EXPERIMENTAL PHONOLOGY

John J. Ohala
University of California, Berkeley

This paper is an advertisement for Experimental Phonology. I will try to promote an approach to phonology that I think will do for the field what Buddhism promises to do for the soul: to permit it to escape the endless and agonizing cycle of birth and death of trendy theories, schools, frameworks, etc. and achieve oneness with the spirit and principles that guide all scientific endeavor, be it physics, chemistry, physiology, psychology, or linguistics. The structure of my paper will be as follows:

--To characterize briefly what experimental phonology is.
--To counter certain myths or misconceptions about experimental phonology.
--To provide examples of phonological experiments.

1. WHAT IS EXPERIMENTAL PHONOLOGY?

Undoubtedly the popular image of an experimental discipline includes complicated procedures or instruments. But this is a false image: first, the complexity is unessential and complex instrumentation by itself does not make an undertaking experimental. William Harvey conducted some of the first physiological experiments with little more equipment than a few tourniquets. What is essential for experimentation is an attitude: first, a keen awareness that the world is not necessarily as it may seem, i.e., that our sense-impressions and therefore the opinions and beliefs based on them may be faulty, and, second, the willingness to actively do something to compensate for or correct these potential errors by making observations under carefully controlled conditions. Pasteur (and others before him) doubted the apparent evidence of spontaneous generation of, for example, bacterial growth in food because he guessed that we might be unable to detect the minute air-borne spores that produced the bacteria. He therefore contrived a situation where foodstuff was left open to the air but where air-borne micro-organisms were prevented from reaching it. He demonstrated that under such circumstances the food remained unspoiled, thus further undermining the doctrine of spontaneous generation. What made his actions experimental was taking pains to make observations whose evidential value on the issue at hand would not be distorted by unwanted sources of error. As a field matures and accumulates more experience with experimentation it (a) learns of more potential sources of error and ways to compensate for them and (b) begins to ask questions of a more detailed sort which requires ever more careful observations. This is where the complicated instruments and procedures may enter in—simply as a natural outgrowth of the epistemological development of the discipline.
Perhaps it is no secret that phonology as an experimental discipline is in its infancy and the techniques used are still relatively simple. So all who would be the Galileos, Newtons, Harveys, Lavoisiers, and Pasteurs of phonology can still get in at the beginning where imagination and a breadth of knowledge counts for more than narrowly focussed technical knowledge.

There might have been some expectation that a characterization of experimental phonology would have made some mention of the specific techniques employed in phonological experiments. I have purposely omitted such a list because specific techniques are--in the grand scheme of things--of purely secondary importance. No such list could hope to be complete: the techniques used are limited only by imagination, of which, hopefully, there is an endless supply in linguistics. As one learns from the history of science, there is a constant evolution of experimental techniques in every scientific discipline. Nevertheless, I will present a bit later some examples of experimental techniques from my own work.

2. Some Misconceptions About Experimental Phonology

"Experimental phonology is just a new label for experimental phonetics."

There exists a well-established tradition called 'experimental phonetics', the founders of which include L'abbé Rousselot (1897-1901) in France and E. W. Scripture (1902) at Yale. Although many of the works of Rousselot and those influenced by him (Roudet 1910, Grammont 1933) were motivated by traditional phonological concerns, notably the causes of sound change, the parallel tradition initiated by Scripture, which eventually was the one adopted by the majority of experimental phoneticians, espoused a strongly positivist philosophy where only the supposedly objective measurements of speech obtained from instrumental records were accorded much value. To a large extent it was this positivist attitude which dominated the phonetic work in what was called 'speech science', i.e., the study of speech in departments of speech pathology.

