Title
Understanding Variability in Speech: A Brief Survey Over 2.5 Millennia

Permalink
https://escholarship.org/uc/item/0j86c7st

Journal
UC Berkeley PhonLab Annual Report, 4(4)

Author
Ohala, John J

Publication Date
2008
UNDERSTANDING VARIABILITY IN SPEECH: A BRIEF SURVEY OVER 2.5 MILLENNIA

John J. OHALA

ABSTRACT

Ever since the earliest attempts to understand the nature, the form, and variability in speech, whether within one language or the variation one finds in the pronunciation of the “same” word in different but related languages, there has always been interest in explaining such variation in physical, physiological, or psychological terms. After millennia of study of this subject this goal is within reach.

1. INTRODUCTION

From distant times there has always been curiosity about the nature and the mechanism of speech and language. This is reflected in myths suggesting a divine origin of speech or positing specific deities as the patron of speech: in Hinduism, Sarasvati, in Greek mythology, Hermes, in Norse mythology, Lodur. The same motivation to understand the workings of speech led to scientific, i.e., empirically-based studies of speech. In this paper I propose to give an interpretive history of phonological science over 2.5 millennia. My purpose in attempting this audacious task is to provide some perspective on modern mainstream phonology as practiced today in the West.

First, what do I mean by ‘phonological science’? I believe the essence of a scientific discipline is to be found in the questions it asks. The answers to the questions (i.e., theories, hypotheses, conjectures) may change or evolve over time and the methods used to obtain these answers may also change. The perennial questions asked in phonological science, I would maintain are:

1. How can one describe a language’s pronunciation to:

   a) lend it permanency in the face of apparent variation and change (there can be many motivations for this: to better interpret or understand important texts – including oral text); simply to establish a language or dialect as standard with respect to competing languages or dialects); it is not the linguists’ business or concern to pass judgment on the motivations behind linguistic description.

   b) for the aid of language learners

2. To establish relationships – typically familial or diachronic relationships between different (mutually unintelligible) languages. Again, the motivation for this can be varied: from “pure” scientific interest to chauvinistic, ethnic, or nationalistic purposes; but the motivation is less important to us than is the scientific support for the claims.

3. To discover the psychological basis of speakers’ language competence in the phonological domain.

4. To discover optimal techniques for language instruction/learning.

5. To discover how humans’ language capacity arose, i.e., how language evolved.
6. To understand speech pathologies such that they can be ameliorated.

7. Understanding variation in spoken speech to the point of being able to mechanize the functions of speaking and perceiving speech, i.e., automatic speech synthesis and automatic speech recognition.

In this paper I will focus on the history of contributions to the first three questions – without denying the importance of the remaining four in which remarkable progress has been made in the past few decades.

When appeals to divine intervention don’t satisfy, to which domains would the interested scholar turn to? The answer – and I believe this is still true to the present day: -- is (1) the physical, i.e., the anatomical, physiological, and acoustic-perceptual aspect of speech, (2) the psychological elements serving speech, and (3) the social influences and functions of speech.

2. DESCRIBING SPEECH

The necessary first step – and this is true of any scientific investigation – is to reduce the huge mass of data to a manageable few entities. This is the pattern in the development of astronomy, chemistry, physics, and biology. Like these domains, speech is highly variable. Paṇini and all the phonologists after him identified what they considered the basic “building blocks” of speech. Whether these correspond to individual phonemes or CV syllables, etc., this is the start of the scientific proposal for an understanding of speech.

The earliest description of speech and its elements that we have is a quite sophisticated one, that of Sanskrit by the Hindu grammarian Paṇini from (estimated) 3rd to 5th c BPE. In his monumental work, the Astadhyayi, the description was in quasi-phonetic terms but it was sufficient to differentiate the various phonemes and their morphophonemic variants. Paṇini was not the first to address the sound system of Sanskrit; he himself refers to earlier works that did not survive.

