SPECTROGRAPHIC ANALYSIS OF SPANISH SIBILANTS AND ITS RELATION TO NAVARRO'S PHYSIOLOGICAL PHONETIC DESCRIPTIONS

by
Gerald Dykstra

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in the
University of Michigan
1955

Committee in charge:

Professor Lawrence B. Kiddle, Chairman
Associate Professor Robert Lado
Associate Professor Gordon E. Peterson
Professor Charles N. Staubach

SPECTROGRAPHIC ANALYSIS OF SPANISH SIBILANTS AND ITS RELATION TO NAVARRO'S PHYSIOLOGICAL PHONETIC DESCRIPTIONS

by
Gerald Dykstra

The purpose of this study was to show some relationships between a portion of a physiological phonetic description of Spanish and a limited body of Spanish data as recorded by the sound spectrograph. This goal led to the complementary task of exploring the reason for the occurrence of non-compatible, though corresponding statements.

Informants and linguistic samples were selected in a manner which allowed, as nearly as possible, reference to the acoustic data as representative of Spanish as opposed to a single dialect of Spanish. Tomás Navarro Tomás' Pronunciación española was selected as representative of and outstanding among works on Spanish physiological phonetics. Comment on Navarro's descriptions was limited to his statements concerning the affricate and sibilant phonemes. The intention was to study precisely that acoustic material which was directly pertinent to the remarks of Navarro and to comment in terms which are communicative to a person with physiological phonetic orientation or training.

The results show the overwhelming mass of acoustic data to be of the type labeled "corroborative" or "sympathetic" in com-

parison with Navarro's statements. They also show considerable amounts of supplementary data. It is seen, for example, that Navarro's "voiceless" is very commonly the acoustic "energy of the type indicative of reduced vocal cord activity." There are also some data which appear to be of the type labeled "contradictory." The data indicate, for example, that Spanish voiceless stops are neither acoustically nor physiologically as long as the Spanish affricate.

The general conclusion is that physiological phonetic descriptions consist not solely of efforts to describe the activities of the speech producing mechanisms. They are done partially, or largely, on an impressionistic basis. In this situation, the limits of ear and nerve activity as well as phonemic bias of brain interpretations enter into the report. These facts, together with the facts concerning lack of simple one to one correspondences between sequent activities in the chain of speech events, serve to explain the occurrence of corresponding but contradictory data in the acoustic and the physiological reports.

PREFACE

In a scientific study, as in a study which deals with the humanities, it is impossible to estimate with any justifiable hope of accuracy the extent of the influence which any individual or school of thought or current fashion in thinking has had upon the investigator.

Though a full and just accounting of the sources of my views and interpretations as shown in this study is obviously not feasible beyond simply saying that they are rooted in the siftings of my total experience as I could best bring it to bear upon the problem at hand, I want and indeed feel impelled to mention certain influences which have been especially congenial to me.

Among those whom I can name here, an expression of sincere appreciation is due Professor Ernst Pulgram for his most welcome attitude of encouragement and support. Dr. Robert Lado was of inestimable help first in advising the use of statistical technique to heighten and sharpen the meaning of some otherwise relatively insignificant measurements and second in stimulating, during a series of luncheon meetings, what I feel to be the major ideas emerging from the study. Professor Charles N. Staubach's rigid insistence on a clear statement of purpose spurred me on to define with clarity the place of this study in the world of human research. I think, too, that his influence is largely responsible for my overall point of view that this is a spectrographic observation of part of the Spanish sound structure with the orientation of the physiological phonetician furnishing the basic framework for all the reporting.

Of particular importance in my total experience are the contacts with Professors Charles C. Fries and Kenneth L. Pike who, though not on

the doctoral committee, have contributed heavily to the content of this dissertation in the form of inspiration, outlook, and thoughts gleaned from their published texts, classroom discussions and private consultations.

Professor Gordon E. Peterson deservesfull credit for all laudable practices that may be found here concerning the gathering, recording and spectrographing of data. He gave freely of his time in directing and helping with all technical aspects of this dissertation from operation of the machinery through interpretation and format for presenting the illustrative spectrograms.

I wish finally to express most sincere thanks to Professor Lawrence B. Kiddle, chairman of my doctoral committee, for his kindness and for his readiness at all times to discuss matters pertaining to the preparation of this dissertation. He was perhaps the primary stimulus behind the choice of topic and the primary guide in its development. I shall hope that my work here may not make his efforts seem wasted.

TABLE OF CONTENTS

		Page
PREFACE		ii
CHAPTER I.	BACKGROUND OF THE STUDY	1
II.	PROCEDURES	11
III.	THE AFFRICATE SIBILANT /č/	17
IV.	THE SIBILANT /s/	65
V.	CONCLUSIONS AND IMPLICATIONS	135
APPENDIX	NAMES AND HISTORIES OF INFORMANTS AND LIST OF DATA	155
BIBLIOGRAPHY		159

CHAPTER I

BACKGROUND OF THE STUDY

John C. Steinberg's hand-drawn spectrogram[1] in 1934 was as futuristic as the Space Patrol stories of today but, of course, more intent on accurate description as befits a scientific paper. Space Patrol reflects a desired goal of travel to celestial bodies. The means developed or dreamed of consists of rocket ships. Steinberg's work similarly reflected a longfelt desire to convert speech into a visible display. The means dreamed of and later developed in the Bell Telephone Laboratories was the sound spectrograph.[2]

This machine, like others before it, presents visual stimuli which may serve as prompters to make an investigator look for causes that he might tend to overlook if supplied only with auditory stimuli. A small illustration of such a prompter with the resulting conclusion on cause is found in this summary statement: "tongue-height differences and frequency differences of formants 1 [formants are spectrographic representations of concentrations of acoustic energy] for different vowels do not correspond when comparing articulatory triangles with acoustic triangles. If we look for some feature that is nearly the same for [i] and [u], and for [e] and [o] [acoustically, formant 1 is similar for [i]

---

[1]John C. Steinberg, "Application of Sound Measuring Instruments to the Study of Phonetic Problems," Journal of the Acoustical Society of America, IV (July, 1934), 23.

[2]Available technical descriptions include: Bell Telephone System Technical Publications, Monograph B-1415; R. K. Potter, G. A. Kopp and H. C. Green, Visible Speech, (New York: D. Van Nostrand Company, Inc., 1947); Martin Joos, Acoustic Phonetics, Language Monograph No. 23, XXIV (1948).

and [u] and again for [e] and [o]] in the articulation, we find it better in the overall opening of the mouth tract, as measured for instance by the distances between the upper and lower incisors or by the distances between the highest point of the tongue and the point of the palate closest to it - in other words, the general width of the strictures. Either of these two measurements show only small differences between corresponding front and back vowels. The use of the term opening instead of tongue-height seems indicated."[1]

The preceding statement gains meaning in the light of tongue-height measurements made by Jones, Kenyon, Parmenter, Carmody, Millet and Russell. If we arbitrarily assign no height or zero units of height to their tongue-height measurements for [a], as comparatively illustrated by Delattre[2], and six units to [i] we find that the average tongue-height for [u], according to these six sources, is only 3.5 units. Apparently, however, the "opening" for [i] and [u] is the same.

It is recognized that great care must be taken in determining the trustworthiness or the usefulness of our interpretations. "For any given unit characteristic of the auditory perception, such as the simple identification of a phoneme, the spectrogram will very often exhibit several features which are distinctive to the eye, and the information which can be obtained from the spectrogram is, accordingly, ambiguous. Even when only one feature or pattern is strikingly evident, one cannot be certain about its auditory significance; unless he assumes that those

---

[1]Pierre Delattre, "The Physiological Interpretation of Sound Spectrograms," PMLA, LXVI (September, 1951), 867-68.

[2]Ibid., 866.

aspects of the spectrogram which appear most prominently on visual examination are, in fact, of greatest importance to the ear."[1]

One of the most necessary considerations for those scholars who are interested in language as a device of communication between separate nervous systems is the consideration of human reaction or "auditory significance."

In spectrographic interpretation adjustments of various kinds are made.[2] These constitute attempts at correlating tangible, measurable observations of recorded activity with perceptual behavior which is relatively intangible and not as extensively measurable at this time.

A comparison of the reactions of that master of Spanish phoneticians, Navarro Tomás, to a limited body of acoustic energy, and my reactions to spectrographic interpretation of a related, but different body of acoustic energy, constitutes the major concern of this study.

The motivation behind the study is personal interest.

The intent of the study is to explore a limited field of acoustic phenomena in Spanish with one setting of the sound spectrograph.

The first goal is to find data which may seem to corroborate, supplement, or contradict Navarro's comprehensive description[3], which is the most generally accepted[4] presentation of Spanish sounds.

---

[1]Franklin S. Cooper, Alvin M. Liberman and John M. Borst, "The Interconversion of Audible and Visible Patterns as a Basis for Research in the Perception of Speech," Proceedings of the National Academy of Sciences, XXXVII (May, 1951), 318.

[2]Martin Joos, Op.cit., 131.

[3]Tomás Navarro Tomás, Pronunciación española, (4th ed.; Madrid: Imp. de la Lib. y casa edit. Hernando, s.a., 1932).

[4]Emilio Alarcos Llorach, Fonología española, (Madrid: Editorial Gredos, 1950), 93. And G. L. Trager, "The Phonemes of Castillian Spanish," TCLP, VIII (1939), 217-222.

A further goal, which developed during the course of the study, is that of explaining the reason for the appearance of the varying physiological and acoustic interpretations loosely referred to above as contradicting data.

The striving toward this further goal lends coherence to the dissertation and results in its main contribution. It demonstrates the relevance of acoustic speech data to human reaction. It shows why this acoustic speech data cannot be described effectively by ear and how we learn of the unobservable by studying the observable as did Newton who spoke of unobserved gravity by studying its effects on observable objects.

Still more important, the further goal forces a focus upon the variations in acoustic speech data which do not correlate on a simple one to one basis with variations in the vocal organ movements or variations in reaction behavior on the part of a listener. And because this is taken into consideration a coherent logical explanation converts to deeper understanding that which would otherwise have to be considered as contradictory material in Navarro on the one side and in the data presented here on the other.

Navarro's work is here taken as the outstanding work of its kind. The assumption is made that all his statements are true from at least one point of view, thus in every instance forcing a search for explanation by careful examination of the dynamics of speech organs, air vibrations, and human behavior patterns.

The field of Hispanic studies is deeply indebted to Navarro Tomás for the description used here and for others of his truly great works. It should be expected that the work of so great an authority on Spanish

speech is not to be easily superseded but rather that it is to stand with, at most, small modifications and some additions. It is accordingly considered justifiable to assume the basic validity of his statements as they apply to vocal organ movements and human behavior or reactions.

This could similarly be true of the work of a man of great acumen who might have synthesized all his own and previous observations on, let us say, cloud formations. Yet with the invention of the camera and aircraft it might be found highly useful to re-study cloud formations since with new equipment it could be done from several angles and distances at once. Moisture content and depth could be measured, too. And still all of this would probably change only slightly the descriptions of the naked-eye ground observations of cloud formations, if those observations were scientifically made.

The important point illustrated here is that eye reaction and ear reaction remain human reactions. If they constitute the core of our interest, then the mechanical reactions of light waves upon light sensitive film, or of air molecule motion indirectly upon electrically sensitive paper, can tell us only what is impinging upon the eye or ear. Such representations as the spectrogram may show sound components not previously perceived. A small segment of these may subsequently be detectable by the trained ear; but all of them add to our understanding of the phenomena in question. Thus, if I hear [s] and [f], and learn that a mechanical reactor shows the one recurrently to be composed of noise in a higher frequency range than the other, I may possibly hear a pitch difference of which I was not consciously aware before, and I will probably feel that I understand just a little more about [s] and [f].

Since a goal of this study is to look at spectrographic data and impressionistic data in juxtaposition, it will provide thorough spectrographic descriptions. It will include facts and figures on, for example, the frequency ranges, in cycles per second, of a given sibilant in various environments whenever it is felt that these facts and figures apply as pertinent commentary on Navarro's statements.

It is recognized that speech consists of a considerable chain of events, which includesat least the brain activity of the speaker, neural impulses from the brain, vocal organ muscle and tissue movements, air molecule activity, ear drum and hair cell movements, neural impulses to the brain, and receptive brain activity. It is recognized also that speech can,in theory, be studied with equal legitimacy at any one of innumerable points along this chain.[1]

A corollary of this, which was not focused upon until the research work was largely completed, is that there is no valid reason to assume a simple one to one correspondence of activity all along the chain. In fact, Peterson has shown that "The degree of correspondence which can be observed decreases as extraneous noise increases in level to obscure the signal."[2] Any point or area along the chain of events in the speech act may legitimately be studied from a scientific point of view; but the point must be specified, and the findings in a study of one point cannot, ipso facto, supersede descriptions which were made at another point or area.

---

[1]This idea of speech as a sequence of events and the problem of practicability and validity of studying the speech sequence at one or another point along the line was foreshadowed at least as early as 1904 in a chapter entitled "Akustisch oder Genetisch?" Otto Jespersen, Phonetische Grundfragen (Leipzig: B. G. Teubner, 1904), 65-104.

[2]G. E. Peterson, "Systematic Research in Experimental Phonetics: 4. The Evaluation of Speech Signals," Journal of Speech and Hearing Disorders, XIX (June, 1954), 159.

Comprehensively speaking, the goal of this study will be to describe data at one point in the chain of speech phenomena. These descriptions will be compared with Navarro's descriptions. This goal covers the finding of corroborating and supplementing materials and is expanded to include explanation for what would be refuting data if a simple one to one correspondence were assumed for behavior between all points on the speech chain.

The apparent inconsistency so far is that if no acoustic data is to be considered to refute or contradict Navarro, then no data can be said unqualifiedly to support or corroborate his conclusions.

To the extent that Navarro may be considered to restrict himself to statements which concern his impressions, it is a theme of this dissertation that acoustic study cannot tell anyone what Navarro's impressions or perceptions were. To the extent that it is not manifestly clear in his statements whether he is describing his impressions as received by ear, his impressions of x-rays or other recordings of vocal organ movements, or his beliefs concerning air vibrations, it is impossible to do more than to make comparative statements. For even when Navarro says, for example, that /č/ is no longer than the /p/, /t/ or /k/, we cannot know beyond doubt that he is speaking of air wave motions; though, in truth, it may seem that he is implying there is no difference between the lengths at any point along the chain of events in the speech act.

When, however, it is found that a description of air wave motion coincides quite well with Navarro's description it is felt that there might be a causal relationship. Although it is fully recognized that Navarro's statement has not thereby been proved, it is seen that there

are interesting correspondences between at least two points along the chain of speech events. It is accordingly referred to as corroborating data though perhaps a better term would be "sympathetic data."

Conversely, it cannot be said that non-sympathetic data in air vibration proves false all the data concerning physiological movements or behavioral reactions, because such a statement rests upon the provably false assumption[1] that there exists a simple one to one cause and effect relationship, now predictable, between the various points along the speech chain. This thesis maintains that with our present limited knowledge we must study the various points one by one and gradually set up correspondences which we may eventually be able to predict to a very considerable extent.

Finally, it would be undesirable to assume a one to one relationship precisely because it would justify statements concerning the great accuracy of the spectrograph machine and therefore its reliability or superiority in determining certain features of behavioral reaction as well as in determining simpler mechanical reaction. There would be no appeal from the verdicts of measurable spectrograms. It seems much more conducive to understanding if instead we ask further: "What is the explanation for the divergent descriptions that appear?"

It is assumed that it is worthwhile to learn things that are of interest. The subject is of interest to many people who assume that linguistic forms are partially similar in form and meaning. Interest alone is sufficient to warrant study.

Furthermore, matters that are at first of purely academic interest, very often become the building blocks for practical inventions, such as

---

[1] Martin Joos, Op.cit., 33, 92, 117.

telephones and release of atomic energy, another day.

