige, or eloquence of the person or group espousing the theory or the perceived elegance or simplicity of the theory. It may seem, then, that in comparison with the latter methods, experimental methods come out the loser: non-empirical methods involve much less effort but the result
is the same, only tentative, temporary "truths." In the short term, especially
during that period when a field is just beginning to employ experimental
methods, this may indeed seem to be the case. But in the long term, the result
is not the same. Changes in the direction of current theory which come about
due to debating methods or the rise and fall of the prestige of theory makers
show little or no coherence. Such is evident, for example, in the periodic
replacement of one dominant school by another in Greek (and subsequent
European) philosophy, as well as in other fields such as political science, litera-
tary criticism, and theology with little or no experimental base. On the other
hand, changes of theory which are motivated by examination of experimental
or empirical results exhibit the property of convergence. We may judge
an experiment (others' or our own) to have been misguided or deficient in
one way or another. With the experimental tradition, though, the proper
response to a "bad" experiment is a better experiment. As a consequence,
as history demonstrates, the experimental results get progressively better in
the sense of giving us power over the world, that is, allowing us to predict
events with more accuracy. Dalton's atomic model had to be revised given
the experimental results of Rutherford and Thompson, and their concep-
tions in turn had to be revised given the findings of Bohr, Heisenberg, and
others. But at each stage of refinement some aspect of the old theory is
retained or, at the very least, is recognized as a necessary conceptual precur-
sor to later theories, as, for example, many historians of science regard the
early, now superseded, "fluid" theories of heat and electricity. The same
convergence, the same ultimate usefulness, of theories in a non-experimental
tradition is not found. It is not the rationalist construction of Cartesian cos-
ology which forms the basis of current space guidance systems. It is not the
a priori model of human physiology based on the four humors that forms
the basis of the medical tradition which today has been able to eradicate such
diseases as smallpox and poliomyelitis. Aside from applications which exploit
phonology's quite successful conquest of the linguistic past and its taxonomy
of elements of speech, the discipline has relatively little to show in the prac-
tical domain for the theoretical turmoil that has absorbed the energies of
phonologists over the last century or so. This is not because, like literary
criticism, society makes no practical demands of our field. If phonology were
making any progress in finding out how speech sounds behave (in their physi-
cal and psychological form), confident recommendations on methods of
second-language instruction, methods of correcting pathological speech, ways
to synthesize high-quality speech, and ways to do automatic recognition of
speech would be possible. It might be expecting too much to look so soon
for solutions to these very difficult problems, but one cannot help wonder
if phonology's apparent lack of progress in these and other areas is due to
its weak empirical base.

The characterization of the nature of experimentation given here may be
accepted, but many would still maintain that the experimental method is not appropriate to phonology or any aspect of linguistics (Itkonen, 1978). Among the reasons for this view might be the quite correct assessment that linguistic behavior, unlike the behavior of physical systems such as pendula, the orbits of electrons, and chemical reactions, is dependent on so many factors that it is virtually impossible to control them all in contrived, experimental settings. But this situation only tells us—trivially—that it is more difficult to do experiments in linguistics (or any behavioral science), not that it is any less necessary. Bernard, who wanted to understand the workings of the human body, faced the same problem and the same objections to experimentation, but answered:

Experimentation is undeniably harder in medicine than in any other science; but for that very reason, it was never so necessary, and indeed so indispensable. The more complex the science, the more essential it is, in fact, to establish a good experimental standard, so as to secure comparable facts, free from sources of error.
(Bernard 1865/1957, pp. 2-3.

The issues of whether phonology should adopt accepted scientific methods of validating theories or whether it should isolate itself from the scientific mainstream, as in earlier days physiology tried to do, is such a fundamental one, on which phonologists hold such strong opinions, that probably no amount of argumentation will change anyone's mind on the matter. The ultimate form of persuasion, if any would succeed, though, should be "existence proofs": There is no more convincing way to show that experiments can help to answer questions in phonology than by answering phonological questions through experiments. Ultimately, it was the success of the experimental method in physiology rather than Bernard's eloquent persuasion which finally transformed that field. We hope the papers in this volume will add to the growing body of literature which constitutes this existence proof.

INTRODUCTION TO THE CHAPTERS
IN THE PRESENT VOLUME

The chapters in this volume span a wide range of subareas in phonology: the first six deal with phonological universals, the next with the way language-specific structure influences native speakers' perception of speech, the next five with the psychological mechanisms underlying speakers' mastery of the sound system of their language, and the last chapter with the detection of ongoing sound change.
INTRODUCTION

Studies on Phonological Universals

Phonetic and phonological universals—whether absolute universals of the sort ‘all languages utilize stops in their sound inventory’ or statistical universals like ‘sequences of labial consonant plus [w] tend to be avoided in syllable onsets’—give us important clues to the fundamental mechanisms underlying speech. They may serve a function comparable to that of Mendeleev’s periodic table of elements in the development of chemistry.