Other notable early descriptions of the sound systems of individual languages include:

- Various Greek and Latin grammarians’ descriptions of and philosophizing about their languages, including, e.g., Dionysius Thrax (Tēkhnkē grammatike, 2nd c. BPE) and Plato’s Cratylus (4th c. BPE; perhaps an intentionally humorous look at etymology).
- The rhyme tables of Middle Chinese (Sui Dynasty, c. 6th-7th c. AD). Although Chinese writing is not in itself phonetically-based, the compilers of the rhyme tables demonstrated a parsing or decomposition of the monosyllabic words into initials and rhymes. These tables have proven to be invaluable in the reconstruction of earlier stages of Chinese.
- Descriptions of Arabic (e.g., by Khalīl ibn Ahmad Al Farāhīdī (الفراهيدي أحمد بن الخليد، الراحم عبد أبو)) and Sībawaihi (صبيح) in the 8th c.
- The “Grammarian” (his name is unknown) who described Icelandic in the 12th c.

I would place in the same category as “description” the development of an orthography for a language. Instead of being an explicit description of a language's parts and how they are arranged, an orthography in an implicit account of the same. The few orthographies that embody some sort of iconism (usually articulatory) actually are bit more than an implicit description. Some linguistically significant orthographies – those that represent constituent sounds – are:

- Egyptian hieroglyphics – phonetically based, as first discovered and deciphered by Thomas Young (and subsequently elaborated and the decipherment finished by François Champollion).
- Devanagari (for Sanskrit and its daughter languages; ult. derived from Brāhmī script, c. 3rd c. BCE)
• The Olmec hieroglyphs from Central America where each symbol represented a CV sequence, from c. 3rd c. BCE. (Justeson & Kaufman 1993)

• The Japanese syllabary (as a systematization in 4th ~ 5th c. of parts of Chinese orthography, kanji, which in itself was not primarily phonologically based)

• Hangul, the Korean writing system, commissioned by King Sejong in the 14th c. (1446). Not only sound based, but also an iconic representation of position of the vocal organs.

• The development of indigenous writing systems for Cherokee, Eskimo & Inuktitut, Amharic, and various other Native American and African languages. (In many cases, the concept of a phonologically-based writing system was previously known and the innovation consisted in developing – a nouveau – a version adapted to a new language.)

It is often claimed that such sound-based orthographies approximated what would be called in modern linguistics as a phonemic description of the language’s sound system. This may be a somewhat exaggerated claim: (1) syllabaries have symbols for CV (and sometimes CVC) but usually not separate C’s and V’s, (2) Japanese kana have a separate symbol for syllable-final /n/ but also separate indivisible symbols for /na/, /ni/, /no/ etc. (3) Egyptian hieroglyphics had symbols for long vowels but not short vowels.

3. GIVING EVIDENCE THAT DIFFERENT LANGUAGES DEVELOPED FROM SOME COMMON PARENT LANGUAGE.

Here I highlight attempts which would satisfy our current sense of (some) scientific validity. (In fact there have been from ancient times wild and disreputable comparisons between one language and others, usually attributing variation to laziness and other unattractive traits of speakers of the “other” languages or dialects.) The more scientifically respectable claims made some attempt to discount borrowing as a cause of similarity and, most importantly insisted on there being many points of similarity: lexical, morphological, etc.

The very earliest attempts made “gestalt” comparisons between lexical items and didn’t take into account all morphological factors. The later accounts (by Rask, Grimm, etc. emphasized point-by-point similarities (e.g., the word initial ‘p’ in Greek and Latin corresponded to “f” in Germanic, etc. I would consider that both approaches exploited a kind of intuitive notion of probability and this is the essence of the COMPARATIVE METHOD. This method is simply an intuitive application of the statistical law that the probability, $p$, of the collection of events, $X = x_1, x_2, ..., x_n$, is the product of the probabilities of the separate events.