Detailed study of language, its forms and its changing, contribute to a number of disciplines. Morris Swadesh, for example, has set forth a provocative formula for determining the separation date of related languages.[1] The study is called glottochronology and its impact could be great in the field of anthropology. In the field of education it has been found that linguistic comparisons have proved useful in teaching and learning a second language.[2]

A tie-in with the fields of psychology, philosophy and literature is also readily found in the study of language function. It will become increasingly obvious that this report, for example, is designed to bring out the sharply limited relevance of infinite variation in physical data to human behavior. It is necessary to keep an open mind concerning the border area between the speech variations which obviously correlate with behavioral variations and those which do not. It is in this area that we are likely to find those nuances of speech which we refer to when we say that one person's voice is unpleasant while another's is mellifluous. It is here, as well as in the field of gesture, that we must look for the differences that make one artist's reading of a poem seem unartistic and cause a similar reading to be acclaimed a true artistic production under the skills of another.

---

[1]Morris Swadesh, "Lexico-statistic dating of prehistoric ethnic contacts," Amer. philosophical soc. Proceedings, XCVI (1952), 452-63.

[2]David W. Reed, Robert Lado and Yao Shen, "The Importance of the Native Language in Foreign Language Learning," Language Learning, I (1948), 17-23.

One man cannot find all the answers, but thousands of investigators over a long period of time have found, are finding, and will continue to find facts which, put together in the fertile minds of trained or gifted people, produce great insights into behavior.

In summary, it may be said that the chief purposes of this dissertation are (1) to show as far as possible, within a limited field of linguistic data, the relationships between Navarro's physiological description of Spanish and data found by spectrographic analysis, (2) to point up examples of spectrographic data which indicate the presence of activity not deducible from Navarro's statements concerning articulation and vice versa, and (3) to explore the reason for the non-compatible, though corresponding statements.

CHAPTER II

PROCEDURES

It was considered essential to limit the data strictly and to explore a maximum of implications rather than solely to accumulate masses of unstructured information. This was accomplished partly by limiting comment on linguistic matters to the Spanish sibilant phonemes.[1] It was implemented by imposing strict limits on the data to be used from the physiological phonetic descriptions of Spanish and also from the acoustic phonetic material which could have been made.

The limit on consideration of physiological phonetic descriptions that were to be used was realized by choosing the single outstanding work of Navarro Tomás. This had the further advantage of eliminating the necessity for discussion of disagreements among phoneticians which, though perhaps interesting in the light of spectrographic analysis, remain relatively insignificant to the purpose of this dissertation.

Navarro's text was carefully combed for descriptive statements concerning sibilants and their influences in neighboring sounds.

Words were chosen from lists, given by Navarro as illustrative of his descriptions, in such a way as to give a representative sampling of phonetic environments for the sounds he described. These words were then included in 18 sentences[2] which were constructed and approved as possible sequences in Spanish by three native speakers of the language.

---

[1]The term "sibilant phonemes" is used, in this dissertation, to include not only /s/ but also the affricate phoneme /č/, which regularly includes a sibilant segment. It excludes sibilant allophones, submembers of other phonemes, that may be found, for example, in some pronunciations of words like yo and correr.

[2]The sentences are reproduced in the Appendix, pages 157 and 158.

They include 247 spots in which there is a likelihood of occurrence of sibilants or other fricatives made at or near the same physiological point of articulation.

The limitation on acoustic phonetic data was accomplished in part by considering only spectrograms produced by one commonly used setting[1] of the spectrograph machine.

It was, of course, impossible to use the same air waves that Navarro did. It was also impossible to secure, for recording, the same informants that Navarro used. And, since he does not tell us which informants he used for which material, it was not possible to simulate the data that he used for each of his descriptions.

The data from the single setting of the spectrograph machine were intended therefore to include a representative sampling of one social dialect of Spanish as found throughout the major portion of the Spanish speaking world. The social dialect chosen is that of educated speakers of Spanish in normal rapid conversational type speech.

It would have been desirable to use the same data that Navarro used for a study of this kind. Since that was wholly impossible, it was decided best to speak of Spanish as a whole in comparison with Navarro's description, thus testing his statements thoroughly yet fairly. Therefore speakers were chosen from a number of geographical areas representing major dialect divisions. Since it was considered desirable, for comparison with Navarro's work, to speak of Spanish in general rather than of an idiolect or dialect, it was considered fortunate that the selection of informants reflected the relative density of population in these areas fairly well.

---

[1] Full band control and the 300 cps. filter.

The refinement over Navarro's presentation is that speakers, linguistic contexts and situations are always known.

The speech of twenty-one educated informants was studied. The selection of these informants was based upon a desire to have a token, if inadequate, coverage of the speech of most of the major dialect areas. This was for the purpose of having a body of material with enough diversity to find some parallelism of description with Navarro's general physiological statements whenever possible since the project aims at enriching with increased meaning his articulatory descriptions.

In this study, therefore, Latin America is represented by eighteen informants who come from the five dialect areas established by Pedro Henríquez Ureña[1] on the basis of the indigenous languages of the areas. There are six representatives from the Nahuatl area, three from the Lucayo area, five from the Quechua area, one from the Araucano area and three from the Guaraní area. It happens that there is one representative for approximately each five million inhabitants of these areas. Spain, however, is represented by only three informants: one from Burgos, one from Madrid and one from Seville.[2]

For the purpose of preliminary recording of the data the informants were accompanied to a sound proof radio recording room. They were asked to speak as they would at home with good friends. In addition they were given a handwritten instruction sheet which read: "Speak at normal conversational speed and with normal pronunciation as you would with friends at home."

---

[1]Pedro Henríquez Ureña, "Observaciones sobre el español en América," Revista de Filología Española, VIII (1921), 357-390.

[2]Histories pertinent to the linguistic background of each informant were recorded and are reproduced in the Appendix on pages 155-157.

Every effort was made to have the situation appear informal. There was casual conversation and an opportunity for practice with the microphone and the list of sentences in every case before the recording took place. This allowed for adjustment to the situation on the part of the speaker and for adjustment of volume mechanisms to the speaker on the part of the recording team.

The body of sentences was recorded twice by each informant.

The name and country of the informant, the number of the tape reel, the recording number, the date and the time were read into the tape recording preceding each repetition of the sentences.

All recording was on tape with the Ampex high fidelity tape recorder.

The recordings were monitored and it was the opinion of the author and two native speakers of Spanish (informants XIV and XV) that the sentences sounded like normal unaffected Spanish. Therefore the first recording of each sentence by each informant was chosen as the data to be recorded on the spectrograph. The relatively few exceptions to this rule consisted of readings which included excessive stuttering or excessively long pauses. In these cases it was the second reading that was considered the more desirable as material to be spectrographed.

The spectrograph machine itself was set on full band control and the 300 cps. filter was employed.

Twenty-six to thirty spectrograms for each informant were found sufficient to cover the data of interest for this study in the 18 sentences.

These spectrograms were to become the core of data which were to be studied in reference to the remarks of Navarro, which center about a physiological production point of view.

After the experiment was thus set up and everything was recorded, the data were thought of as constituting a series of spots with potential for occurrence of various bits of behavior that language students have come to call phonemes, allophones, sounds, etc.

Fricatives and surrounding environments in all the data were studied and measured as a step toward gaining general acquaintance with the data. Navarro's text was re-studied for the purpose of finding statements on physiological phonetics which allow parallel or similar statements in acoustic phonetics. These statements were taken from Navarro's text and subjected to comparison in the acoustic field. Prerequisites to this were, first, thorough study and measurement of details of the spectrographic markings and, second, statistical procedures to provide a basis for statements of reliability and predictability as deduced from the sampling.

There was always a conscious effort to explain all matters with commonly used terminology from the field of physiological phonetics. This precluded an exhaustive treatment of speech phenomena on the basis of spectrographic analysis alone. The intention and the attempt was to study and comment on precisely that acoustic material which was directly pertinent to the remarks of Navarro and understandable to a person with physiological phonetic orientation or training. Thus, for example, over 1000 measurements and over 100 additions and averagings plus some full scale statistical procedures were performed in connection with Navarro's statements concerning pronunciations corresponding to *x*.[1] Several times that number were performed in connection

---

[1] Navarro's italics, representing written forms, are here shown by underlining. His bold face type, representing phonetic transcription, will here be enclosed in brackets.

with pronunciations of the affricate /č/. These were reported on insofar as they showed or failed to show, in the field of acoustic phonetics, parallelism of description with the field of physiological phonetics.

This rigid limitation of comment to phenomena directly pertinent to Navarro's work carries to the point of following, in the division of topics, his spelling oriented approach to pronunciation. Thus, for example, although it may be undesirable for some purposes, pronunciations are, in this dissertation, freely spoken of as "corresponding to x in the written language."

CHAPTER III

THE AFFRICATE SIBILANT /č/

For the purposes of a comparison of spectrographic analysis with the statements made by Navarro concerning Spanish /č/, a somewhat unconventional format suggested itself. In this format, Navarro's statements appear first. They are accompanied by superscript capital letters used to signify specific portions of the text that will be commented upon in appropriate succeeding sections. Accordingly all words or phrases marked with a superscript A will be treated in section A where these words and phrases will also reappear, but not in the full context which is available here for reference.

The superscripts refer to words and phrases which can be outlined as in the following key:

A. Points of articulation	H. Variation
B. Affricate nature of /č/	I. Timbre
C. Voiceless nature of /č/	J. English [š] and [tš]
D. Gradual vs. abrupt movement	K. French [š]
E. Length of /č/	L. Italian [tš]
F. Phonetic similarity	M. German [tš]
G. Simplicity	N. Sharpness

"Palatal[A] africada[B] sorda[C] .....

".....el predorso[A] se separa gradualmente[D] de los alvéolos[A] y del prepaladar[A], formando con éstos durante un instante[E] una estrechez por donde el aire se escapa, produciendo una breve[E] fricación[B], semejante[F] por su timbre[I] al sonido de la sh inglesa[J]. Tanto esta fricación[B], como la oclusión[B] que la precede son momentáneas[E], y se efectúan entre los mismos órganos[A] y en el mismo punto del paladar[A];

su duración[E] total viene a ser como la de cualquier oclusiva simple[G]....

"En pronunciación dialectal[H], la articulación de la ch española presenta multitud de variantes[H].....

".....el elemento fricativo[B] de nuestra ch, aun careciendo de la labialización que caracteriza al sonido francés[K], tiene, en efecto, cierta semejanza[F] con dicho sonido;.....Tienen sonido más o menos semejante[F] a nuestra ch, la c en ital. cento[L]; la tx en cat. butxaca; la ch en ingl. church[J], y la tsch en alemán deutsch[M]; pero la parte fricativa[B] del sonido español es más breve[E] y más aguda[N] que la que generalmente presenta dicha articulación en los demás idiomas citados."[1]

A. Points of articulation: "Palatal"; "predorso"; "alvéolos"; "prepaladar"; "mismos órganos"; "paladar".

Direct observation, palatograms, kymographs and x-rays have long served useful purposes in speech study and the advent of the spectrograph by no means invalidates other approaches. This becomes especially clear if the investigator's purpose is to determine with precision the points of articulation of specified portions of speech behavior. It is impossible, for example, to state that either the frequency characteristics of the friction or the shape of following and preceding vowel formants can be used to determine the point of articulation of the [š] in Spanish /č/ with the degree of accuracy that can be attained with palatograms.

---

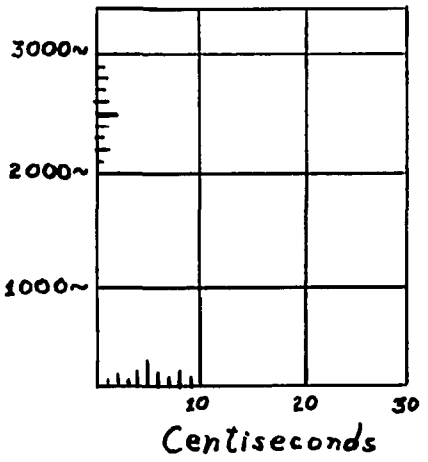
[1]Tomás Navarro Tomás, Op. cit., 125-27.

Friction caused by air directed against the teeth through a narrow opening between the alveoles and prepalate above, and the predorsal portion of the tongue below, appears at a characteristically high position on the spectrogram[1], usually heaviest above 2000 cycles per second[2], as in figure 1. Friction caused by air passing through a narrow passage between the velum and the back of the tongue shows a characteristic position which is lower; usually heaviest below 2500~as in figure 2. But figure 3 demonstrates the variability of these frequency characteristics as it corresponds to alterations in lip position or cavity size. Since neighboring sounds inevitably affect the movements of the vocal organs, including lips and cavities, it is obvious that differences in sound environment will affect the characteristic position of the friction on the spectrogram. It follows from this that variations in spectrographic marking do not necessarily indicate a shift in the so-called point of articulation and similarly that an actual shift in the point of articulation may be compensated for by a change in cavity shape

---

[1] R. K. Potter, G. A. Kopp and H. C. Green, Ibid., 54 and 143-49.

[2] The term "cycles per second" will, on subsequent occasions, be abbreviated as ~ . The following diagram illustrates both frequency and time scales used in the figures:





[š ɛ]

Figure 1. The first portion of the English word "share" as spoken by informant XXVI is spectrographically represented here.[1] The relatively high frequency of nearly all the friction in [š] is apparent.

in such a manner as to obviate precise representation of the point of articulation. Since the fricative sounds have so much free and conditioned variation in the shapes of the resonating cavities, the supplanting of palatograms and x-rays by spectrograms for the purpose of accurately determining point of articulation is not expedient at present.

If we broaden the definition of "point of articulation" to include characteristics of the resonating or damping chambers we can expect spectrograms to be looked upon as sources of valuable information in determining these characteristics.

---

[1] The informant and the linguistic material will be identified in the descriptive text following each figure. The script under the illustrations is supplementary and will be given in phonetic, phonemic or traditional orthography according to which of these seems most appropriate. Phonetic notations are enclosed in brackets; phonemic, in slant lines.



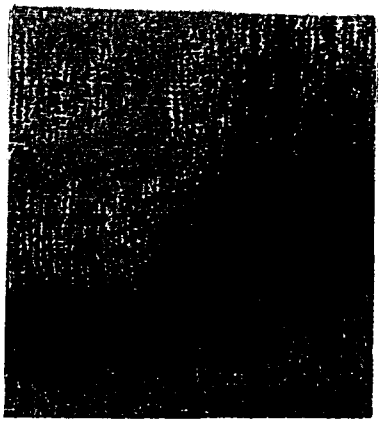
[×]

Figure 2. Friction caused by air passing through a narrow opening between the back of the tongue and the velum is seen to be stronger in frequencies which are lower than for [š]. This sample was produced without linguistic context by informant XXVI. In the context of Spanish words like juzgar, embajada and cónyuge, the voiceless velar fricatives given by the Spanish speaking informants consisted of friction which was most intense at even lower frequencies than in this sample. In the Spanish data the highest frequency velar friction occurred regularly in the word embajada where formant 2 of the contiguous /a/ phonemes is often seen to continue, at the same height, as a formant throughout the voiceless velar fricative.

B. Affricate nature of /č/: "africada"; "fricación"; "oclusión"; "elemento fricativo"; "la parte fricativa".

Nineteen of the twenty-one informants pronounced four clearly affricate segments when presented with "muchacho" and "salchichón" in the prepared scripts. One informant pronounced two clearly affricate segments and two segments in which the occurrence of a stop element is doubtful. The remaining informant pronounced one affricate and three non-affricate sibilants.

The spectrograms of the pronunciations of Spanish /č/ show eighty-one portions of blank or near blank followed by irregular blackness which is heaviest in the top half of the spectrograms and often occurred there



[x̆ x]

Figure 3. The frequency characteristics of velar friction may change considerably without a change in the so-called point of articulation. A shift from rounded to unrounded lip position is recorded here as produced by informant XXVI. The unrounded lip position is the same as that used in the production of figure 2.