Experimental phonology is quite distinct from this latter version of experimental phonetics in two respects. First, it is different in philosophy since it is not bound by a positivist attitude. Philosophically, it is very similar to mainstream science which perhaps is best labeled as 'hypothetico-deductive', i.e., driven equally by hypotheses, a product of reason, as well as by data, a product of the senses. Simply put, it attempts to reconcile what we think with what we see. Ladefoged's term 'linguistic phonetics' was probably coined in order to differentiate his practice, motivated by linguistic theory, from traditional, positivist experimental phonetics. Second, experimental phonology encompasses much more than phonetics; it incorporates certain areas of psychology as well as the sociolinguistic experiments of
the type performed by Labov (see, e.g., Labov 1966). For that matter, if experimental studies of animal vocalizations or even facial expressions could shed light on the behavior of speech sounds—which I believe is true (Ohala 1984)—then they, too, would be part of experimental phonology. Whatever is expected to yield reliable answers to traditional questions in phonology belongs in experiment phonology.

"Experimental phonology, per se, does not include theory construction."

It may also be thought that experimental phonology consists exclusively of testing of theories without theory construction and is less exciting because, as everyone knows, theory spinning is "where all the action is". But this is not the case. A more mature view of science, including phonology, sees it as a continuous cycle of theory—test—revised theory—revised test, etc. One does not just throw one's speculations to a doting public, take a bow, and then retire. A theory should contain within its statement the seeds of a test; the results of a test more often than not require the modification or even the complete abandonment of the theory and the formation of a new one. Rarely, if ever, are these activities farmed out to separate individuals. Claude Bernard, for example, regarded as the founder of experimental physiology (now simply 'physiology'), was the source of a good many theories over a century ago, e.g., concerning digestion, the function of the liver, etc.,—theories which survive to the present because they have survived attempts at falsification through tests—in many cases, his tests.

The inseparability of theory and experiment is based on the simple fact that if one can form a belief about something, that is, formulate a hypothesis or a theory, then one can also critically examine the origin of that belief and actively try to refine and control the observations that prompted it.

3. Some Examples of Phonological Experiments.

A. Levels of Phonological Representation.

It is widely assumed in modern phonology that there exist at least two distinct levels of representation of words: the 'surface' level which represents the actual pronunciation or something very close to it, and an 'underlying' level which may be closer to the lexical level, the form of pronunciation in the mental lexicon. For example there has been some discussion in the literature regarding the epenthetic stops which may appear in words such as warm[p]th, team[p]ster, young[k]ster, prin[t]ce. Some would argue that these stops are surface phenomena, not present at more underlying levels. It is possible to explain how these stops could be created as consequences of the assimilatory denasalization (and devoicing) of the latter half of the nasals. Neverthe-
less, this by itself is no proof that these stops are not present at more underlying levels; through sound change such stops have become integral parts of other words—if we can take consistency of spelling as a rough guide to the lexical status of the stops, e.g., glimpse, Thompson, dempster, bumpkin, Hampshire, resumption, thunder, and such spellings of something as 'sump'n'. Nevertheless, it is risky to put too much confidence in spelling since it may reflect simply the conventional, usually conservative, spelling before the sound change occurred. In the case of the stem glimpse, the OED records the following divergent spellings, where the starred years given entries with the 'p':

<table>
<thead>
<tr>
<th>Year</th>
<th>Spelling</th>
</tr>
</thead>
<tbody>
<tr>
<td>1386</td>
<td>glimsing</td>
</tr>
<tr>
<td>1400</td>
<td>glymsede</td>
</tr>
<tr>
<td>1540</td>
<td>glimsing</td>
</tr>
<tr>
<td>1551</td>
<td>glimsinge</td>
</tr>
<tr>
<td>1557</td>
<td>glimsing</td>
</tr>
<tr>
<td>1583</td>
<td>glimsinge</td>
</tr>
</tbody>
</table>
*1592 | glimpse |
*1598 | glimpse |
*1601 | glimpings |
1602 | glimsing |
*1603 | glimpings |
*1633 | glimpst |
*1635 | glimpsed |
*1657 | glimpings |
1663 | glimsing |
*1671 | glimpse |

(All subsequent entries contain 'p'.)