Although not used in a quantitative way by the historical linguists, it is a method capable of yielding a level of objectivity and accuracy of evidence comparable to that found in the “hard” sciences. Charles Darwin – some centuries after the linguists – used this same method of marshalling evidence in support of his theory of evolution, i.e., proof by the sheer weight and interconnection of the accumulated bits of evidence. Any discipline faced with data that is difficult to quantify and whose data is subject to distortion by a great many uncontrollable extraneous factors – and this is a characteristic of the behavioral, social and some of the biological sciences – must resort to this method of orchestrating the empirical support for its hypotheses. Linguistics was among the first, if not THE first, to do this.1

1 As a somewhat curious footnote on the history of the application of the comparative method: it was applied in a strictly mathematical way by the great English scientist Thomas Young to support the possible relatedness of Basque and Ancient Egyptian. He claimed to find enough cognate words to judge that they were related. His error: he “cherry-picked” the cognates and used a questionable model for estimating similarity of candidate cognates. (See Poser, http://www.hillposer.org/Papers/young.pdf)
Marcus Zuerius van Boxtorn. One of the earliest works linking the European languages historically was authored by Marcus Zuerius van Boxtorn (1612-1653) from the University of Leiden. Prompted by the discovery of an ancient statue which Boxtorn identified as from the “Scythian” culture, he identified Scythian as the language ancestor of the Germanic (including the Scandinavian), Romance, Greek, Celtic, Slavic, and Persian languages, based on whole-word “gestalt” similarities of vocabulary. Unfortunately he also claimed Turkish, Georgian and other languages to be part of the same family which we know by the light of subsequent research is not the case. And although he pre-dated Sir William Jones in claiming a family relationship of these European languages with a member of the “Indo-“ part of Indo-European, namely, Persian – his evidence on this latter point was erroneous.

Boxhorn himself cites Ioannes Aventinus (1477-1534) and Sigismundus Gelenius (1497-1554) as having earlier tried to demonstrate family relationships between various European languages.

János Sajnovics (1733-1785), a Hungarian, demonstrated the relationship between Sami (a.k.a. “Lapp”) and Hungarian (1770). This was taken further by Samuel Gyarath in 1799.

Lorenzo Hervas y Panduro (1735-1809) was a Jesuit missionary, who, when the Pope rescinded the missionary “license” of the Jesuits, attempted to consolidate and archive the observations and discoveries of his fellow Jesuits about the life, lore, and the languages they encountered from around the world. Hervas is credited with discovering the Austronesian family of languages: the language family with the widest distribution geographically, stretching from Madagascar on west to Easter Island on the east and encompassing scores of languages in the Philippines and Oceania. Again the method of comparison was of the “gestalt” type.

Jonathan Edwards, Jr (1745-1801) was an American churchman who has been credited with the discovery (1788) of the Algonquian family of languages in the Eastern regions of North America.

The crucial breakthrough in the history of historical linguistics and in linguistics in general comes from the fusion of phonetics and historical linguistics.

This is exemplified most notably in the work of Lambert ten Kate (1674-1731) a Dutch scholar who published an account of the relation between Dutch (at that time) and the well-known European classical languages (ten Kate 1723). Another pioneer in this domain was Charles de Brosses (1709-1777) who, in a 1765 work, using an original iconic feature-based phonetic transcription, demonstrated sound-to-sound and feature-to-feature correspondences in the vocabulary of related European languages.

This method was applied on a significantly larger scale by Rasmus Rask (1818) and Jacob Grimm (1822) who also used point-by-point comparisons between the parts of words in different languages. They are generally credited with sparking the revolution in historical phonology which still flourishes.

A crucial element of the method – part of what has become known as the Comparative Method - for establishing the phonological element of a language’s history was to insist that sound change tended to be regular; exceptions had to have their own regularity caused by a differing environment (this latter point furthered especially by Grassmann (1863) and Verner (1875).

Some remarkable innovations and discoveries were made using the comparative method: Schleicher reconstructed a parent language no longer extant (contra to previous practice of identifying only known languages – including dead ones – as possible parent languages); Saussure reconstructed speech sounds – the so-called ‘laryngeals’ in Proto-Indo-European –
known only by their effect on adjacent sounds, the laryngeals themselves no longer extant in any language known at the time (except that decades later Jerzy Kuryłowicz found them in Hittite texts).