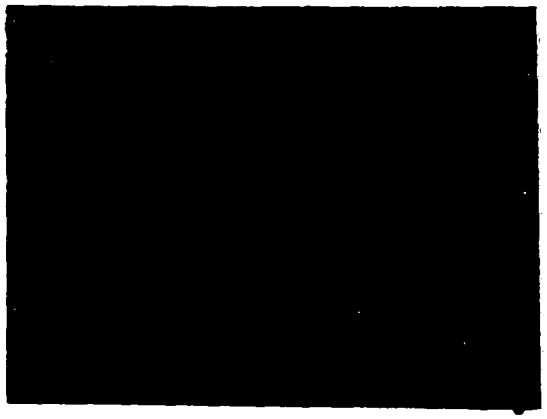
only. These segments are interpreted as stoppages of air flow followed by resumptions of egressive air flow causing friction.

All eighty-four occurrences of /č/ show some friction varying in intensity between the extremes illustrated in figures 4 and 5. Nearly



[s a l t š i t š ó n]

Figure 4. This represents the word salchichón as pronounced by informant I in the context of sentence 3 (Appendix, page 157). The first /č/ illustrates an affricate type which occurs with almost total lack of friction.



[m u t š á t š o]

Figure 5. Informant VIII supplied samples of especially high intensity friction in the fricative portions of the /č/ phonemes in the context of sentence 1.

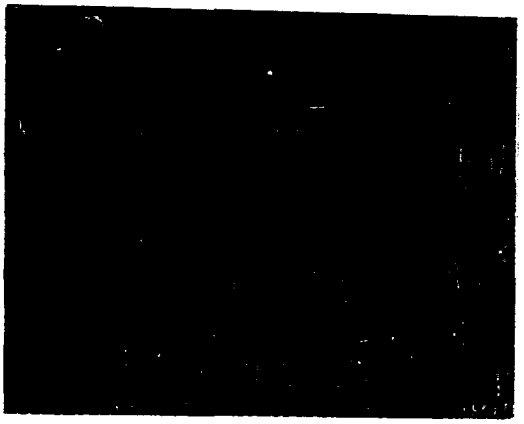
all of them resemble the high intensity friction of figure 5 more nearly than they resemble the low intensity friction of figure 4.

The situation is quite different with respect to the stop portion. In fact the variant pronunciations which may fill one of these spots occasionally forces this speech segment out of the phonetic classification of "affricate" as illustrated in figure 7. Figures 6 and 7 present



[s a / t š : t š ó n]

Figure 6. Informant VIII, sentence 3. The second /č/ contains the longest stop element in the data for this phoneme.



[sa / š i š ó n]

Figure 7. Informant XXI, sentence 3. The word salchichón, though regularly found with two occurrences of the phoneme /č/, is here shown with absence of any affricate portion.

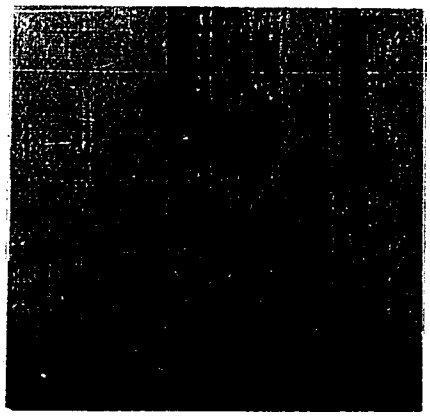
evidence of "stop" portions of the "affricate" varying in duration from a high of 8 centiseconds (fig. 6) to complete absence (fig. 7).

C. Voiceless nature of /č/: "sorda."

As with the other characteristics, the spectrograms show a wide range of possibilities for vibration or lack of vibration of the vocal cords in the speech spots that are under attention in this chapter.

Of the eighty-four potential spots in which we might expect vocal cord vibration to be absent, i.e., during the spans of time in which the phoneme /č/ was being pronounced, forty show from two to six centiseconds in which no evidence of any vocal cord activity was recorded. The other forty-four show evidence of varying degrees of vocal cord movement throughout the affricate span.

This does not mean that these pronunciations were made with maximum vocal cord vibration throughout, not that "voicing" would be perceived throughout these pronunciations. The very light vertical striations in figure 8 probably represent vocal cord activity which is only partial.



[m u t š á d š o]

Figure 8. Informant VI, sentence 1. The first /č/ shows a five centisecond span of very light vertical striations at the base.

It is called partial in the sense that the glottal lips approach and recede but do not come near enough to each other to cause the strong variation in pressure which is perceived as "voicing."

The heavy blackness which almost eliminates the vertical blank sections or lines in the stop portion of the first affricate of figure 9 indicates a great deal of turbulence or noise which recurs with each

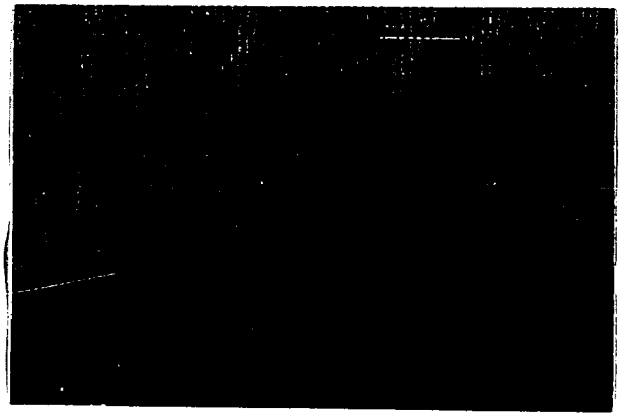


[m u d š á t š o]

Figure 9. Informant V, sentence 1. Illustrated here is a "voice" bar extending through the stop element of the first /č/. The same phenomenon is seen in figure 8 in the second /č/ of the same word.

release of the vocal cord lips. This quantity of turbulence in the vocal cord area is equivalent to the amount that is found in clearly perceived "voicing" in voiced stops as shown in figure 10.

Figure 9, then, shows a sound which has a stop portion with vocal cord vibration as strong as in voiced stops and a fricative or sibilant portion which lacks strong vibration. Therefore the ordinary phonetic symbolization should be [dš].



[s o l d á d o]

Figure 10. Informant VII, sentence 2. The first /d/ is a clearly voiced stop.

Figure 11 illustrates the effects of strong periodic air turbulence at the vocal cords or, in other terms, clear vocal cord vibration with complete or almost complete opening and closing of the vocal cords throughout the duration of the affricate. The clearly indicated ordinary phonetic notation is [dž].

D. Gradual vs. abrupt movement: "gradualmente."

The spectrographic data indicate that the separation of the tongue from the roof of the mouth does proceed relatively gradually in about half of the cases under study. Figure 12 is a sample of that data in which evidences of friction are first indicated weakly in the very high



/s a/ č i č ó n /

Figure 11. Informant VI, sentence 3. The voicing bar continues throughout the duration of the first /č/.

frequency range of the spectrogram.[1] As time progresses the friction spreads to lower and lower frequency levels. This indicates that some characteristics of the resonating chamber are indeed changing. It in-



[s a / d t š i d t š ó n]

Figure 12. Informant VII, sentence 3. Sibilant noise first appears in the high frequencies in this sample and spreads over a wider range as time progresses.

---

[1] If the machine indicated the presence of energy in frequencies higher than 3600~ it might well be found that evidences of noise or friction were present even earlier in time.

dicates also that the change is not instantaneous and it can be presumed further that the tongue is probably playing an important role in the relatively gradual changes that are taking place.

There remain, however, a large number of cases in which there is nothing gradual about the onset of the friction. The initial spike,[1] contiguous with the following symptoms of noise or friction in figure 13, indicates a sudden burst of acoustic energy and this is here associated with a relatively rapid separation of the tongue from another surface.



[m u t š á d š o]

Figure 13. Informant IV, sentence 1. Here sibilant noise is seen to begin simultaneously over a wide range of frequencies.

The immediate appearance here of friction at all frequencies, found in figure 12 only after the tongue has had ample time to move some distance, argues also for the view that in roughly half of the data the separation of the tongue from the superior surface is not gradual but is sudden and explosive.

---

[1]This term is borrowed from Potter, Kopp and Green, Op. cit., 33, 34.

E. Length of /č/: "instante"; "breve"; "fricación...oclusión...momentáneas"; "duración...como la de cualquier oclusiva simple"; "más breve... que...en los demás idiomas..."

It has become evident that most attempts at a sharp segmentation of physical phenomena have not been in accordance with the actual fact.[1] In this study, it is again very clear that any portion of linguistic behavior influences contiguous portions to such an extent that it is extremely difficult to make clear cut divisions between the various physical manifestations of phonemes or portions of phonemes which occur next to each other in time. An attempt at accurate measurement of the length of the segments as shown on the spectrograms quickly convinces one of the rather arbitrary nature of the exact placement of the lines that are to serve as beginning and ending points for the measurement.

The border characteristics of the [š] and [i], for example, may overlap to such an extent that it becomes impossible to draw a dividing line without a margin of error or disagreement which comes to one or two centiseconds or even more. Still it is clearly possible to point to portions of the spectrogram which are characterized predominantly by markings we have come to associate with [š] or [i] and it is necessary to conclude that the vague border lines on the spectrograms reflect vague borderlines in the reality of speech.[2]

Yet we want to measure length because of the desire to make a statement, concerning spectrograms, that will be meaningful in the light of Navarro's statements concerning length. Certain abritrary criteria can

---

[1] Kenneth L. Pike, Language, (in preliminary format lithoprinted by Braun and Brumfield, Ann Arbor, 1954) Chapters I, IV, and V.

[2] Kenneth L. Pike, loc.cit.

be set up in such a way as to give analogous starting points for the measurement of all stops and affricates. In this problem, measurement began at the point where all evidence of formants 2 and 3 ceased. Measurement ended, (1), in the case of the stops followed by a vowel, at the point where any evidence of formants 2 or 3 superimposed upon regularly recurrent vertical striations was resumed; (2), for stops followed by fricatives, at the point where evidence of friction was present; (3), for the affricates, at the point where there was no more evidence of [š]-like friction in the high frequency range.

The effect of these rules was to tend to allot relatively more length to the stops than to the affricates because the beginning of formants 2 or 3, which constituted the end measurement of the stops, sometimes did not begin until several centiseconds after the last traces of high frequency [š] friction were recorded.

This was an effect that was both desirable and necessary. It was desirable because Navarro had said that the length of /č/ is like that of the simple stops. A series of early measurements had indicated that, acoustically, there might be a significant difference. If the early guess were later borne out in fact, it would be a more clear cut case if it could be shown that the rules of measurement seemed to assign less than due length to the occurrences of /č/. But it was also a necessity to set up the rules in this way. It was obvious that irregularly striated fill, which usually indicated voiceless aspiration after the stops, and which continued to the onset of the following vowel, had to be included as part of the stops rather than of the vowels. It was equally obvious that the measurement of affricates, which did not always show friction noise continuing to the onset of the vowel, had to stop at

the point where no more evidence of [š] friction was present. The reason was that, in the word "salchichón", the [i] was very often so much reduced in speech that formants 2 and 3 did not appear at all. And, of course, the onset or discontinuance of voicing could not be considered as a suitable criterion because many of the potential spots for "voiceless" stops and affricates contained evidence of vocal cord vibration throughout.

Statistical analysis relevant to the length of /č/ and its segments shows that /č/ as a whole is significantly longer than any one of the voiceless stops. Figure 14 shows the distribution and lengths of four stops and four affricates each of which occurred twenty one times in the data. Figure 15 is a composite of the eight histograms in figure 14. Whereas the average length of /č/ in the data is 10.7 cs., the average length of voiceless stops is 8.5.

The mean length of the first /č/ in muchacho in the sampling is 11.1 cs. Statistical procedures[1] show that two-thirds of the means of like samplings will not vary from 11.1 cs. by more than .5 cs. and that 95 per cent of the means of like samplings will fall between 10.2 cs. and 12 cs. This compares with a mean length, in the sampling, of 9.3 cs. for the second /č/ in muchacho.

Further statistical procedures show that there is a probability of less than one in a thousand that this difference is due to chance. It seems to be a significant difference of length which can be correlated with stress and intonation and possibly with the position in the utterance of the syllable in which /č/ appears. It would appear that,

---

[1] J. P. Guilford, Fundamental Statistics in Psychology and Education, (1st. ed.; New York and London: McGraw-Hill Book Co., 1942), 135 ff.

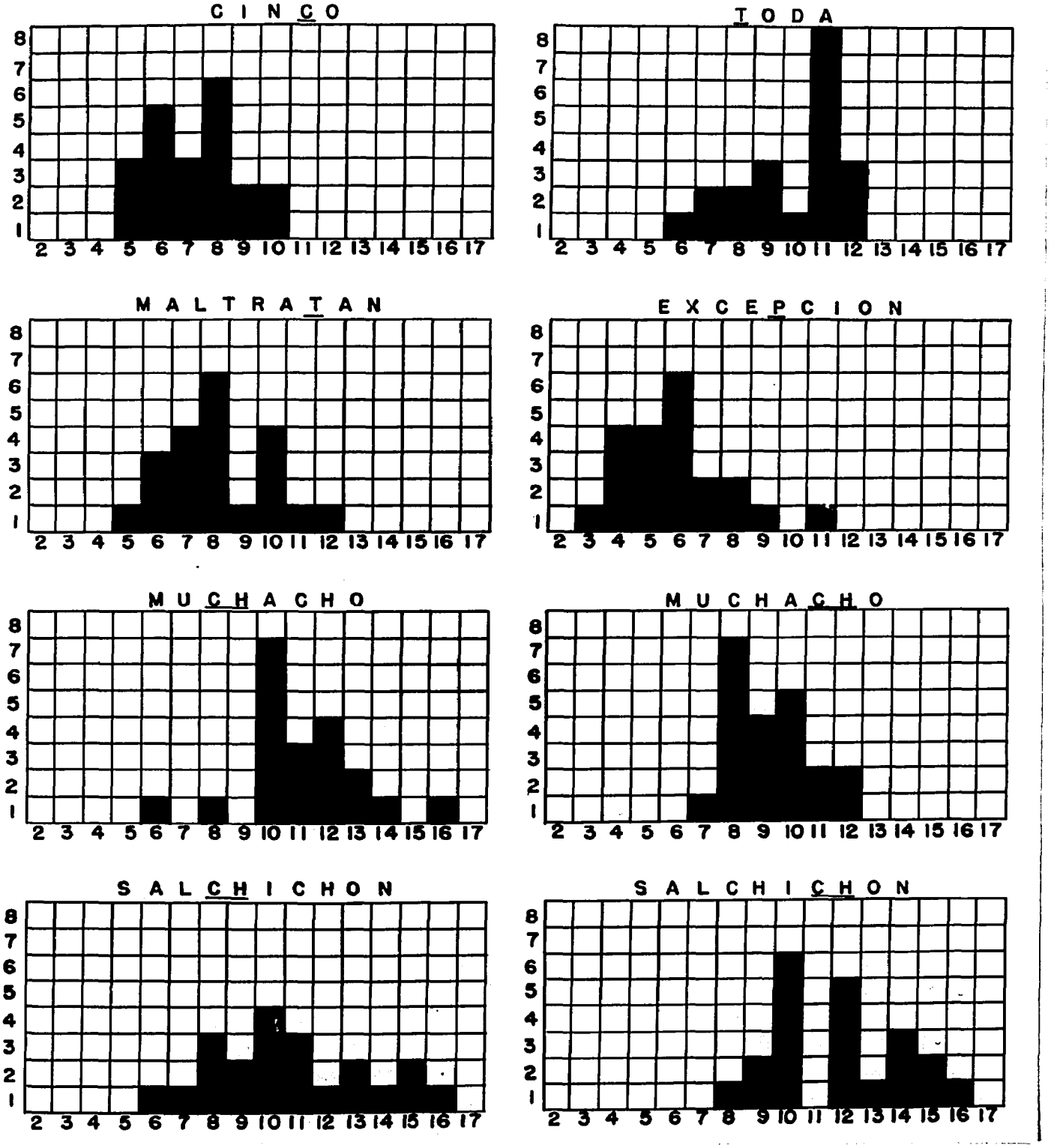


Figure 14. The four histograms in the upper half of this figure illustrate the number of informants (on the vertical scale) who pronounced stop segments with lengths varying from three to twelve centiseconds (on the horizontal scale). The lower half of the figure illustrates the lengths of /č/. A definite shift to the right, or to the longer lengths, is apparent.