Björn Lindblom, in his chapter “Phonetic Universals in Vowel Systems,” offers a further refinement of the model presented in Liljencrantz and Lindblom (1972) which attempts to predict, on the basis of articulatory, acoustic, and perceptual constraints, the spacing of the vowels in the vowel space for languages with a given number of vowels. He elaborates the perceptual components of the model with current findings from psychophysical models of auditory processing, achieving a significant improvement in the agreement of the model’s predictions with data on the vowel inventories of languages. Lindblom argues persuasively for the elimination, where possible, of axiomatic postulated constructs in phonology, for example, distinctive features and natural processes, and their replacement with comparable deductively derived constructs.

A test of a hypothesis about how nasalization affects vowel quality, which was offered to account for certain sound changes, is reported in the paper by James T. Wright, “The Behavior of Nasalized Vowels in the Perceptual Vowel Space.” His study, which uses the method of multidimensional scaling of perceptual judgments, attempts with considerable success to reconcile phonological facts and phonetic predictions, the latter from the articulatory, acoustic, and perceptual domains. Wright’s results are particularly important in that they illustrate the role of the listener in the implementation of sound change and, like Janson’s and Kawasaki’s chapters (discussed below), that these things can be studied in the laboratory.

Investigations of phonological universals require cross-language comparisons of phonetic and phonological data. There are many areas where it is not very clear exactly how to make these comparisons. Vowel quality is one such case. The measured acoustic properties of vowels are influenced to a major extent by the dimensions of the speaker’s vocal tract as well as by the language- or dialect-specific vowel norms. Modern theories of the forces that shape vowel spaces have now become sufficiently detailed that it is important to be able to know if it is true, for example, that, as claimed, the Danish /i/ is higher than the English /i/. Sandra F. Disner, in her chapter “On Describing Vowel Quality,” offers a solution to this problem by showing the advantages of an analysis of variance of acoustic measures of vowel quality and applies her results in an evaluation of current theoretical work on universals of vowel systems.
Haruko Kawasaki, in her chapter "Phonetic Explanation for Phonological Universals: The Case of Distinctive Vowel Nasalization," presents the results of a perception experiment which helps to explain universal (cross-language) tendencies between distinctively nasalized vowels and nasal consonants, specifically, the fact that the two tend not to appear adjacent to each other. She shows that listeners hearing the same nasalized vowel in the context of a nasal versus non-nasal consonant "discount" some amount of the nasalization when a nasal consonant is nearby, because it can be "blamed" for some of the nasalization. They thus fail to hear the vowel as having as much nasalization as it actually does. This finding has significant implications for a wide range of other universal constraints on the sequencing of sounds as well as for mechanisms of sound change.

In a chapter using a quite different technique from the others, Ian Maddieson demonstrates in "The Size and Structure of Phonological Inventories: Analysis of UPSID" how a statistical analysis of the phonemic inventories of a balanced sample of hundreds of languages (as derived from the linguistic literature) can be used to test certain claims about phonological universals. He looks at a number of hypotheses about phonemic inventories, including questions of whether there is an optimal number of segments in a phonemic inventory, the relation between the size of an inventory and its segmental content or structure, and the familiar hypothesis that complexity in one area of the phonology of a language (e.g., vowels) is compensated for by simplicity in another (e.g., consonants or suprasegmentals); on this last question he finds that, contrary to the common belief, "complexity of different kinds goes hand in hand."

The revolution in electronic and digital technology offers phonology unprecedented opportunities to acquire high-quality data with considerably less difficulty than was possible in the past. Sandra Pinkerton, in her chapter "Quichean (Mayan) Glottalized and Nonglottalized Stops: A Phonetic Study with Implications for Phonological Universals," describes how she was able to travel through Guatemala with a miniature phonetics lab and collect aerodynamic data from several speakers of various Mayan languages. The results help clarify phonetic descriptions of the stops in these languages which figure as exceptions to certain claims about universal correlations between voicing, place of articulation, and the ingressive-egressive character of glottalized series of stops. In addition, she discovered a glottalized stop with an unusual biphasic oral pressure impulse that might be the link between the ejectives and implosives, which are found to alternate in certain Mayan dialects.

The Influence of Language Structure Upon Perception

The idea that the way people perceive speech is biased by the sound system of their native language, particularly by its phonemes, was explicitly
stated, on the basis of empirical evidence, by Sapir in the 1930s. Although modern perceptual phonetic studies have provided more convincing empirical evidence of this phenomenon, it is not yet possible to predict precisely, given the phonological inventory of the speaker's language, what form this "warping" of the perceptual space will take. Terrance M. Nearey and John T. Hogan, in their chapter "Phonological Contrast in Experimental Phonetics: Relating Distributions of Production Data to Perceptual Categorization Curves," apply to this problem two mathematical models originally developed for general-purpose pattern recognition or categorization of input signals. They demonstrate the potential of these models for the categorization of stops, fricatives, and vowels. To the extent that these are useful models of the native speaker's perceptual processes, they may be applicable as well to automatic speech recognition.