Two parallel movements further strengthened the work of Historical Phonology: first, the description of more and more languages of the world, especially by those trained in phonetics and, second, the application of laboratory phonetics and other empirical studies to questions in historical phonology. In these areas there were significant contributions from Liu Fu (1925), Li Fang-Kuei (e.g., 1944), Wu Zongji (e.g., 2000), Y. R. Chao (1976), and William S-Y. Wang (1969). There is now a steady tradition of the integration of laboratory phonetics and historical phonology (Ohala 1981, 1993).

Applications of the Comparative Method in Historical Phonology to other languages led eventually – by the 20th c. – to the assignment of virtually all of the 6000 or so known human languages to one or another of some 20–30 families. This must surely count as one of the most significant contributions of linguistics to science, putting linguistics at the forefront of the behavioral sciences.

4. GIVING AN ACCOUNT OF THE PSYCHOLOGICAL BASIS FOR PHONOLOGICAL PATTERNS IN A GIVEN LANGUAGE, I.E., THE PSYCHOLOGICAL GRAMMAR.

An interest in the psychological aspect of language was stimulated by the writings of Saussure, Baudouin de Courtenay, Sweet, among others. A notable example was the work of Meringer & Meyer on speech errors (“Versprechen und Verlesen” 1895. Originally designed as a study to find out if dissimilation could be explained as due to speech errors (“slips of the tongue”), they gathered hundreds of speech errors in unrehearsed speech – to their credit they decided that cases of dissimilation were unlike naturally-occurring speech errors.

But there was much speculation at this time (and even today) that observed sound changes were somehow implemented purposefully – that they were a product of the speaker’s will – that they were teleological (e.g., to make speech easier to produce or easier to hear and understand).

My belief – the results of much of my work on sound change – is that once we understand the phonetic basis of sound change it can be understood as listener error – comparable in the auditory domain to scribal errors when copying manuscripts. i.e., we do not have to posit complex psychological mechanisms.

Around the end of the 19th c and the beginning of the 20th the theory of the phoneme was developing: the idea that underneath the plethora of speech sound types that the phoneticians were discovering, there was really just a relatively small number of ‘phonemes’ which were the psychologically distinctive elements. The rest, the ‘allophones’ were contextually-determined variants of the phonemes.

My observation: unfortunately the principal method for discovering these alleged psychological entities was again the comparative method applied over a shallow time depth. Although the target was a psychological entity, there were no appropriate psychological methods to find them.

In the mid-20th c., generative grammar came on the scene and proposed to be able to discover the psychological underpinnings of sound variations manifest in related words in different morphological and syntactic environments.

E.g., \[\text{[aj]} : [i]\]

\[
\begin{align*}
\text{divine} : & \quad \text{divinity} \\
\text{sine} : & \quad \text{sinuous}
\end{align*}
\]

4 Chao and Wang were once the author’s colleagues at University of California, Berkeley.
Criteria for deciding on the architecture of these psychological mechanisms were: generality and simplicity. This is what justified abstract underlying form, ordered phonological rules changing features of the underlying phonemes, markedness, and a host of other devices.

The result of mainstream phonology’s abstract, imaginary conception of the elements of speech: a separation of the work of phoneticians (and other empirically-oriented studies of speech) and the “phonologists”. And there has been a bewildering proliferation of phonological “theories” with an average a “half-life” of 7.5 years:

In short: there is no evidence that the machinery proposed to account for phonological variations like English *divine* ~ *divinity* or even ‘want to’ ~ ‘wanna’ has any relation to speakers’ psychological mechanisms. The sole criterion of positing mechanisms with maximal generality (via maximum simplicity) has never been justified psychologically. The methods used really resemble the comparative method applied at time depth shallower than that appropriate for, say, the Romance or Germanic languages, but deeper than that applicable for finding the ‘phonemes’ of a language.

This is not just my interpretation of generative phonology and its offshoots. S. R. Anderson (2000) remarks:

> When we examine … early MIT theses in phonology such as those of Foley (1965), Harris (1967), Kiparsky (1965), McCawley (1965) and Schane (1965), as well as Halle’s own work on Russian and Latvian (e.g., Halle & Zeps 1966), there is a common thread apart from the focus on rules in descriptions. In each case, these works draw heavily on the existing results of historical linguistics, either directly (as in Kiparsky’s case) or as a source of antecedent analysis whose bearing on synchronic phonology is quite direct in a ‘morphophonemic’ context.