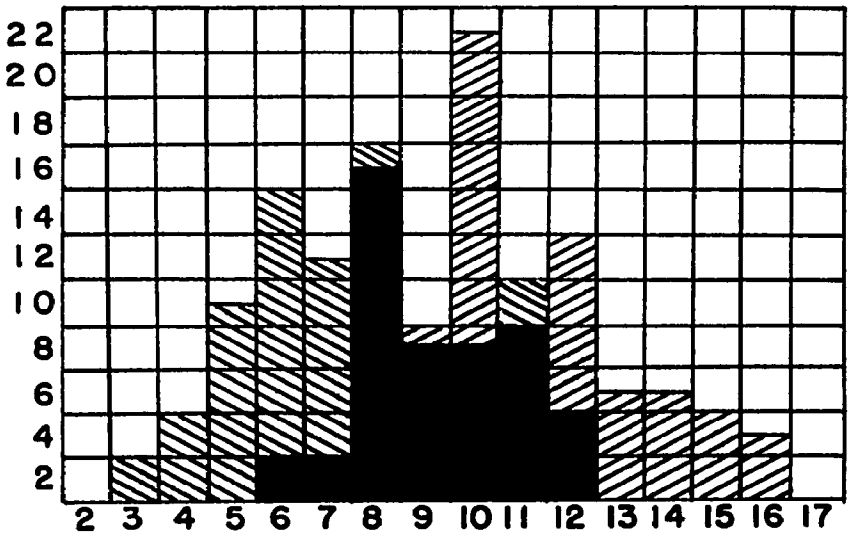


Figure 15. This is a composite histogram showing all the data recorded in the eight histograms of figure 14. Number of informants is shown on the vertical scale and length of the phonetic segment, in centiseconds, on the horizontal scale. Lines slanting downward to the right represent pronunciations of stop segments. Lines slanting to the left represent pronunciations of /č/. The solid black fill represents overlap of the two. Very significant is the exclusive extension of the stops toward the left, representing shorter lengths, and the near-exclusive extension of /č/ toward the right.

other things being equal, the syllable with the greatest intensity (stress) and with the greatest frequency change (intonational variation) will be the syllable that contains the longest mean occurrence of /č/. This is supported by the data since the utterance final syllable -chón in salchichón contains the longest /č/ in the data with a mean length of 11.8 cs. The other stressed syllable with a /č/ is the sentence medial cha in muchacho, with less of an intonation contour and also less syllable length. This /č/ is the second longest in the data with a mean length of 11.1 cs. The difference in this sampling is insufficient to warrant a statement concerning statistically valid significance; it may be only a chance difference.

Both unstressed syllables are shorter than either stressed syllable. The first /č/ in salchichón has a mean length of 10.8 cs. which is not low enough, and the standard deviation from the mean is not low enough, to make this figure demonstrably significant as a smaller figure than that for /č/ in stressed syllables. The second /č/ in muchacho has a mean length of 9.3 cs. which is statistically significant not only in contrast with the first /č/ in muchacho but even in contrast with the other unstressed /č/.

There is 5 per cent probability that the differences between the unstressed occurrences of /č/, as measured according to the rules described above, are chance differences. In addition there is reason to believe that, in this instance, the rules of measurement may have been faulty, for it was found that formants disappear earlier from /l/ in salchichón than from /a/ in muchacho. It follows that part of the reason for obtaining longer measurements for the first /č/ in salchichón than for the second /č/ in muchacho may be due to these difficulties in measurement. This hypothesis can be tested. If difficulties of measurement are causing this particular difference it will appear in the measurements of the initial element of the affricates immediately following /l/ and /a/. The [t] after /l/ will appear longer than the [t] following /a/ relative to the succeeding sibilant portions of the /č/. In comparing measurements, it is found that, in fact, the /č/ following /l/ was originally measured as having a stop element averaging 2.5 cs. longer than its sibilant portion. The /č/ following /a/ was measured as having a stop element .9 cs. shorter than its sibilant portion. This is not a great difference from an absolute point of view but explaining and discounting it as a difference is considerably

more than enough to take the previously mentioned difference in length of unstressed /č/ out of the statistically valid classification of 95 per cent probability of significance.

Except for the second /č/ in muchacho all means of /č/ are longer than the longest mean of all measured occurrences of /p/, /t/ and /k/ as found in seventeen environments in the data. In addition, the shortest mean length of /t/ is 2.5 cs. longer than the longest of /k/. In sum, the average length for /č/ is 10.7 cs.; for /t/, 9.2 cs.; for /p/, 8.2 cs.; for /k/, 7.3 cs.

Of twenty-one individuals from many parts of the Spanish speaking world, not one had a mean length for /p/, /t/, or /k/ that exceeded his own mean length for /č/. Only one informant had a mean length of three /p/ phonemes which equalled his mean length of /č/. An additional measuring of twelve other occurrences[1] of /p/ spoken by this informant showed an average length of 8.5 cs. This compared normally with his 8.6 cs. average for nine occurrences of /t/ and his 9 cs. average for four occurrences of /č/.

Despite Navarro's statements it is clear that acoustically /č/ is longer than the voiceless stops.

It is also true, however, that the stop element of the /č/ is shorter than the stop element, minus aspiration, of the voiceless stops. The average stop portion of those stops which recurred twenty-one times in the speech of all informants is 6.3 cs. That of /č/ is 5.6 cs.

The comparison of /č/ with /t/ is more striking since, of the stops, /t/ has the longest stop portion and the shortest period of aspiration.

---

[1]This data was recorded and run on the spectrograph machine at the same time as the other data. Fifteen /p/ phonemes represent the total data of this type recorded from this informant.

The shortest /t/ has a longer average stop portion than the longest /č/. The longest /t/ which was measured in the data had a mean stop portion of 8.5 cs. compared with a range in mean stop portions of /č/ from 4.6 to 6 cs.

Discussion on the implications of these findings is found in chapter V.

F. Phonetic similarity: "semejante por su timbre"; "cierta semejanza"; "más o menos semejante."

Similarity of segments of behavior - a basic tenet of phonemic theory - is something which in the first instance is assumed.[1] Absolute physical data, considered without correlation to human behavior, show only infinite variation. Yet our reactions show that we consider certain gross acoustic features to recur in phonemes, the individual manifestations of which seem to cluster around nuclei of features.

It is of course possible to obtain extremely accurate and detailed information on variations of elapsed time, pitch, resonance or damping, and intensity by comparing spectrographic recordings of sounds. It is even more significant from a human reaction point of view that, within this diversity, the spectrogram presents data in such a way that similarity of pattern can usually be seen for phonemic entities.[2] Figures 16 and 17 exemplify one facet of this in that the frequencies of vocal cord vibrations vary in a ratio of three to one (figure 16 represents a falsetto voice and figure 17 represents a low tone) while the resonation pattern remains constant.

---

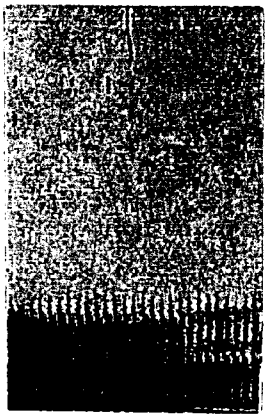
[1]Leonard Bloomfield, "A Set of Postulates for the Science of Language," Language, II (1926), 154, 155.

[2]This is felt so clearly that a spectrographic phonemic alphabet is presented in R. K. Potter, G. A. Kopp and H. C. Green, Op. cit., 65 ff.



[ú]

Figure 16. Informant XXVI produced this spectrogram by pronouncing a sustained [u] in a falsetto voice. The fundamental frequency of the vocal cord vibration is about 360~ as shown by the vertical striations.



[ù]

Figure 17. Informant XXVI produced this spectrogram by pronouncing a low pitched [u]. The fundamental frequency of the vocal cord vibration is about 120~.

Hence it is possible to delimit the spectrographically recorded equivalents of the assumed (or demonstrated) phonemic entities with emphasis on the physical sameness within these entities as opposed to the differences between entities. Although many voice features may vary widely, the resonation features in speech sound signals seem to remain relatively constant.

It is also possible, and not necessarily less profitable from the point of view of language study, to emphasize the differences that do occur between various spectrographic representations of a phoneme.

Detailed study of this "variations-on-a-theme" portion of language structure may prompt some very interesting observations. It appears from the present study, for example, that a parallel can profitably be drawn between analysis of the structure of the sea anemone and analysis of language structure. A slug may consume the internal structure of the anemone, but the anemone can rebuild the organs consumed by the slug. Analogously, a desire for speed or the principle of least effort or something else consumes some of the phonemes, though a potential for their occurrence in that spot is not consumed. On a different level of analysis, something seems to consume or alter certain features of a phoneme. Voicing may be added where an expected and usually present characteristic is a contrastive voicelessness (Figure 18). Expected

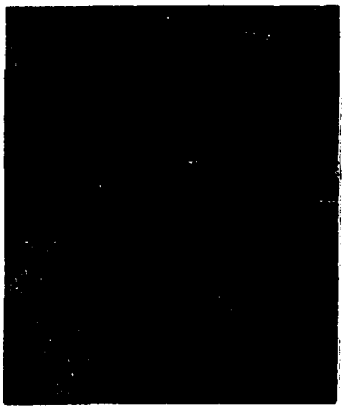


Figure 18. Informant XIV, sentence 10. The velar stop, rendered as a voiceless stop by most informants, is here characterized by continuous vocal cord vibration. Even the aspiration that is regularly shown in voiceless stops is absent. The resulting representation is what would be expected for the velar stop in a word like mango.

vocal cord vibration may be absent from an occurrence of /d/ (Figure 19) or the otherwise regularly recurrent spectrographic gap may fail to appear in some actualizations of Spanish /č/ (Figure 20).

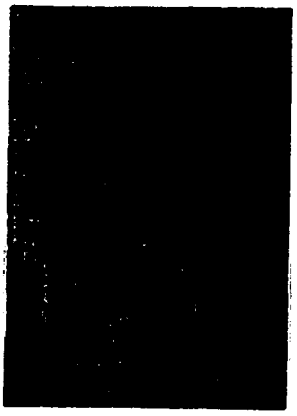


Figure 19. Informant VIII, sentence 6. The dental stop, regularly voiced in this environment, is here shown to be lacking strong vocal cord vibration. There is also clear evidence of aspiration not otherwise found in voiced stops.

The very serious question which arises is this: "Can it be said, when these altered forms occur, that there was an occurrence of voiceless /k/, voiced /d/ or affricate /č/?" Two further theoretical ques-



Figure 20. Informant XXI, sentence 1. The second /č/ is clearly not an acoustic affricate since there is no stop portion.

tions are implied here. First, "How similar must sounds be in order to be considered recurrences of a behavior cycle[1] or a phoneme?" and second, "How different must a sound be from all actualizations of all other phonemes in order not to be excluded from one phoneme by virtue of being a member of another?"

Further discussion of these and related questions is found in Chapter V which deals with a summary of implications and problems.

G. Simplicity: "oclusiva simple."

At one time scientists, quite understandably, praised simplicity of description. It was felt that all science was an attempt to generalize. Generalization was equated with classification into groups whose members the scientists responded to as "alike." Classification was assumed to be a desirable goal probably because scientists have a desire to understand. And, perhaps because the brain is incapable of comprehending very much of the whole universe in terms of infinite variety, they desire the sensation of much greater understanding obtainable by fitting large quantities of phenomena into a few simply described groups.

Now, however, it becomes apparent that simplicity itself is relative to a system. Pre-Hellenic people probably considered extremely complex certain computations which to a Greek mathematician were easily solvable with a simple algebraic formula. A few decades ago mathematicians probably considered extremely complex certain computations which a modern electronic calculator solves in a tiny fraction of the time it might have taken the Greek mathematicians. It may have been subconscious considerations of this type that led some scientists to reject

---

[1] Term borrowed from K. L. Pike, Op. cit., 57.

the fashion of attaching optimum value to the simplest descriptions.

A parallel situation is apparent in descriptions of speech activity. There is infinite variation but a community reacts to certain ranges as constituting recurrences of the same units. These are the phonemes. A simple description, in this situation, is one which lists the phonemic units. A single phonemic unit is a range of behavior to which a native speaker responds "alike" or "alike unless you want to quibble." It is a simplified description in that it groups the total data in such a fashion as to make it more easily comprehended. But, significantly, this is done on the basis of the function of speech behavior or the human reaction to speech behavior and not on the basis of absolute physical similarities of the sound waves.

The idea that simplicity itself is relative to a system serves as the clue to the next development. Theoretically, that which is simple in terms of a native speaker's reaction may be complex in terms of sound waves and vice versa. Practically, some segments of speech behave like single words but are reacted to as compound words because the hearer has had previous experience with smaller segments which go to make up the compound words. "Housetop" serves as an example. Other words, not considered to be compound, may be acoustically or morphologically more complex as, for example, "unreliability".

It seems, for a comparison of the spectrographic evidence and Navarro's words "oclusiva simple," that a situation of this type has been encountered. The phoneme /č/ is phoneto-phonemically complex relative to the voiceless stops while the total make-up of the voiceless stops is acoustically complex relative to /č/.

An attempt at setting up arbitrary boundary lines between the spectrographic representation of vowels and voiceless stops presents the same kinds of problems as does dividing preceding vowels from the voiceless affricate /č/. Acoustic complexity of /č/ and the voiceless stops is equal in this area. But the transition from [t] to [š] in the phoneme /č/ is relatively abrupt and easy to "locate" arbitrarily in such a fashion that all [tš] sounds can be analogously divided relative to analogous transition points between stop and sibilant as in figure 21. Finding an equivalent transition point from the stop

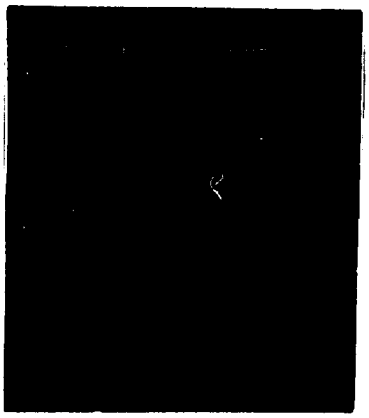


/s a / č i č ó n/

Figure 21. Informant IV, sentence 3. This spectrogram illustrates the average difficulty that one would have in determining analogous transition points between stop and sibilant portions of /č/. There could be only the slightest variation in measurement if it is assumed that the sibilant begins and the stop ends at the left edge of the spike in the first /č/ and at the initiation of the high frequency noise, in the absence of a spike, as in the second /č/.

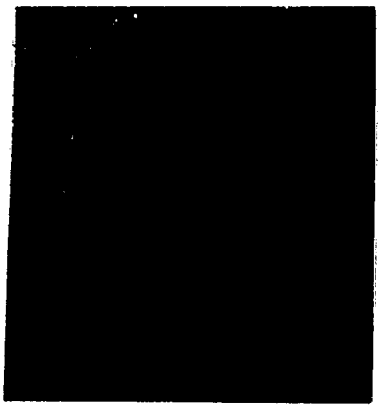
to the aspiration of voiceless stops may present more difficulty, as illustrated in figures 22 through 26. The representation immediately following the gap for voiceless stop may be clearly voiceless friction or aspiration. This aspiration may be initiated gradually or it may occur with a sharp initial spike which in turn may or may not be follow-

ed by one or more additional spikes irregularly spaced. The aspiration may have clear formants indicating the resonance areas characteristic of the following vowel and it may also have characteristics of voicing. Finally, the gap may be followed by certain consonants without any intervening friction or noise at all. Figures 22 through 26 serve as token illustrations of the range in "simple stop" types.