Given this much variation, who would be willing to say that the 'p' was underlying only around its first attestation of 1592 or that all 'p' forms subsequent to that gave sure testimony to its underlying character? In a search for more reliable evidence on the status of such stops I proposed that if certain durational characteristics of sounds are determined by their underlying syllable structure, it may be possible to differentiate surface from underlying stops by looking at their influence on the duration of surrounding segments (Ohala 1981a, 1981b, 1986a). It is well established that sonorants appearing before tautosyllabic voiceless obstruents are considerably shortened in English vis-à-vis those not closed by voiceless obstruents (Lovins 1978). Therefore, if a [p] in a word such as teamster is a purely surface entity, the vowel and nasal should be relatively long; if it present at the underlying level, however, they should be relatively short. Simplifying somewhat, I tried this idea out first on non-existing but possible English words clam + ster and clamp + ster. (Instructions: "add 'ster' to 'Clam', etc."; elicitation of these two forms done several minutes apart and mixed in with other distractor items.) Fig. 1 shows the measured VN durations from recordings of subjects producing these neologisms.
Fig. 1. Ordinate: Duration (in msec) of VN sequence in, from left to right, clamp + ster, clam + ster (with no detectable epenthetic stop), and clam[p] + ster (with epenthetic stop). Standard deviation indicated by vertical lines.

The results were consistent with the idea that the durational characteristics of the VN before the [p]'s in clamster were different—longer—than those in clampster, presumably because the former were more "surfacy" than the latter.

There are other ways to demonstrate the surface or non-surface character of sounds and they need to be used more widely in phonology than they are at present. Differentiating such levels on purely impressionistic grounds or, worse, on the basis of the thoroughly discredited and meaningless criterion of 'simplicity', is foolhardy.

B. Recognition of the Connection between Words.

Since the time of Panini and Plato it has been the task of phonologists to demonstrate the relationship between words based in part on their sound. Recently, the phonological similarity of words has been cited as evidence of their psychological one-ness.
Thus, given the existence of pairs of words such as those in (1)

(1) obscene obscenity
grain granular
code codify
South Southern

known to be related historically, semantically, and orthographically, it is assumed that for each pair the speaker knows of a common underlying form from which the different pronunciations are derived by the application of certain phonological rules. But what evidence do we have that the native speaker of English knows this never-pronounced form? Well, first there is our knowledge derived from linguistic study of the history of English words. But we know this is an unsure guide: beef and cow are historically related, as are nerve and snare, but average speakers are so unaware of the connection that they typically display mild surprise at being informed of this. Well, in addition, there should be a fair amount of phonetic similarity between the words in order to have common underlying forms. But this is an unsure guide, too, because such pairs as admire / admiral, mar / marriage are phonetically similar but few, I think, would posit them as having common underlying forms. How about also insisting that there be a semantic relationship, too? This also is an unsure guide, as beef and cow would indicate, as well as such pairs as leap / leopard or penis / penetrate. The answer, as is well known, is that word pairs that have a common underlying form should have both semantic and phonetic similarity and, moreover, should exhibit a common, widely-attested, phonetic relationship. Thus, extreme / extremity meet this requirement since there are many other pairs showing the same kind of alternation, but pope / papal would not since there are few other pairs showing the same relationship (in fact, to my knowledge, only one: nose / nasal). Creating a common underlying form for extreme / extremity would take advantage of a generalization—the phonological rules which relate tense and lax vowels—creating a common underlying form for pope / papal could not take advantage of any such generalization. This argument, then, is based on the assumption that the native speaker is good at recognizing and squeezing out all the general sound patterns permitted by the surface forms of words. But is the native speaker as good as making such generalizations as we linguists are (Derwing 1977)? Our judgements on the matter are unreliable because our background biases us to see general patterns where the non-linguist may not. There are countless 'generalizations' in the visible universe waiting to be made but history reveals that nevertheless they only became "obvious" after someone points them out, e.g., that outgrowths of facial hair in primates, including humans, is always located at the periphery of the head, i.e., where it would be most obvious to a viewer facing the individual and thus help to make the head (and presumably the head's owner)
look larger (Guthrie 1970).

Manjari Ohala and I designed an experiment to test whether native speakers can differentiate general from particular sound patterns that connect pairs of words (Ohala & Ohala, in press). The experiment, an elaboration of one conducted by Derwing and Baker (1977), was carried out as a class project by students in one of my graduate courses. Briefly, we drew up a list of 20 words thought to exhibit various common phonological relationships and 20 words exhibiting various isolated phonological relationships; see (2).