Psycholinguistic Studies of Phonological Issues

The chapters by Lyle Campbell ("Testing Phonology in the Field") and Jean-Marie Hombert ("Word Games: Some Implications for Analysis of Tone and Other Phonological Constructs") show how word games can be used to shed light on a variety of phonological issues: the psychological reality of certain phonological rules; whether phonotactic rules or sequential constraints require reference to purely phonological units such as syllables or whether morphological information is needed as well; whether tone (and other phonetic properties of words) can be analyzed in a segmental way or as prosodies that "float" on the word, that is, which need not always be temporally tied to the segments of the word. On the latter point, Hombert presents evidence that African and Asian tone languages behave somewhat differently. Both chapters illustrate how a technique that has long been used by some fieldworkers to clarify points of phonological analysis can, with the introduction of certain controls, become a valuable experimental technique.

In "Experimental Phonology at the University of Alberta," Bruce L. Derwing and Terrance M. Nearey give an account of a wide range of pioneering psycholinguistic phonological studies they and their co-workers have undertaken since the early 1970s. These range from the problem of trying to discover the form of the rule by which native speakers of English generate the phonetic shape of the plural suffix, to the basis of subjects' judgments of phonetic similarity or dissimilarity between two phoneme strings, to the psychological reality or accessibility of various phonological rules. Their work illustrates the advantages of having a continuing program of experimental studies in order to be able to refine data and techniques and thereby converge on a common result.

The power of the concept-formation experimental paradigm for psycholinguistic studies in phonology is demonstrated by Jeri J. Jaeger in her chapter, "Concept Formation as a Tool for Linguistic Research." The technique,
which involves, in essence, teaching subjects linguistic concepts or categories by induction and then giving them an opportunity to apply these concepts to phonological entities whose categorization is controversial, is appropriate for both laboratory and field work. She reports the results of several concept-formation experiments that clarify the psychological reality for English speakers of the phonemic membership of the phone [k] in word-initial [sk] sequences, the status of [tʃ] and [dʒ] as unit phonemes versus clusters, and the phonetic feature [voice].

Language-specific morpheme structure constraints (MSCs) represent a construct that has survived the many overhauls in phonological theory since the 1950s. Investigations of the psychological representation of MSCs are presented by John J. Ohala and Manjari Ohala in “Testing Hypotheses Regarding the Psychological Manifestation of Morpheme Structure Constraints.” Using experimental techniques previously applied to studies of MSCs by Greenberg and Jenkins (1964) and Zimmer (1969), they pit the Greenberg and Jenkins model for MSCs against those proposed in early generative phonology and in The Sound Pattern of English (Chomsky & Halle 1968). The Greenberg and Jenkins model, which is based on a general method for comparison of candidate words with words stored in the mental lexicon, is more in accord with the experimental results.

The Detection of Ongoing Sound Change

Sound change—the steps by which distinct daughter languages arise from a common parent language—was the first topic that phonologists succeeded in analyzing in what may be called a scientific way, that is, with a clear understanding of the possible sources of error in such an undertaking and the development of rigorous procedures to compensate for those errors (Rask 1818; Grimm 1822; Schleicher 1861–1862). In part, though, the comparative method relies on an intuitive sense of what kinds of sound changes are likely to occur and which direction they might take, for example, that /ki/ → /tʃi/ is more common than the reverse. It remains to put this aspect of the comparative method on as firm a basis as the other components. It is in this context that Tore Janson’s paper, “Sound Change in Perception: An Experiment,” is particularly important. He demonstrates that sound change can be studied in the laboratory. He finds a difference between older and younger speakers of Stockholm Swedish for the perceptual norms of the vowels /a:/ and /oː/, a difference that can be attributed to an ongoing sound change. In addition, his results help resolve the issue of whether sound changes are phonetically gradual or abrupt; he shows that the sound change he studied must be counted as phonetically gradual since it is doubtful that listeners of either generation would ever have a chance to exhibit a clear
category shift on the vowel continuum in a way that would be detectably different from the other generation.

CONCLUSION

In every discipline the number of questions asked always exceeds the number of questions successfully answered. In phonology there have really been only two kinds of questions that have received answers that have demonstrated, time-tested validity: (1) What is the physical nature of speech sounds? Panini, some 2300 years ago, gave initial answers to this question, and phonetics research up to the present time continues to refine this knowledge; (2) What is the history of languages and language families? In the nineteenth century the classical grammarians, using and refining the comparative method, successfully reconstructed the linguistic past. These questions bore fruit because they were married to suitable methods of investigation. Currently phonologists have given much attention to many other questions: What psychological structures underlie language use? How and why does sound change take place? What forces shape segment inventories? To obtain answers to these questions it is necessary to develop appropriate methods to study them. This volume represents a collection of studies that employ candidates for these needed methods and at the same time offer candidate answers to some of these important questions. We hope that phonologists will find these papers stimulating—whether they agree with the conclusions reached or not—and go out and do some experiments of their own.

REFERENCES


Paris: J. B. Bailliere et Fils, 1865)