/t u v í m o s k e/

Figure 22. Informant VIII, sentence 15. The /k/ illustrated here, like the /p/ in figure 18, may be divided into stop and aspirate protions as readily as /č/ can be divided into stop and sibilant.



/t u v í m o s k e a l i á ř n o s/

Figure 23. Informant XIV, sentence 15. This example illustrates the difficulty of setting up a general rule for dividing stop and aspirate portions of voiceless stops in such a way as to promote general agreement in measuring. The linguistic context is the same as for figure 22.



Figure 24. Informant XI, sentence 1. Figure 18 shows a /p/ with two clear spikes in the aspirated portion. This illustration shows three spikes more widely spaced and thereby portends the difficulties of measurement that arise in still more exaggerated cases of spike separation as seen in figure 25.



Figure 25. Informant XIII, sentence 2. The problem of separating and measuring stop and aspirate portions is evident here. The complexity of voiceless aspiration relative to that of the sibilant in /č/ is also apparent in this figure and in the immediately preceding figure.

From the point of view of spectrographic analysis the "simple stop," as a cover term for all voiceless stops in Spanish, is complex also in the sense that there are systematic variations between /p/, /t/ and /k/



/ɪnyɛksyón/

Figure 26. Informant VI, sentence 12. The vertical blank section, the rapid fading of indications of voicing at the base line and the intersection of preceding formants 2 and 3 attest to the presence of a voiceless velar stop in this spectrogram. The indication of high frequency noise following the /k/ is the evidence of a phoneme /s/ occurring in this area. There is no way of justifiably assigning a portion of the /s/ to the aspiration portion of /k/. Therefore /k/ cannot be said to have aspiration here as it does elsewhere whenever a vowel follows an actualization of /k/ in which vocal cord activity is absent or sharply reduced as it is here.

in the duration of the closure and the duration of the aspiration. Spanish [t] is significantly longer than [p] and [p], in turn, is significantly longer than [k]. In spite of this, it was found that the average[1] length of aspiration is shorter in the [t] (1.9 cs.) than in

---

[1]Measurements of aspiration were averaged only for sounds which were found to recur twenty-one times in the data. The averages were taken from measurements of aspiration in the [k] of cinco and in the [t] sound in Measurements and toda. The aspiration in the first [t] of of maltratan was discounted because it could not be separated from the following [r] so as to give measurements that might be considered comparable to the others. The averages of aspiration occurring in the other two [t] sounds were equal at 1.9 cs. each.

the [k] (2.8 cs.).[1] Still further, the twenty-one occurrences of [p] in excepción showed a complete lack of measurable aspiration comparable to aspiration elsewhere. The gap was immediately followed by [s] noise.

It is only this [p] in excepción that regularly looks "simpler" than the occurrences of /č/ since in excepción the gap is immediately followed by friction that must be assigned to the following phoneme and not to the stop. The friction or aspiration that appears after the voiceless stops in the other environments in this data, however, has much less in common with, and can therefore be easily separated from the following voiced vowels. This friction is assigned to the preceding voiceless stop phoneme.

Sibilant friction constitutes 52 per cent, or an average of 5.6 cs., of the total average length of /č/ whereas aspiration friction constitutes 26 per cent, or an average of 2.2 cs., of the total average length of the voiceless stops in this part of the data. The data, however, contain more samples of /t/ than of /p/ and /k/. And since /t/ is the longest as a whole yet has the shortest aspiration element, it can be said that an equalized sampling would almost certainly show more than 26 per cent of aspiration in the total average length of voiceless stops.

Contrary to the relatively uniform and unobstructed friction which occurs in all the /č/ samples and which follows [p] in excepción, the friction or aspiration at the end of the voiceless stops is, though

---

[1]The aspiration was included in the length measurements of the stops. Thus the average length of [k] in cinco was 7.2 cs., but since the aspiration was 2.8 cs., the average gap for [k] was only 4.4 cs. The [k] sounds in pecadores and in que were even shorter than the [k] in cinco. This compares with an average gap of 5.1 cs. in eighty-four occurrences of /č/ and an average gap of 8 cs. in twenty-one occurrences of /t/ in toda.

only about one half as long, usually accompanied by irregular striations which in a few cases could indicate irregular vocal cord activity and elsewhere probably indicates multiple pops due to interaction of relatively unrestricted release of air pressure with tissues and saliva. This hypothesis is supported by the fact that a voiceless stop + [s] does not show these striations. The air pressure is released more slowly through the obstruction of the sibilant friction. The release of the stop is less precipitous or violent and any succeeding onset of vocal cord vibrations is also made under less rapidly shifting air pressure conditions. The /č/ similarly has pressure which is released during the voiceless sibilant segment and then voicing can begin under conditions which are apparently more favorable to an unperturbed initiation of vocal cord vibration. The voiceless stops occurring before vowels never have these conditions favorable to a smooth onset of vocal cord vibration.

The range of total length variation seems to be about the same for /č/ and for the voiceless stops. In this sampling /č/ varies from 6 to 16 cs.; /t/ from 5 to 16 cs.; /p/ from 3 to 15 cs.; and /k/ from 4 to 12 cs.

In reference to regular vocal cord activity throughout the full length of the occurrence of the phoneme, however, the stops once more show complexity in the sense that "voicing" is found much more frequently. Of eighty-four spots for occurrences of /č/, none shows evidence of strong voicing throughout its full length and only 4 per cent show evidence of weak vocal cord activity throughout. Of 259 spots for occurrences of voiceless stops, 8 per cent show evidence of strong voicing throughout and an additional 5 per cent show evidence of weak vocal cord activity throughout.

It is interesting to note that the human computer, in this case the person who understands Spanish, handles very simply these data which are tremendously complex for the spectrograph machine, or, more accurately, complex for the person who observes and tries to interpret the spectrographic interpretation.

The fact that the human computer may react to /č/ as complex and to the stops as simple proclaims clearly once more that simplicity is relative within a system. In this case the system does not require any reaction to behavior following a velar stop and preceding a following vowel even when that intervening behavior may be more prominent than the stop portion itself in terms of time consumption (figures 27 and 28).

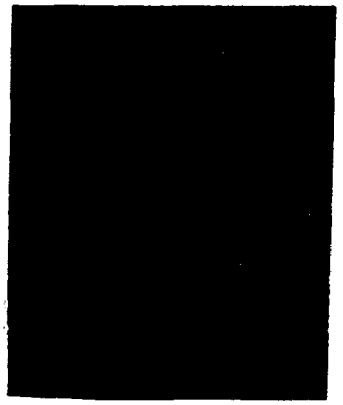


Figure 27. Informant I, sentence 1. That the aspirate portion of /k/ may overshadow the voiceless stop segment is clear from this spectrographic representation in which there is scarcely over 1 cs. of voiceless velar stop but 3 cs. of velar aspirate.

Since no reaction is demanded in the system, it goes by unnoticed as to its complexity. On the other hand the system does demand separate reaction to behavior consisting of dental stop, [t], plus a specified type of release, [š], which, if absent, would cause this behavior unit to be equivalent to some varieties of /t/. This is in specific contrast to

/p/, /t/ and /k/ which, except in limited environments, regularly occur with aspirated release yet have a potential for occurring without that release (as in the limited environment illustrated in figure 26) without danger of becoming equivalent to forms of another behavior unit or phoneme.

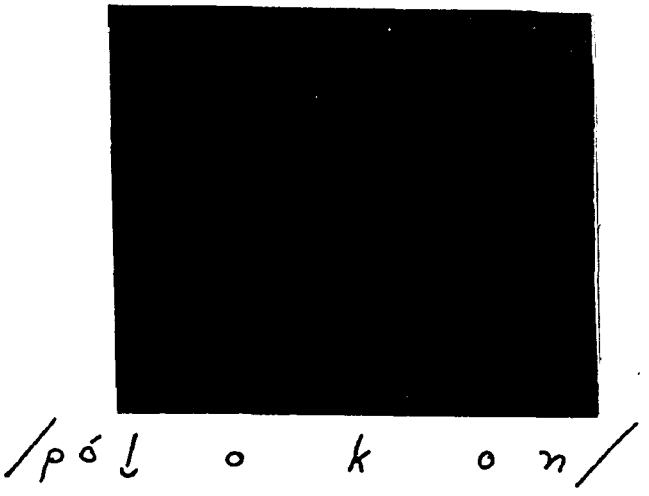


Figure 28. Informant XVII, sentence 18. It is demonstrated here that the aspirate portion of /k/ may remain relatively long in the longer or slower pronunciations as well as in the very rapid pronunciations as illustrated in figure 27.

Speakers of a language probably are not generally aware of greater complexity in one phoneme than in another; but those who specialize in the study of the sounds of their language quite easily observe, for example, that /t/ and /s/ are significant behavior units in Spanish and that /č/ contains elements which include major characteristics of both of the former. This parallels the linguist's separate consideration of the morphemes "sloth" and "ful" when he is dealing with the word "slothful."

/č/ may seem complex to some observers because of a native phonemic bias. The voiceless stops, which are usually acoustically more complex and less complex only in limited environments, are considered "simple"

for two reasons that seem apparent. First, systemically, the "simple" stops potentially occur with or without aspiration and when the aspiration does occur it usually occurs without characteristics as closely resembling those of another total phonemic unit as is true with the sibilant portion of /č/. Second, physically, the average length of aspiration in those environments where aspiration regularly occurs is roughly half that of the average length of the sibilant release in /č/.

H. Variation: "dialectal"; "variantes."

Implicit in the tone of Navarro's writing is the meaning of "non-Castillian" when he writes of "dialectal" pronunciation. In the light of attitudes and approaches generally accepted by linguists this becomes an unnecessary and unwanted restriction of meaning for the word "dialectal," at least for the purposes of this thesis. It is useful, in this instance, to maintain a separation, wherever possible, between phenomena of air vibration and evaluative phenomena.

The selection of informants for this study was such that very few reliable statements on dialect can be given. However, since the informants were selected from the various major geographical areas of the Spanish speaking world it may be assumed that various dialect forms do occur in the data.

Some dialect variants can be readily observed in the spectrograms, as, for example, the occurrence of [l], [y] and [ž] in the word "pollo." It seems quite probable, though, that for /č/, even after further investigation, personal variations of length, articulation, sharpness, etc. will retain a prominence on the spectrograms that will overshadow the differences one might hope to find in dialect types.

This is due to the fact that the perception of dialect variations of /č/ consists largely of the perception of the following factors listed by Navarro: the (time) extension of the contact between the tongue and palate; the point of articulation; the position of the tip and blade of the tongue; and the duration of the fricative element.[1] It was pointed out in the comments in section A of this chapter that, except for duration, these factors, as here worded, are less subject to accurate measurement on the spectrogram than by other means. Therefore, though dialect variations of the type mentioned by Navarro may occur, they will, in these data, be obscured on the spectrograms by the spectrographically more prominent personal variations.

Still less subject to spectrographic determination are evaluative measurements which categorize behavior as "pronunciación española correcta" as opposed to "pronunciación dialectal."[2] And it is at present manifestly impossible to determine which pronunciation on a spectrogram is "castellana sin vulgarismo y culta sin afectación."[3] These are generalizations which are not stated or defined in such a way as to be readily and productively experimented upon and assayed with the spectrograph, and certainly not with the limited data used in this study.

The relative length and the presence or absence of the fricative and the stop elements would seem to be the primary sources of dialect discrimination on the basis of spectrographic analysis. A further source might be the relative intensity of the fricative element, i.e., fortis and lenis distinctions which are discussed more fully in section N below.

---

[1]Navarro Tomás, Op. cit., 126.

[2]Loc. cit.

[3]Navarro Tomás, Op. cit., 8.

I. Timbre: "timbre."

As a musical term timbre refers to quality or tone identity of a sound as determined by the number and character of its overtones. As used in phonology, the term has a much less specific usage. Phonetic descriptions speak of [h] timbre, vowel timbre, whistle timbre, percussive, implosive, glottalized, click and belch timbres.[1] Timbre, as used in these examples refers to the characteristic sound of one setting of a sound producing mechanism.

Since the fricative part of /č/ has no overtones, nor even a fundamental frequency in the conventional sense, it is safe to say that the word "timbre" is used by Navarro in the general phonetic sense and "por su timbre" could be omitted without loss of communication value since the information is repeated in context with the words "al sonido de."

The sound spectrograph analyzes speech sounds into physical components or acoustic characteristics, including those which are assumed to cause perception of timbre, more satisfactorily than any other single device now in common use.

In harmony with the broad interpretation of timbre, the following comparisons with [tš] or [š] in English, French, Italian and German may be considered a continuation of this discussion.

J. English [š] and [tš]: "sonido de la sh inglesa"; "la ch en ingl. church."

The pertinent material in English was taken from one speaker's pronunciation of "church," "share," "sh," and the nonsense syllable [aš].[2]

---

[1]Kenneth L. Pike, Phonetics, (Ann Arbor: University of Michigan Press, 1943), 91, 93, 100, 103, 131, 142, 147.

[2]Spoken by informant XXVI.

Of the five examples of [tš] and [š] found here, the lowest frequency of fill begins at 800~ and continues strong through the highest frequency range in the machine (approximately 3500~). At one point in one occurrence only frequencies above 2300~ occurred. Since this variation falls within the range outlined by the Spanish samples it may be said that these acoustic phenomena in English are similar in this respect to the comparable phenomena in Spanish.

Resonance bars within the fricative elements here are neither as clear as in the French sample nor as clear as in 75 per cent of the occurrences of Spanish /č/. But the intensity of the fricative element in English is greater than in 80 per cent of the Spanish data. The length of both occurrences of [tš] is greater in the English sample than in 90 per cent of the Spanish data concerning /č/.

There are no data which flatly contradict an acoustic statement, parallel to Navarro's impressionistic statement, that Spanish /č/ is phonetically similar to related sounds in English, French, German and Italian. Conversely, there is evidence which supports it.

K. French [š]: "sonido francés."

It is apparent that the spectrographic representations of some occurrences of the fricative element of Spanish /č/ include features which bear a strong resemblance to comparable features found on the spectrograms of French [š] in the word "chambre."

The effects of the labialization of which Navarro speaks in reference to the French sound may be seen in the strength of the resonance bars of figure 29. The reason for connecting labialization and strong resonance bars in fricatives is made clear in figures 30 through 33. Figure 30 shows the markings that result from the pronunciation of a

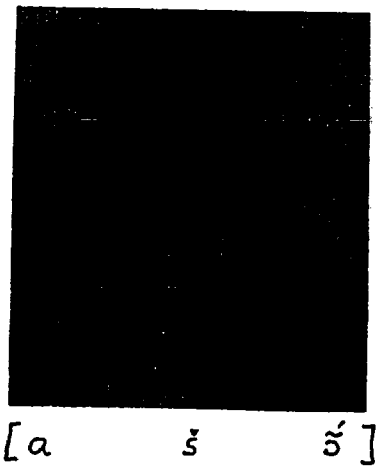


Figure 29. Informant XXIII pronuncing "la chambre." Acoustic energy is high in the [š] but also concentrated in visible formants.