(2) Common Patterns

<table>
<thead>
<tr>
<th>particle / particular</th>
<th>thumb / thimble</th>
</tr>
</thead>
<tbody>
<tr>
<td>substance / substantial</td>
<td>strong / stringent</td>
</tr>
<tr>
<td>extreme / extremity</td>
<td>pope / papal</td>
</tr>
<tr>
<td>resume / resumption</td>
<td>applaud / plausible</td>
</tr>
<tr>
<td>abstain / abstention</td>
<td>peace / pacify</td>
</tr>
<tr>
<td>regal / regicide</td>
<td>nose / nuzzle</td>
</tr>
<tr>
<td>comprehend / comprehensive</td>
<td>slay / slaughter</td>
</tr>
<tr>
<td>erode / erosion</td>
<td>price / precious</td>
</tr>
<tr>
<td>permit / permission</td>
<td>mouse / muscle</td>
</tr>
<tr>
<td>proper / propriety</td>
<td>toad / tadpole</td>
</tr>
<tr>
<td>secret / secretary</td>
<td>confer / confession</td>
</tr>
<tr>
<td>Peter / petrify</td>
<td>live / liver</td>
</tr>
<tr>
<td>magnet / magnesia</td>
<td>linger / lingerie</td>
</tr>
<tr>
<td>vine / vinegar</td>
<td>page / pageant</td>
</tr>
<tr>
<td>fable / fabulous</td>
<td>promise / promiscuity</td>
</tr>
<tr>
<td>glass / glacier</td>
<td>tame / timid</td>
</tr>
<tr>
<td>vocal / vociferous</td>
<td>leap / leopard</td>
</tr>
<tr>
<td>marine / marinate</td>
<td>male / malicious</td>
</tr>
<tr>
<td>slipper / slippery</td>
<td>risk / rescue</td>
</tr>
<tr>
<td>sect / section</td>
<td>haste / hassle</td>
</tr>
</tbody>
</table>

Some of the word pairs in the latter set are not, in historical fact, related to each other ('confer / confession' to 'haste / hassle'). However, native speakers may not know this, and were, in any case, free to express their own opinion on the matter. (Here I simplify the description of the procedures; for details see Ohala & Ohala, in press.) We presented these pairs orally, randomized, to 16 English speakers, and asked them rate them on a 5 point scale, first as to their derivational relationship ("could they have a common ancestor?"), then on a second run, on their semantic similarity, and finally on a third run on their phonetic similarity. Before each run we offered and discussed a smaller practice set of different word pairs whose ratings would not be controversial, e.g. 'parasol / umbrella', 'lamb / lamp'. Subjects were told that their judgements would help us select items that would be used in an aptitude test for high school students and that their answers should simply reflect their intuitions as educated adults.
We hypothesized that, for word pairs of comparable phonetic and semantic similarity, the generality of the sound pattern relating them would make subjects see a closer derivational connection. That is, that after taking into account the various semantic and phonetic judgements, subjects would give significantly higher derivational closeness ratings to pairs of the sort 'extreme' / 'extremity', than they would for those like 'pope' / 'papal'. It turned out that the degree of judged phonetic similarity counted for very much less than the semantic similarity in determining or correlating with the derivational judgement, so that final analysis was done just with semantic vs. derivational judgements. This result may be seen in Fig. 2.

The solid regression line, a logarithmic function derived by the least squares method, gives the best prediction of the derivational ratings of all 40 test words based on their semantic rating. The dashed line gives the function for the 20 words showing subjects' judgements of derivational relationship as a function of their judgements of semantic similarity of the word pairs in the experiment. Solid line: all 40 word pairs; dashed line: 20 words exhibiting common sound patterns; dotted line: 20 words exhibiting isolated sound patterns.
exhibiting the common patterns and the dotted line, the function for the 20 words showing the isolated patterns. The latter two functions are not significantly different, however, from the first one, i.e., they do not account for significantly more variance.