What this means is that modern mainstream phonology in the 20th c. and continuing into the 21st c. – insofar as it purports to discover the mental, i.e., the psychological mechanisms underlying speakers’ competence, – is a *scientific fraud*. Rather than uncovering what is in speakers’ heads, it is conducting the same kind of descriptive studies and using the same methodology developed in the centuries before the 20th – for the sake of establishing languages’ history, not their psychological representation in the brain.

On one occasion, Chomsky has admitted the empirical weakness of the program of research that he advocated:

> "... I do not think anybody actually working on language can doubt … that sooner or later … it is going to be necessary to discover conditions on theory constructions, coming presumably from experimental psychology or from neurology, which will resolve the alternatives that can be arrived at by the kind of speculative theory construction linguists can do on the basis of the data available to them. That is, there will come a point, no doubt, and I think in some area of linguistics it may already have been reached, where one can set up alternative systems to explain quite a wide range of phenomena. One can think that this or that system is more elegant and much deeper than some other, but is it right? … It seems to me that in phonology that point may have been reached."

This admission was made in 1967 at conference on brain mechanisms underlying speech when there was skepticism expressed by other participants about the basis for Chomsky’s claims about the psychological foundation of language competence. But more than 40 years have passed and mainstream phonology is still engaged in ‘speculative theory construction’. Rather, an earlier view of Chomsky’s (1964), made *à propos* of syntactic research, seems to have prevailed:

> … at the present stage of the study of language, it seems rather obvious that the attempt of gain some insight into the range of data that we now have is likely to be far more fruitful than any attempt to make this data more firm, e.g., by tests for synonymy, grammaticalness, and the like. Operational criteria for these notions, were they available and correct, might soothe the scientific conscience, but how, in fact,
would they advance our understanding of the nature of language or of the use and acquisition of language?

We can be grateful that Lavoisier in the late 18th c. did not adopt a similar view with respect to the weighing of the products of combustion – results from which led to an overthrow of the phlogiston theory of combustion and the ushering in modern chemistry.

5. WHAT TO DO?

As mentioned at the start of this paper, all scientific disciplines consist, roughly, of questions about the world, answers to these questions, and the methods used to evaluate the candidate answers. There is no argument that the program of inquiry about the phonological aspect of language that Chomsky (and his associates, notably Halle) initiated has been, to use a slangy term from the 1960’s, “consciousness-raising”. But methodologically it represents a step backwards from the impressive progress phonological studies had made previously. But what can be done to promote a truly scientific pursuit of the psychological basis of speakers’ ability to produce and perceive speech? To some extent the answer has already been given: there is already a strong tradition of psycholinguistic studies that illuminate what speakers know about the sound patterns in their language and what use they make of that knowledge: Meringer & Meyer (1895), Esper (1925), Greenberg & Jenkins (1964), Zimmer (1969), Ohala & Jaeger (1986), Ohala (1981), Ohala & Ohala (1995), Cutler (2005) to mention just a few relevant works. I do not claim that these works represent the truth: rather they show how questions of the psychological aspect of phonology can be evaluated in the empirical arena. The ascendancy of this tradition is essential for our field. It will not happen overnight: It is for the emerging generation not yet contaminated by the effete practices and delusions of the past that will usher in the new, stronger phonology.

6. ACKNOWLEDGEMENTS

I want to thank Mark Sicoli, Maria-Josep Solé, Yao Yao, and especially Natasha Warner, for giving me valuable assistance in the preparation of this paper. But the opinions expressed are mine.

7. REFERENCES


ten Kate, Lambert 1710 Gemeenschap tussen de Gottisch spreke en de Nederduytsc. Amsterdam, Jan Rieuwertsz.


Sajnovics, Janos. 1770. Demonstratio idioma Ungarorum et Lapponum idem esse. Tyrnavia.


Young, Thomas 1819. Remarks on the Probabilities of Errors in Physical Observations, and on the density of the earth, considered, especially, with regard to the reduction of experiments on the pendulum. Philosophical Transactions of the Royal Society of London 109:70-95.