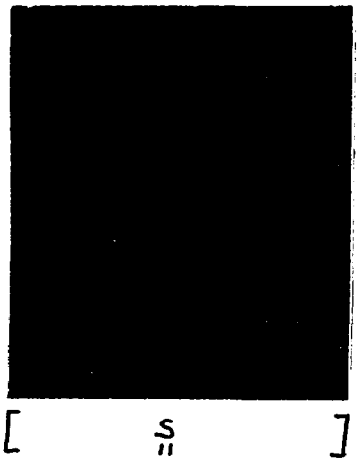
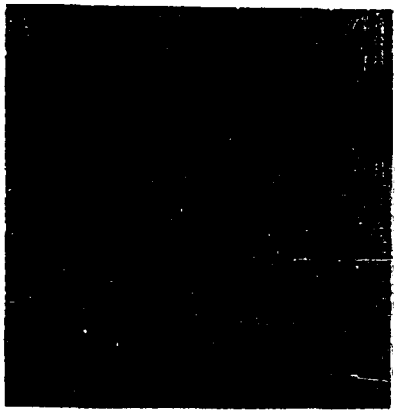


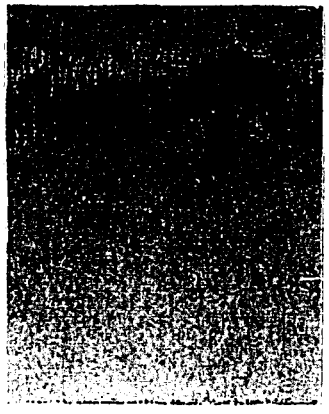
Figure 30. Informant XXVI pronuncing a fortis [s]. High frequency noise is general and not concentrated in formants.

fortis [s] with lips unrounded. Figure 31 represents the same pronunciation but with hands cupped cylindrically in front of the mouth so as to form a considerable chamber beyond the point of sound production. In figure 32 the process is repeated except that the previously open-end cylinder is now closed at the end away from the lips and near the microphone. We should expect that [s] would have less sharply defined



[S " with cylinder]

Figure 31. Informant XXVI pronouncing a continuation of the same fortis [s] seen in figure 30 but with hands forming a cylinder in front of the mouth.



[S " with cup]

Figure 32. Informant XXVI, a continuation of the conditions found in figure 31 except that the manual cylinder is closed to form cupped hands in front of the mouth. It is seen that, by diminishing the front opening of a chamber, an additional quantity of acoustic energy may be filtered from that which was recorded in the two preceding figures. Analogously, lip rounding might be expected to show results of this general type.

formants or resonance bars than [š] which is both farther back and labialized. Similarly [s] would be less often expected to include formants than would an equivalent retroflex [s]. Figures 30 and 33 attest to the validity of those suppositions.



Figure 33. Informant XXVI pronouncing a retroflex or back [s]. Formants similar to those found in figure 29 are less clearly discernible here but can be distinguished nevertheless. Conversely, figure 30 does not present visible formants.

The effects of labialization or a lengthened damping chamber (which Navarro properly implies is not a necessary part of the Spanish sound) can be seen in many of the spectrograms which recorded the acoustic phenomena in the pronunciation of "muchacho." This probably comes about as a result of contiguous phonemes /u/ and /o/ which have a usual feature of lip rounding. Of all samples of /č/ between /u/ and /a/, 86 per cent show, through the appearance of formants, evidence of a relatively long resonance chamber; no samples are clearly without that evidence and the remaining 14 per cent are open to question. 71 per cent of the occurrences of /č/ in the environment of /a-o/ show evidence of formants while 29 per cent do not. In the environment of /l-i/, 48 per cent show evidence of formants. And in the environment of /i-o/, 76 per cent show evidence of formants. It must, of course, be recalled that the spectrograms used in this study do not show any energy that may occur above 3500~.

It is not at all clear from these data that the spectrograph can conclusively bear out Navarro's implication that characteristic labialization of French [š] contrasts with a lack of labialization in Spanish [tš]. The most promising attack seems to be in the area of damping phenomena as registered in spectrographic formants, but since the French data are far too limited, conclusions must be postponed.

L. Italian [tš]: "la c en ital. cento."

The Italian [tš] (figure 34) presents nothing startlingly different from what appears in the Spanish data. The length of 14 cs. more or

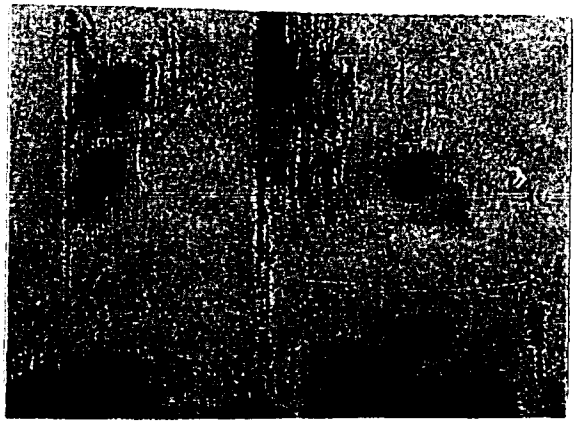


Figure 34. Informant XXIV pronouncing "de cento." The gross acoustic characteristics of this Italian [tš] fall within the range of those found in the data for Spanish [tš].

less equally divided into stop and sibilant portions, the frequency characteristics of the sibilant noise, the intensity, the formant characteristics, the gentle but quick onset of friction, and the transitions to and from contiguous vowels all bespeak an acoustic similarity to Spanish /č/.

M. German [tš]: "la tsch en alemán deutsch."

The [tš] of one German speaker's pronunciation of "Deutschland" is reproduced in figure 35.



Figure 35. Informant XXV pronouncing "Deutschland." The non-transitional parts of the sibilant portion of German [tš] give measurements not much different from Spanish averages except possibly for length relative to the stop portion.

Using the same criteria that were used in measuring the Spanish sounds it is found that this [tš] measures 17 cs. Of this total the gap representing the capped air stream is 6 cs. and the clearly fricative portion is 8 cs. The remaining 3 cs. consists primarily of a light frictional element indicating less acoustic energy in this transition period which may indicate the presence of a short period in which a voiceless [l] is occurring.

The fricative noise begins abruptly, occurs most prominently above 1500~, and shows signs of having its energy concentrated in formants. These characteristics also occur in the Spanish data.

This spectrographic representation of German [tš] is, like Navarro's impression, more or less similar to the Spanish /č/.

Navarro says also that the fricative element of the Spanish sound is shorter and sharper. Though, again, the German data are wholly insufficient for valid statements concerning phonetic norms in the lan-

guage, statements can be made illustrating types of comparison which are difficult or impossible without spectrographic equipment.

As already stated, spectrographic analysis can give rather definitive statements on length. In this particular case the average length of eighty-four potential occurrences of the fricative element of Spanish /č/, using the border criteria described earlier, was 5.6 cs. compared with 11 cs. for the German fricative element in this sample.

A comparison of relative lengths of stop and fricative components is found to be in agreement with Navarro's statement. The average of eighty-four stop components in Spanish is 5.1 cs. - only a little less than the average length of the fricative element - while the 7 cs. stop portion of the German example is not much over half the length of the 11 cs. fricative component.

As to sharpness, it does not seem that the spectrograph will bear out a great deal concerning the validity of Navarro's impression. Figure 36 shows two examples of Spanish /č/ with energy scattered in approxi-

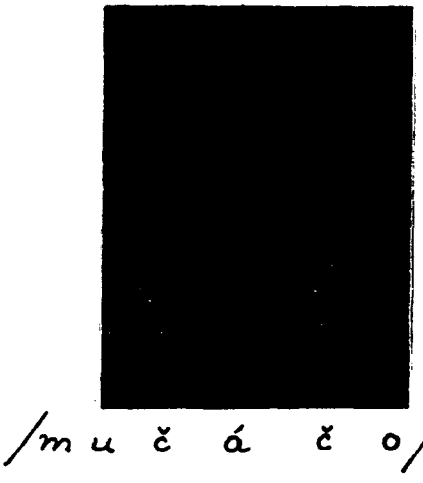


Figure 36. Informant X, sentence 1. It is clear from this picture that energy distribution on the frequency scale could not be the cause for a sensation of greater sharpness from the Spanish sound. The energy distribution for the sibilant in the first /č/ is approximately the same as that for the midpoint of the German sibilant segment.

mately the same frequency range as for German [tš]. Figure 37 shows that Spanish may exaggerate even more than the German example the strength of a low fricative noise formant caused by resonance cavity shapes conditioned by following sounds.

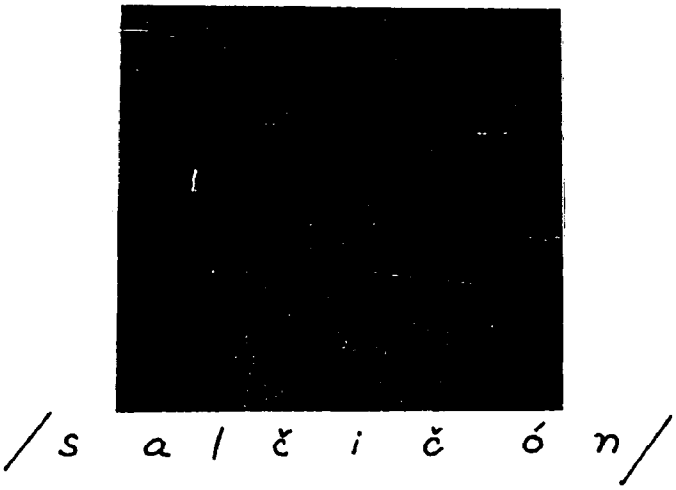


Figure 37. Informant XVII, sentence 3. If sibilant noise in the high frequencies is equated with sharpness this illustration shows that Spanish can and very regularly does show sibilant noise in lower frequencies than does the German sample. The second /č/ shows a low noise formant which corresponds to the somewhat higher formant, also conditioned by its following environment, in the German sample.

The next and final section of this chapter deals further with the efficacy of the spectrograph in terms of gathering material that may be interpreted as sharpness of a fricative or sibilant element.

N. Sharpness: "más aguda."

Perceptual sharpness of sounds is not as clearly defined as, for example, the feature of pitch. If it is said that one sound has a higher pitch than another it can be tested and the conclusions of a proper experiment will be accepted as valid. When it is said, however, that one sound is sharper than another, it is difficult to find a satisfactory, commonly accepted basis for experiment. The problem itself is not pre-

cisely grasped. It seems rather to be an artistic matter in which reactions are as yet relatively unpredictable.

Assuming that Navarro's observations are correct, it has been impossible so far to find one particular characteristic of the spectrograph markings which can be definitely equated with the sensation of sharpness related to perception of sibilant sounds. However some tentative statements can be made concerning various characteristics which may bear a relationship to what Navarro calls "aguda."

In the preparation of the preceding four sections of this chapter it was not observed that there was any regularly recurrent difference in the frequency level of any of the noise found in the stop or sibilant portions of [tš] as found in the data of four languages compared with Spanish. If Navarro means "high-pitched," the data do not seem to bear him out; however, neither are they extensive enough to refute him definitively.

It is possible that perceptual "sharpness" might have some relationship to what Navarro called "la mayor o menor fuerza con que se produce el sonido."[1] If this were true then we could expect degrees of "sharpness" in the sibilant portion of /č/ to be related in some way to fortis and lenis. Fortis shows on the spectrograms as greater relative blackness of the marking; lenis as relatively lighter blackness. Again, the data give no conclusive evidence to show Spanish [tš] as either more or less fortis than [tš] or [š] in the other four languages.

Abrupt versus gradual onsets and offglides of friction might serve as contributing factors to perceived gradations of sharpness. Perception of sharpness might also be partially influenced by the relative length

---

[1] Navarro Tomás, Op. cit., 126.

or shortness of a sound. But neither rate of onset nor relative length of the sibilant element contrasts significantly between Spanish and the other languages for which affricates were tested.

All of the factors mentioned above occur in the data in varying combinations. And, in comparison with the English, Italian and German data, there seem to be more spectrographic indications of the similarity that Navarro mentioned earlier than of differences that might be connected with sensations of increased sharpness in the Spanish sound.

Before making any meaningful new conclusions, however, it would be necessary to get impressionistic reactions concerning sharpness from more speakers of Spanish and from speakers of the other languages involved and perhaps from speakers of some "neutral" languages. In addition it would be necessary to get more spectrographic data.

In summary, even the limited spectrographic analyses used as a base for this comparative study imply some things of importance in a review of the topics listed at the beginning of the chapter.

In reference to the first section - points of articulation - no basis for direct refutation of Navarro was found. The spectrograms used in this study show characteristics of the resonating cavities but, without further experimentation, precise tongue positions of the informants studied cannot be stated at this time.

In reference to the "affricate nature of /č/," the spectrograms showed almost four per cent occurrence of /č/ without a stop element. Navarro says that contact of the tongue with the prepalate and the alveoles may be narrow, but that it is always enough to stop the air stream.[1]

---

[1] Ibid., 125.

Concerning the next section, Navarro speaks only of the "voiceless nature of /č/." This section is devoted largely to illustrating the fact that there is a wide range of possibilities in regard to the possible appearance of acoustic signals that are indicative of vocal cord vibration. Assimilation of strong voicing characteristics from preceding vowels may continue part way or completely through the duration of the stop element of the affricate. In addition, modified vocal cord vibration continues throughout some occurrences of the affricate. These are perceived as voiceless. The reason for this is obviously not the lack of vocal cord vibration. Such physiological vibration seems clearly indicated by the acoustic vibration. The reason might possibly lie in the failure of the ear to react to a below threshold stimulus; or it might well be further along the chain of speech reactions as discussed further in the summary statement concerning "length of /č/."

In discussing "gradual vs. abrupt movement" Navarro is explicit concerning the abruptness of the tongue movement between the stop and sibilant elements of the affricate. He is undoubtedly correct if his statement is limited to the proper frame of reference. The spectrographic data indicate that the proper frame of reference is not acoustic, and probably not physiological, since in these areas wide variation is tolerated.

Concerning the "length of /č/," it was found that Navarro's statements are not supported by the acoustic evidence. Moreover, it seems probable that, in this area of the study also, there is a fairly close correlation between acoustic activity and physiological activity. The statistics indicate strongly that the simple stops are not equal in length with the affricate, acoustically, in Spanish. It seems indicated

that they are not equal from a physiological descriptive point of view and it seems warranted to say that probably Navarro's informants did not supply him with physiological stops equal in length with the affricates. The explanation seems to be that the difference in length is below Navarro's threshold of perception. Therefore, though Navarro's statements concerning length are clearly invalid as an acoustic or physiological description, they might possibly be valid as a description of a type of brain reaction, or phonemic reaction, as found in a Spanish speaking person.

The sections on "phonetic similarity" and "simplicity" corroborated, from a slightly different point of view, the most significant implications of the preceding three sections, namely: statements concerning linguistic phenomena must vary as the frame of reference varies.

The contribution of the remaining sections was primarily one of showing that more spectrographic work would be required to shed much light on the topics discussed therein.

CHAPTER IV

THE SIBILANT /s/

There are five major divisions, following pertinent paragraphs of Navarro, in this discussion of Spanish /s/. They occur in the following order:

A. Fronted [s] and seseo (§105), (§108).

B. Pronunciations corresponding to x in the written language (§129).

C. [gs], [ks] and [s] (§128).

D. [z] and variations (§94), (§107).

E. Assimilation and "defects of pronunciations" relative to /s/ (§109), (§106).

A. Fronted [s] and seseo.

> §105...La s final de sílaba, en contacto con una t siguiente, ...toma el punto de articulación de la t, formándose con la punta de la lengua contra la cara interior de los incisivos superiores, y no contra los alvéolos de estos mismos dientes, como ocurre en los demás casos, § 106, esta s dentalizada la representamos con el signo [ṣ]:....
>
> En contacto con una interdental [θ] la s es atraída hacia los bordes de los dientes, un poco más que ante la consonante t, llegando a ser en parte absorbida por la fricación de dicha [θ]; pero la representamos también en este caso con el mismo signo [ṣ]: ...
>
> De un modo semejante, la s sonora, [z], en contacto con una [d] o [đ] siguientes, deja su articulación alveolar, § 107, formándose contra los dientes, como estas mismas consonantes; ...[1]

> §93...Sabido es que en la pronunciación hispanoamericana, en la de Andalucía y Canarias y en la de las clases populares de Valencia, Mallorca, Cataluña y Vasconia, cuando hablan español, es corriente sustituir la [θ] (c, z) por la [s]: ... A esta sustitución se le llama seseo. El seseo andaluz e hispanoamericano se distingue, por el timbre de la s, del valenciano y del catalán. El seseo vasco, por su parte, es también distinto del de las otras regiones.

---

[1] Ibid., 104-5.

> ...Son muchas, sin embargo, las personas de dicho origen que teniendo que viajar o vivir fuera de su país adoptan el uso de la [θ], cuyo sonido, por su carácter culto, borra todo indicio de procedencia.[1]

Navarro says that [s] is fronted before [t], [d] and [θ]. Because of the prevalence of seseo, of course, [θ] occurs only in the speech of people in a limited dialect area and, according to Navarro, in the speech of many others who must travel or live outside of their country.