We took these results to show that, contrary to what is commonly assumed, native speakers do not necessarily differentiate general from particular sound patterns, at least in derivational phonology. This should not be a very surprising result: what payoff to the native speaker is there for the psychic energy expended in noticing the general sound patterns and working out the common underlying forms for pairs related by them? Remarkably little, it would seem. Memory space is certainly not at a premium and in any case everyone recognizes that each member of pairs such as 'extreme'/ 'extremity' has to be stored separately since they have at the very least idiosyncratic semantic and occasionally syntactic features—the same is true for most such pairs: 'divinity' is commonly used both as a noun and an adjective although this is not true of 'serenity', in spite of their both showing the same phonological pattern. The generalizations may have some value in spelling and reading but it is hard to evaluate this because most pairs related, for example, by vowel shift, have other orthographic cues to the tense and lax distinction, i.e., the silent 'e' at the end of 'extreme' 'divine' and the like. More attention should be given to a cost/benefit analysis of forming phonological generalizations. There may be some surprises.

C. Dissimilation.

The preceding experiments tested long-standing theories or assumptions. I turn now to a novel theory, of my own, which provides an explanation for dissimilation. Again, I give only a brief account; for further details see Ohala 1981c, 1983, 1985, 1986b.

Dissimilations such as Grassmann's Law (3) or the dissimilation of glottalization in Quechua (4), and Salish (5) are puzzling

(3) Dissimilation of Aspiration in Sanskrit and Greek (Grassmann's Law; data from Brugmann 1886:355ff).

<table>
<thead>
<tr>
<th>Proto-Indo-European</th>
<th>Sanskrit</th>
<th>Greek</th>
</tr>
</thead>
<tbody>
<tr>
<td>*bhagha-ti</td>
<td>bōḍhāti</td>
<td>peúthomai</td>
</tr>
<tr>
<td>*dha-dha-ti</td>
<td>dāḍhāti</td>
<td></td>
</tr>
</tbody>
</table>
(4) Dissimilation of Glottalization in Quechua (Orr & Longacre 1968).

Proto-Quechumaran  Aymara  Proto-Quechuan
*p'amp'a  p'amp'a  p'ampa "cover w/ dirt"
*t'ant'a  t'ant'a  t'anta "bread"


Proto-Interior  Shuswap  Salish
*k'íp'  kíp'-m  "pinch"
*q'wác'  qwác'  "full"
*p'úx'  s-pút'-nt  "fog"

for at least two reasons: first, it involves what seems to be 'action at a distance', i.e., influence of one segment upon another skipping over intervening unaffected segments. Second, there must be some natural reluctance to admit dissimilation into the canon of recognized natural sound changes if we also include assimilation. How can we suggest with a straight face that it is natural both for sounds unlike to become more similar and also for similar sounds to become less similar?

My solution to this is relatively simple: first, these are not instances of action at a distance. Rather, it must have been the case that the aspiration which dissimilated in the Grassmann's Law cases and the glottalization which dissimilated in Salish, originally spilled over onto adjoining segment. Normally, listeners "correct for" such predictable assimilations, factoring out the predictable breathiness or tense voice when they occur non-distinctively on adjacent vowels and sonorants. Dissimilation occurs when they apply these 'corrective' rules inappropriately. That is, they engage in a kind of phonetic hypercorrection. Another way of looking at it is to say that the distinctive use of a given feature at one site in a word camouflages its distinctive function at another site in the same word. When these hyper-correcting listeners speak such words themselves, they omit the feature at the place where they thought (erroneously) that it was purely non-distinctive or predictable.

There is evidence for the posited 'spillover' of the features like glottal constriction in glottalized segments; see Keller (1959). Furthermore, this theory predicts that there should be rather strict constraints on the kind of features subject to dissimilation at a distance: namely, those known to spread several tens of milliseconds beyond the segment they are properly
Surveying the literature on dissimilation, this is borne out: the features that dissimilate are labialization, retroflexion, aspiration, glottalization, pharyngealization, uvularization, nasalization, place of articulation, etc. Features that do not spread, such as stop, fricative, or affricate, should not dissimilate—and it seems that in general they don't or in the cases where they are said to, the evidence is equivocal, sometimes by the testimony of the author making the claim; see Posner (1961:93, 99).

Dissimilation seen as hypercorrection is motivated by the kinds of distortions of speech sounds caused by assimilation. Sound changes due to assimilation, then, may be viewed as hypo (under)-correction. This allows us to see that although dissimilation is the opposite of assimilation in many respects—this was the source of phonologists' discomfort—they are not mirror images in all of their characteristics. It has been noted that the product of assimilation is often a new segment or series of segments in a language's sound inventory, for example, when Ancient Tibetan dropped syllable final consonants on it way to becoming Modern Lhasa Tibetan, back rounded vowels became front rounded vowels when the dropped consonant was a dental (6). This is

(6) Tibetan fronting of back vowels before final dentals, but not before non-dentals (Michailovsky 1975; transcription simplified).

<table>
<thead>
<tr>
<th>Written Tibetan</th>
<th>Lhasa Tibetan</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>drug</td>
<td>þhuu</td>
<td>&quot;six&quot;</td>
</tr>
<tr>
<td>thog</td>
<td>thoo</td>
<td>&quot;roof&quot;</td>
</tr>
<tr>
<td>nub</td>
<td>nuu</td>
<td>&quot;west&quot;</td>
</tr>
</tbody>
</table>

**BUT:**

<table>
<thead>
<tr>
<th>Written Tibetan</th>
<th>Lhasa Tibetan</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>bod</td>
<td>phϕϕ</td>
<td>&quot;Tibet&quot;</td>
</tr>
<tr>
<td>ston</td>
<td>tϕ</td>
<td>&quot;autumn&quot;</td>
</tr>
<tr>
<td>lus</td>
<td>lyy</td>
<td>&quot;body&quot;</td>
</tr>
<tr>
<td>spos</td>
<td>pϕϕ</td>
<td>&quot;incense&quot;</td>
</tr>
</tbody>
</table>

transparently due to the well-documented fronting influence that dental consonants have on back rounded vowels (Lindblom 1963, Stevens & House 1963). That is, we must assume that even before the final dental was dropped these vowels were phonetically but non-distinctively much like front rounded vowels. This front-rounded character became distinctive or phonologized when listeners no longer detected the final consonant and had no basis for predicting and therefore correcting the frontness. Thus hypo-
correction led to a new class of vowels. However, it has been observed that dissimilation never (or seldom) leads to a new segment. This follows directly from my account since if dissimilation is hyper-correction its action would be to normalize or undistort something thought to be abnormal or distorted. Therefore the changed segment will always be one which is already a familiar element in the segment inventory.

Furthermore, although in assimilation the conditioning environment may drop out at the same time as the conditioned change—as in the case of the Tibetan vowels and countless other cases one could cite (the development of distinctively nasal vowels from loss of a post-vocalic nasal consonant, the development of distinctive tone on vowels with simultaneous neutralization of voicing on preceding consonants)—in dissimilation, on the other hand, the conditioning environment may not be lost at the same time as the conditioned change occurs. This is because the conditioning environment must be there for the listener to attribute the imagined distortion to, i.e., to be the source of the camouflage. As I read the historical record, this is precisely what happens: the conditioning environment is not lost at the same time as the dissimilation itself takes place.

Finally—and this is why I chose to discuss this example—there is experimental evidence which supports this theory. The evidence for dissimilation at a distance is not very impressive yet, but dissimilation of contiguous segments can be easily demonstrated. As part of an experiment addressing a somewhat different issue, Ohala and Feder (1986) obtained listeners' judgements of the identity of synthetic vowels from the continuum between and including /i/ and /u/, when presented in isolation and the contexts of a following /b/ and a following /d/. The results, expressed as percentage of identification of /u/ is shown in Fig. 3. The solid line shows how the vowels were identified in isolation, the dashed line, in the /d/ environment, and the dotted line, the /b/ environment. It is clear that more /u/ judgements were obtained in the /d/ context. What this means is that some vowels which in other contexts were identified as /i/ were instead regarded as /u/—that is, they were 'backed'—in the consonantal context known to give rise to fronting of /u/ (see (6) above and the discussion accompanying it). I interpret this as a case of a fronting environment camouflage—dissimilating—some of the frontness of a vowel so that it is taken to be a back vowel.