Study and measurement of the data produced no definite rules for correlation of the acoustic data with Navarro's descriptions concerning dentalization or fronting of /s/. The occurrences of /s/ used as a control group included, before consonants, español, despacio, juzgar, escollos, after [p], excepción, and intervocalic, exceso. Also studied were the occurrences of /s/ in está, hasta, escollos de and luz dorada in which the sibilant occurs before [t] or [d] and is presumably fronted. The three words with z and c spellings were included out of personal interest. Non-seseo speech was, of course, discounted in these words. No recurrent pattern of variation was found in these data.

The fact that these data presented an unstructured picture lends weight to two opinions: first, that matters relating to physiological "point of articulation" as traditionally understood are perhaps not best measured by spectrograph alone; and second, that the existence of any correlation with Navarro's statements on this problem cannot be definitively handled by a study of these data which, though broad in scope, are too limited for this specific problem.

In reference to matters pertaining to seseo and spectrographic identification of points of articulation it is necessary to note that,

---

[1] Ibid., 93-4.

in these data, it was found difficult to predict whether a specified portion of a spectrogram represented [θ] or [s].[1] It follows that the unnamed differences which Navarro finds within seseo practice as produced by various groups (Andalusians, Spanish-Americans, Valencians, Catalans and Basques) would not, on spectrograms, be easily distinguishable from variation that might be personal, accidental or conditioned by linguistic context.

As impressionistically noted at the time of recording, 90per cent of the data collected for this study show seseo influence. In only one or two words, and even there not systematically repeated in the second recording, was any attempt shown by a person of seseo speech to avoid seseo. These limited data therefore would seem to question Navarro's statement that many people who have to travel or live outside their countries adopt the use of [θ]. However, proper perspective requires that this statement be seen in the light of geographical context and time context as well as linguistic context. Navarro was writing a quarter of a century ago. The prestige of linguistic forms has had time to change since then. In addition, Navarro was not writing about great numbers of Latin Americans travelling and studying in the United States. The conclusion therefore is not a rejection of Navarro's statement but rather a reinterpretation. It can be said that the data gathered for this study indicate that at the present time the use of [θ] by Spanish speakers in the United States does not erase all indications of origin of the speaker. In fact the regular use of it gives a very definite

---

[1]The illustrations of [θ] and [s] in Potter, Kopp and Green are abstracted from English speech which was apparently not read at a rate of speed comparable to that of the Spanish data used in the present study. R. K. Potter, G. A. Kopp and H. C. Green, Op. cit., 127, 135.

indication of the origin of the user with, apparently, only a very small probability of error.

Since over 90 per cent of the acoustic data show seseo influence and since Navarro says only that seseo consists of substituting [s] for [θ], two possibilities emerge in reference to Navarro's [s] + [θ] sequences: either they become [s] + [s] in most of these data, i.e. [s] with two morae of length, or the product of seseo fuses with the preceding [s] and results in the omission of an element.

In an attempt to resolve the problem, the words ascender and cerro were studied first.

The mean length of the fricative element between the first two vowels in ascender is 10.3 cs. If the data of the two informants who used [θ] are omitted, leaving nineteen occurrences, the mean is found to be only 9.7 cs. This alone might seem at first to indicate an easy answer. For while individual seseo speakers varied from 7 to 14 cs. in the production of this intervocalic fricative, non-seseo speakers averaged 15.5 cs. for [s] + [θ]. It seems likely that the seseo speakers were pronouncing one phoneme in contrast to two for the speakers who pronounced [θ].

This conclusion was not wholly warranted, however. There were nineteen seseo speakers but only two non-seseo people. In addition there were some data that seemed somewhat suspicious. One of the non-seseo informants recorded [s] + [θ] with a duration of only 13 cs. while recording [θ] alone in cerro also with a duration of 13 cs. The mean duration of the fricative in cerro as recorded by all informants was 11.7 cs. - not much greater than the 10.3 mean for the fricative segment in ascender, but it proved to be statistically significant with less than

1 per cent probability of error. By itself this was easily explained. Pronunciations corresponding to c in cerro could be expected to be relatively long by virtue of the fact that these were initial in a stressed syllable.[1] However it was found at the same time that those records of the unstressed sibilants in tuvimos que and aliarnos a which were accurately measurable showed, respectively, means of only 6.4 cs. and 8.6 cs. - considerably less than the unstressed fricative segment in ascender.

It seems reasonable, on the basis of the data, to hypothesize the occurrence of a long sound in the stressed syllable of cerro, a short one in unstressed tuvimos que and two short ones in unstressed ascender.

Additional measurements were gathered on the remaining suitable examples that could be found in the data. These included the words excepción, exceso and necesito. In the word excepción [s] and unstressed [θ], or its seseo form, were found contiguous and they could be measured relative to the stressed [θ], or its seseo form, in the final syllable. In exceso [s] and stressed [θ] could be measured relative to the unstressed [s] in the final syllable.

For excepción this test showed the mean length to be only 8 cs. for [s] + [θ], or hypothetical [s] + [s], at the border of the first two syllables; [θ] or [s] alone, in the third and stressed syllable, reached 11.2 cs. This paralleled, and exceeded, the difference found between unstressed [s] + [θ], or its seseo form, in ascender (10.3 cs.) and stressed [θ], or its seseo form in cerro (11.7 cs.). The 8 cs. measurement of pronunciations corresponding to xc in excepción is lower than that for unstressed [s] in aliarnos a (8.6 cs.).

---

[1] Statistics in the preceding chapter show that /ð/ in stressed syllables averaged 1.4 cs. more of duration than those in unstressed syllables. That constitutes a 14 per cent length addition for the initial consonants in stressed syllables.

The mean length for the first frictional segment in exceso is 11.8 cs. for all informants. This is just barely higher than the analogous syllable initial stressed sounds in cerro (11.7 cs.) and excepción (11.2 cs.). Excluding the non-seseo speakers, the mean length of the first sibilant segment is only 11.2 cs. This is exactly equal to the pronunciation corresponding to c in the stressed syllable of excepción.

It may also be noted that non-seseo informants, who unquestionably have two signaling units or phonemes in the first frictional portion of exceso, had average durations of 17 cs. for this portion.

The mean length of the unstressed intervocalic sibilant in exceso is 9.5 cs. This is distinctly more than the 8 cs. length of unstressed [s] + [θ], or its seseo form, in excepción.

Thus evidence began to accumulate. It was becoming clear that [s] plus a seseo form did not produce additional length and also that whatever the length of /θ/ in comparison with /s/ (the data are insufficient for a statement), it was indicated that seseo sibilants had the same relative length as non-seseo sibilants.

Perhaps the strongest evidence came near the end of this portion of the study. In the spectrograms of necesito it was found that the first frictional element, syllable initial but unstressed, has a mean duration of 7.7 cs. while the second has a mean duration of 9.3 cs. This follows a pattern that had been obvious.

But another matter was of greater importance as a clue to interpretation. During the measurements it became apparent that the first vowel in necesito was about the same length or, in many cases, longer than the stressed vowel. Careful measurement showed a mean length of 6.6 cs. for the first [e] and only 5.7 cs. for stressed [i].

It was thus shown that a Spanish vowel in a syllable hitherto considered unstressed or only secondarily stressed may nevertheless be longer than another vowel in a stressed syllable. It is not believed that [e] is inherently significantly longer than [i] since the second [e] in necesito had a mean duration of only 3.7 cs.

These figures indicate that the relatively high mean duration of the unstressed frictional element in ascender did not necessarily indicate the presence of two structural entities when compared with the low mean duration of word final unstressed /s/.

Also the proportional difference of length between those stressed and unstressed pairs which occurred nearly contiguously in the data showed that the length of the sibilant in ascender cannot be considered too great to allow the interpretation of a single phoneme here in the case of seseo speakers. The stressed [s] is longer than the unstressed by 1.4 cs. in the case of ascender al cerro; by 1.6 cs., only slightly more, in the case of necesito; by 3.2 cs. in the case of excepción; and by 2.3 cs. in exceso.

One final test for a statistically significant correlation between the individual variations in length of the first and second [s] of necesito on the one hand and those of ascender al cerro on the other hand proved negative.

Therefore there is no reason to believe that the fricative in ascender - [s] plus the seseo form of [θ] in Navarro's description - is any longer than the simple seseo form in necesito. It is strongly indicated that seseo of [θ] does not add a mora of length when it follows a

sibilant. It is rather the same, from a descriptive point of view, as though the [θ] were omitted.

B. Pronunciations corresponding to x in the written language.

> §129...Históricamente, la x de nuestra actual escritura equivale al grupo cs; pero su pronunciación solo se ajusta al valor literal que este grupo representa en casos muy marcados de dicción culta y enfática. En la conversación corriente, la x ante consonante se pronuncia como una simple s: ...Entre vocales se pronuncia como [gs] con una [g] débil y relajada que a veces, como la del grupo cc, resulta también en parte ensordecida: ...La [s] en estos casos tiene siempre en español sonido sordo. Franceses e ingleses, influídos por sus idiomas respectivos, incurren de ordinario en el error de dar a dicha [s] sonido sonoro, pronunciando [ẹgzámen], [ẹgzímjo], etc. Ante una h, la x se pronuncia como si fuera intervocálica: ...El habla vulgar pronuncia la x intervocálica con el mismo valor de [s] que la x final de sílaba: ...La pronunciación correcta admite, generalmente, la s por x intervocálica en exacto [esákto], auxilio - [ạusíljo] y auxiliar - [ạusiljár].[1]

Nine words were included in the data with a view to examining the matter of pronunciations corresponding to x in the written language. The words, from among those listed by Navarro, were chosen in such a way as to provide an opportunity to study pronunciations in various environments.

Navarro says first of all that, historically, the oral pronunciations corresponding to x in the written language were equivalent to those which corresponded to cs in the written language. If s in the group cs is to be rendered by [s] in pronunciation, and if seseo consists of the substitution of [s] for [θ], then we may also deduce from Navarro that historically the pronunciations corresponding to x in the written language are equivalent to the seseo pronunciations corresponding to cc in the written language. Navarro goes on to say that today this correspondence of cs and x holds good only when those words,

---

[1]Navarro Tomás, Op. cit., 140, 141.

which in written form contain an x, are pronounced very emphatically or learnedly.

A compilation of spectrographic data shows marked similarities between the pronunciations corresponding to cc in inyección and the averages of those corresponding to x in excepción, exceso, éxito, exacto, examen, auxiliar and exhaló. For the purpose of this compilation, the records of nineteen informants were used. The two using [θ] in their pronunciations were not counted. The resulting comparison of averages on durations of velar elements, voicing and sibilant elements showed striking similarities. The average of velar elements, whether stop, fricative, voiced or voiceless, was 3.4 cs. for pronunciations corresponding to cc and a very similar 3.7 cs. for the average of velar elements in pronunciations corresponding to x in the words listed above. The average amount of voicing in the initial part of cc was computed to be 2.8 cs. and the amount of initial voicing on the pronunciations corresponding to x also averaged 2.8 cs. for nineteen occurrences of each word. The amount of sibilant corresponding to x averaged 10.5 cs. and that corresponding to cc - slightly more since it occurred only in a stressed syllable - averaged 11.9 cs. The minuteness of the variation indicates that Navarro would classify these data on pronunciations corresponding to x as "casos muy marcados de dicción culta y enfática."

Navarro says in effect that in ordinary conversation, which this data supposedly simulates, pronunciations corresponding to x before a consonant are regularly phonetic [s]. Spectrographic data for extraña and excursión do indeed bear this out in general. Only one of twenty-one informants used velar stops in these words. That is to say that less than 5 per cent of the forty-two potential occurrences of velar

elements materialized in these two words. This contrasts with the situation found in the five words which have no consonant after x. In those five words 80 per cent of some 100 occurrences showed a velar element.

Navarro goes on to say that the [s] is always voiceless in words whose written forms contain x. That seemed to be true in most of the acoustic data but not in ten occurrences of extraña or excursión in which the indications of voice resonance or relatively strong vocal cord activity continued beyond the very lightly recorded indications of brief sibilant friction (figures 38 through 42).

Figures 38 and 39 show very hazily a small bit of sibilant friction and an indefinite onset of the following stop with indications of voicing at the base line continuing slightly beyond the striated indications

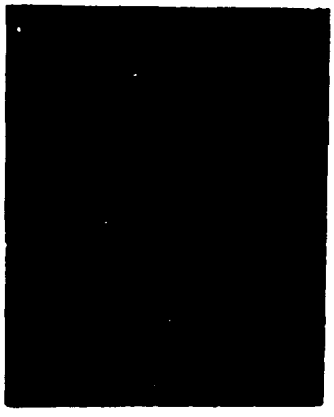
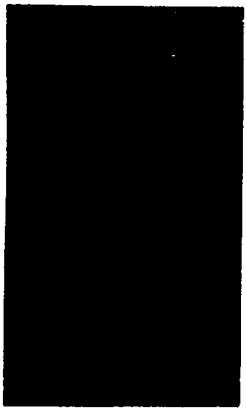


Figure 38. Informant I, sentence 5. The approach of formants 2 and 3 is like that found in approaches to velar stops and fricatives. The light indications of friction at this frequency level probably originate farther front in the mouth, though any semblance they may indicate to the production of voiceless /s/ is dispelled by the fact that baseline indications of voicing continue at least as long as any of the higher frequency noise.

of friction at the level of formants 2 and 3. Figure 40 shows a very sharp onset of the following [k] but the baseline is dark, as it is for



e x t r a ñ a

Figure 39. Informant XIII, sentence 5. The formants here show no velar activity. The light indications of friction without doubt represent only sibilant type noise. It is seen that voicing is predominant over and continues beyond any detectible sibilant friction.



la e x c u r sión

Figure 40. Informant IV, sentence 7. The [a] is represented clearly by three formants. The sibilant is here characterized by the rapid decrease in energy in formant 1 and the approach of formants 2 and 3 as is usual before [k]. The black baseline and the vertical striations above indicate voicing throughout almost all, if not all, of the sibilant segment.

a voiced stop consonant, to within one half centisecond of the onset of that [k]. Markings in the higher frequencies also indicate strong vocal cord vibration throughout the production of the sibilant. Figure 41 shows the same segment as recorded from a woman's voice. The striations

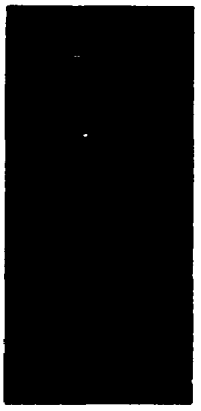


Figure 41. Informant XI, sentence 7. This illustration of a high pitched voice shows sibilant friction in which the vertical striations are not as apparent, but the presence of voicing is detected from the solid bar at the base.

indicative of voicing are much less pronounced and even disappear in the frequency range above 1000~. There remains little more than an indication of friction noise in the high frequencies. But the base retains the heavy black marking so characteristic of voiced sounds. Figure 42 contrasts a sibilant, corresponding to written x, before [k], where

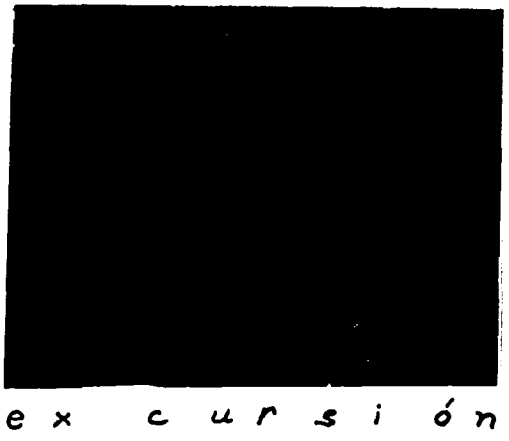


Figure 42. Informant XII, sentence 7. Contrasted here are two sibilant sounds. The first, corresponding to x, has only slight sibilant friction and clear voicing. The second has strong sibilant friction and is predominantly voiceless.

indications of voicing outweigh indications of unvoiced friction, with a following sibilant in which the reverse is true.