4. Conclusion.

The ultimate argument for experimental phonology is the accumulation of successful and insightful applications of experimental methods to traditional questions in phonology. In my judgement that argument has been made. The literature of such applications goes back at least to 1901 and includes work by such distinguished linguists or psycholinguists as Edward Sapir (1929, 1933), W. Freeman Twaddell (1935:11), Stanley Newman (1933), Joseph Green-
Fig. 3. Ordinate: Percentage identification of synthetic vowel stimuli as /u/; abscissa: vowel continuum /i/ to /u/. Solid line: vowels in isolation; dashed line, in the environment /d/; dotted line, in the environment /b/ (from Ohala & Feder 1986).

berg (Greenberg & Jenkins 1964, 1966), Roger Brown (Brown & Hildum 1956, Brown & Nuttall 1959), Sandford Schane (Schane & Tranel 1970, Schane, Tranel, & Lane 1974), and James McCawley (1986) to name a few.

Furthermore, I take comfort in the fact that all phonologists who operate in an academic environment—and that includes the vast majority—give open testimony to their belief in the experimental method, if not in their practice of phonology then in the way they evaluate students. Academics are called upon periodically to assess the knowledge or skills of their students. It might be possible to take a non-experimental approach and assume that simply because students have been exposed to a given body of material that they therefore have absorbed or "internalized" it. But the academic community is not satisfied with that and demands evidence. What evidence? The evidence of tests or their equivalent, e.g., the production of original papers where the insights presented by the student could not have been the result of rote memorization. I don't think it is unreasonable to expect that this practice of relying on tests should be used not only in teaching but in the subject matter which is taught.

Antonie Cohen (personal communication) of Utrecht University tells an anecdote relevant to this. He says that in a certain university (which shall remain nameless) certain faculties do in fact, practice academic evaluation in way that most phonologists
practice phonology—or at least a variant on it. In these cases a final examination consists of all students enrolled in a given course filing into a room and the teacher asking questions out loud; if any individual student can answer the question it is assumed that they all know the answer. He did add, however, that this practice is only followed in schools that teach the so-called liberal arts, e.g., languages, history, philosophy, and not in those that teach brain surgery, bridge building, or car repair.

I leave it to the reader to decide whether phonological theories should be evaluated the way philosophers or brain surgeons are in this school. It is not a facetious question. In fact, the non-academic world is making increasing demands on phonologists to explain the workings of speech: in speech technology (synthesis of speech from text and automatic speech recognition), in language teaching, in speech pathology, in advertising, etc. (van den Broecke, Lindblom, & Ohala 1985). If we don’t satisfy these demands there could be adverse consequences: phonologists will get the reputation of being an elite ivory-tower enterprise, good for little else than to keep practicing phonologists off the streets—as long as indulgent tax-payers or tuition payers are willing to foot the bill. Furthermore—and this is, in my view, the worst outcome—inferior answers to phonological questions will supplant better answers in the areas that need them. Not only phonologists will lose in that case, but all of society. Experimental methods can solve this problem. It has worked for other disciplines, from physics to physiology to pedagogy. It can work for phonology, too.

5. Acknowledgements.

I thank Debbie Feder and Manjari Ohala for their collaboration on some of the research reported here. The following students in my graduate seminar assisted in the conduct and some aspects of the design of the ‘word relatedness’ study: Mariscela Amador, John Cherry, Hazel Corcoran, Barbara DeMarco, Debbie Feder, Randy LaPolla, Kiki Nikiforadou, Jing Wang, and Barbara Weldon. This research was supported in part by grants from the Committee on Research and the Cognitive Science Program (through a grant from the Sloan Foundation), both of the University of California, Berkeley.

References

Brugmann, K. 1886. Grundriss der vergleichenden Grammatik der
Berlin: Springer-Verlag. 447-475.
Twaddell, W. F. 1935. On defining the phoneme. Lg Monog. No. 16.