Figure 43 illustrates a different extreme in the pronunciation of the word extraña. Here the sibilant corresponding to x shows very much more of friction than of vocal cord activity.



Figure 43. Informant II, sentence 5. As illustrated here, the preceding figures cannot be considered to represent the only acoustic type corresponding to x. It is very common to find some clear representation of sibilant friction predominant over indications of voicing.

Navarro writes that x before a consonant is like a simple s. In writing about s in turn, he states that it may become voiced only in syllable final position before a voiced consonant. Figure 44 is a sample (taken from the recorded material of the same informant who provided figure 43) which illustrates a predominantly voiceless sibilant corresponding to s in the written language and occurring in syllable final position before a voiceless consonant. However this contrasts with the evidence presented in figure 45 where s in the final position of a stressed syllable shows all the characteristics of what we normally call a voiced sound. The material for this illustration was purposely chosen from the records of the same informant who in different phono-

logical environments pronounced uncommonly long and clearly voiceless sibilants as illustrated by figures 43 and 44.

The records of the words extraña and excursión also showed a marked difference from the others in the duration of the sibilant friction that was present. The average was approximately 6 cs. for these words while

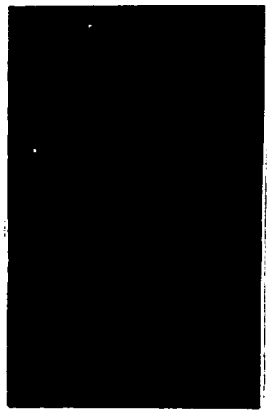


Figure 44. Informant II, sentence 15. Sibilant friction is clearly more conspicuous than indications of vocal cord activity throughout most of the sibilant preceding a voiceless stop.



Figure 45. Informant II, sentence 15. Sibilant friction is accompanied throughout by indications of vocal cord activity in this sibilant preceding a voiceless stop.

that corresponding to written x in the other seven words was approximately 10 cs. Even in the words auxiliar and exhaló where the occurrence of the sibilant had no connection with the stressed syllables the average length of the sibilant portions was 10.5 cs.

The explanation would seem to be that, in the former, the following voiceless consonant cuts the sibilant short much as in English a voiced continuant, especially a vowel, is shorter in a syllable closed by a voiceless sound. It will be recalled also that the average length of unstressed [s] in the sequence aliarnos a was 8.6 cs. while that in tuvimos que, with a voiceless stop following, was only 6.4 cs.

In the words exceso and excepción, if the hypothesis concerning fusion of [s] and the seseo form of [θ] is correct, we can expect a consonant to be pronounced after the pronunciation corresponding to x only in the case of speakers who use [θ]. It was found that the spectrograms bore this out strongly in the word exceso. In this word a velar element averaging 5.9 cs. was found in seventeen out of eighteen seseo speakers. This was comparable to the 80 per cent occurrence of velar elements present in pronunciations corresponding to intervocalic x. Significantly, it contrasted sharply with the 5 per cent occurrence of velar elements in pronunciations corresponding to x before any non-suspiciously actualized consonant. In other words, it is indicated by this evidence that seseo speakers have no phonetic actualization corresponding to the written c after x in exceso. Their pronunciations corresponding to written x in exceso parallel those corresponding to intervocalic x and not those corresponding to x in words like extraña which have a phonetically actualized, non-homophonous consonant following the pronunciation corresponding to x.

The picture was different in reference to the word excepción. Here it was found that only eight of nineteen seseo speakers, or 42 per cent showed clear spectrographic evidence of a velar element. The reason once more might be the difference in stress placement since the unstressed intervocalic pronunciation in auxiliar also shows only 65 per cent occurrence of velar elements. Excepción obviously occurs with divided usage in the acoustic data. The velar element tends most often to be absent.

It is strange indeed that the very informants who use [θ] also had velar elements in all four occurrences of pronunciations corresponding to x in exceso and excepción. Figure 46 illustrates the word extraña as spoken by one of the non-seseo informants. It is clear that there is no velar element, much less a full voiceless velar stop before the sibilant.



Figure 46. Informant XVIII, sentence 5. There is no velar element preceding the sibilant which is followed in this spectrogram by a stop consonant.

Figure 47, by the same informant, shows equally clearly the presence of a stop or very light fricative element before the onset of the sibilant. The fact that it is a velar element is demonstrated by the uniting of formants 2 and 3 as is true elsewhere before [k].

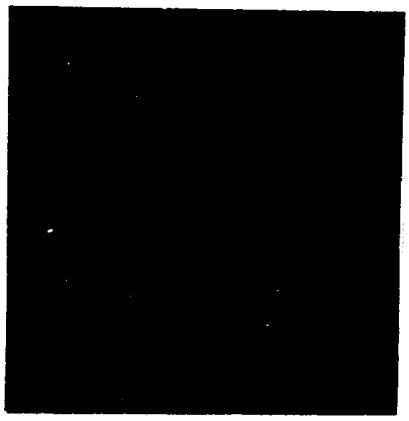


Figure 47. Informant XVIII, sentence 15. Formants 2 and 3 approach each other at the left in this spectrogram. A velar stop is indicated. The relatively intense friction following the velar element corresponds to a physiological sibilant which is followed by lighter markings corresponding to interdental friction.

It would not be in full harmony with the tone of Navarro's writing, which would have Castillian as the focal area of linguistic influence, yet, in this case, it seems reasonable to suggest that the Castillian dialect speakers were influenced, in their pronunciations corresponding to written x before ce, by seseo speakers, who have fused contiguous sibilants and who have, on the basis of historical retention, preserved a velar element whenever it was followed by not more than one contiguous consonant sound.

The [s] element in these words is significantly longer than in extraña and excursión. In exceso the average length of the first sibilant is 10.6 cs. and in excepción it is 8 cs. This is not, however, any longer than for the words in which the sibilant precedes a vowel. We find that in the word éxito the average duration of the [s] is 8.9 cs.; in exhaló, 10.6 cs.; in auxiliar, 10.5 cs.; in examen, 10.3 cs.; and in exacto, 11 cs. It can still be said, therefore, in support of the earlier

hypothesis, that spectrographic measurement of the length of sibilants among seseo speakers indicates that there is almost certainly no additional mora of sibilant length in the words whose written forms show x followed by ce.

Pronunciations which correspond to intervocalic x in the written language, according to Navarro, will be [gs] or, with [g] "en parte ensordecida," [gxs]. As is very often true, he once more proves his extremely keen observation. It is not fully clear, however, whether Navarro meant this as a statement of averages or of limits. The words "a veces," as quoted from Navarro at the beginning of this section would indicate that it was meant to be a statement comprehending the occurrences of "correct" pronunciations of words whose written forms contain an intervocalic x.

However that may be, the spectrographic data under study show, in the word éxito, on the average, 3.4 cs. of voicing at the onset of the segments corresponding to x. Eighteen of the twenty occurrences have a velar element which has an average duration of 4.4 cs., or 1 cs. more than that of the voicing. Since the sibilant segment is, on the average 8.9 cs. long, this means that in the 18 cases with velar elements, an average pronunciation corresponding to x in éxito consists of 3.4 cs. of voiced velar [g] or [ɣ], 1 cs. of voiceless velar [x] or [k] and 8.9 cs. of voiceless sibilant [s].

Very important in another respect are the two cases of éxito which did not show velar elements. Here the "average" pronunciation would be quite different. It would consist of 3.4 cs. of voiced sibilant [z] and 5.5 cs. of voiceless sibilant [s]. That this hypothetical "average" pronunciation may occur as an actual pronunciation is attested by figure 48 which shows the sibilant in auxiliar partly voiced and partly voice-

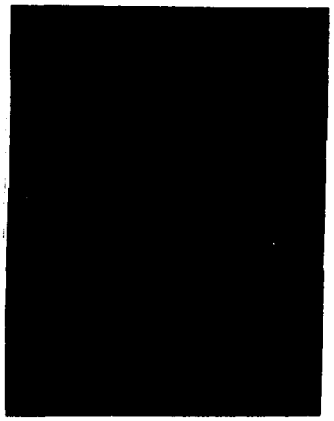


Figure 48. Informant XII, sentence 14. The pronunciation corresponding to written x consists of approximately 8 cs. of sibilant. Of this, approximately 3 cs. are voiced and 5 are voiceless.

less in almost the precise ratio of the "average" mentioned above.

Navarro has observed this initial voiced element in pronunciations corresponding to x when that pronunciation includes a velar element. He says that it is regularly a weak and relaxed [g] which at times is partly voiceless. Aside from the fact that these data show that to be true not only at times but almost invariably, there is also the fact that they consistently show a very nearly equal amount of voicing in all pronunciations corresponding to x. This includes, significantly, the pronunciations of extraña and excursión in which velar elements were all but absent. The implication, therefore, is that Navarro's statement about voiceless sibilant segments corresponding to written x does not cover the evidence of these data. If the initial voicing in velar elements is of any significance then the initial voicing on the sibilants where the velar is lacking would seem to be of equal significance.

Figure 49 illustrates the three types of combinations that occur in the averages of the data. High and low extremes in absolute amount of voicing are shown by the graph in the words exhaló and auxiliar.

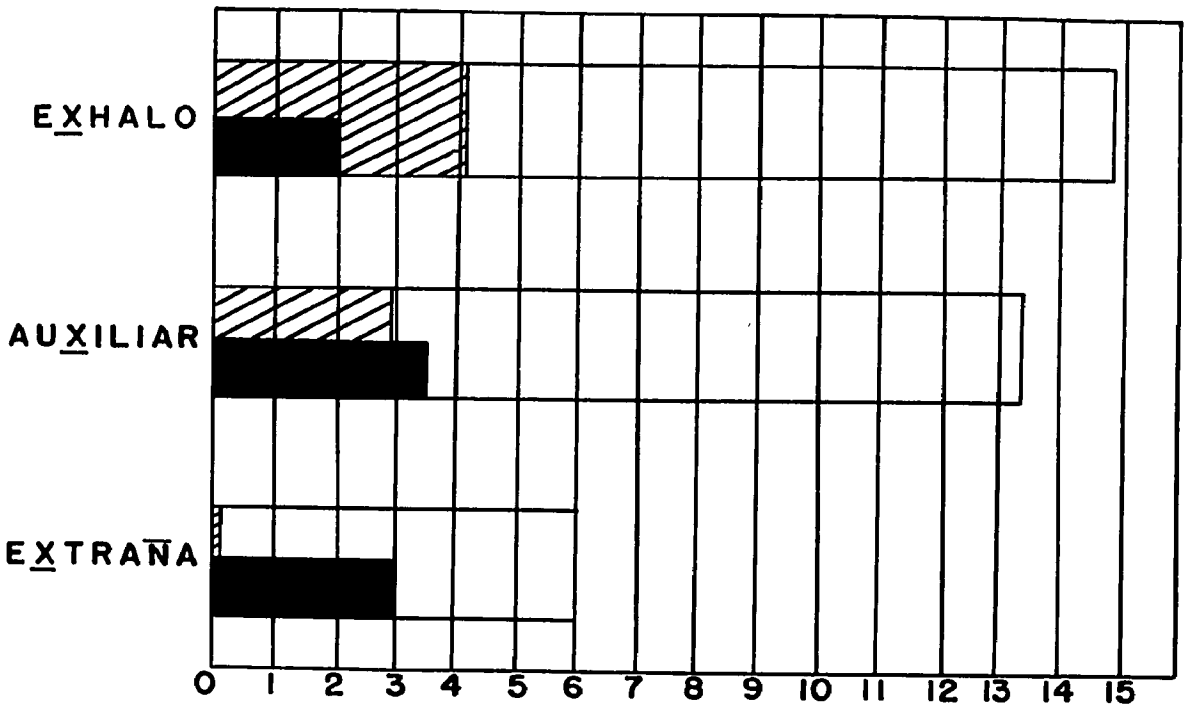


Figure 49. This graph illustrates three types of phonetic combinations that occur in pronunciations corresponding to x in the written language. Each bar represents the mean for twenty-one occurrences of pronunciations corresponding to x in the word shown at the left. Length is shown on the horizontal scale in centiseconds. Time progresses from left to right. Each bar represents the total mean length of pronunciations corresponding to x in the word indicated. The solid black indicates voicing; lines slanting downward to the left indicate a velar element; and enclosed blank space indicates a sibilant element.

The average or composite picture of the twenty-one pronunciations of exhaló showsthat voicing extends approximately half way through the durations of the velar element but not into the sibilant element. Individual pronunciations vary from this, of course, but in only two cases of the twenty-one does it appear that any voicing extends into the sibilant and then only 1 or 2 cs.

For auxiliar the graph shows voicing covering the entire duration of the velar segment. Individual pronunciations, however, show that just over half of the samples with velar elements have that element entirely voiced. In the remaining cases the velar elements are often almost entirely voiceless. The seven samples which do not contain velar elements retain an average of 2.7 cs. of voicing in the sibilant. This is hardly less than the amount found in the pronunciations that do contain velar elements.

The third type graphically illustrated in figure 49 is represented by the averages of the measurements of pronunciations corresponding to x in extraña. Here the voicing continues throughout approximately half the duration of the sibilant although the total amount of voicing is actually less than in the case of auxiliar. Again, individual pronunciations vary considerably. In ten of twenty-one pronunciations, voicing is spectrographically detected as lasting throughout, or even beyond, the recorded evidence of sibilant friction. In only two of twenty-one samples does the voicing last less than 30 per cent of the duration of the sibilant and one of these is the sample which includes the clear voiceless velar stop before the sibilant (figure 50).

Excursión patterns almost exactly like extraña. Excepción very closely parallels auxiliar. Exceso and examen are very much like exhaló. Exacto and éxito fall between the patterns for exhaló and auxiliar.

These data therefore clearly bear out Navarro's belief that Frenchmen and Englishmen who pronounce [ęgzámen], with a voiced sibilant, are in error. Certainly if only two of twenty-one representative pronunciations occur with any identifiable voicing on the sibilant of examen and if that voicing averages only two thousandths of a second on sib-



Figure 50. Informant XX, sentence 5. This spectrogram represents a relative rarity in the manner that it takes exception to the impressions given by the bar graph in figure 49. The intersecting of formants 2 and 3 followed by a gap represents the occurrence of a voiceless velar stop. The spike and other bits of blackness represent the relative position of a voiceless sibilant - one of only two which were found to show less than 30 per cent of voicing in this word.

ilants which themselves have an average total duration of over one tenth of a second, it may be assumed that a non-native would be detected as having a foreign accent if he consistently voiced the entire sibilant - which averages over fifty times as much of it as a native speaker voices.

However, ten of forty pronunciations corresponding to x in the words extraña and excursión showed no indications of voiceless sibilant at all. The spectrographic indications of voicing surpassed in intensity and length those of sibilant friction in the word extraña as pronounced by informants I, III, IV, X, XI, XIII, and XV, and in the word excursión as pronounced by informants IV, XI and XII.

As Navarro says, the h in the written form of exhaló does not correspond to any variation in pronunciation which affects the intervocalic consonant or consonant cluster. The average amount of voicing (2 cs.)

is low but no lower than in exacto. The average duration of the sibilant (10.6 cs.) is comparable to that in auxiliar (10.5 cs.). The average length of the velar element, when it occurs in exhaló (twenty times of twenty-one possibilities), is exactly the same (4.5 cs.) as that of auxiliar.

The last statement also brings to attention an observation by Navarro that seems amazing in the light of these data. It was shown that examen, exhaló and éxito had velar elements in approximately 95 per cent of all occurrences. Yet, the word exacto showed velar elements in only about 50 per cent of the pronunciations and the word auxiliar in only 65 per cent. The words extraña and excursión showed approximately 5 per cent occurrences of velar elements while exceso showed 95 per cent occurrence and excepción registered 42 per cent. An expected correlation between occurrence of velar elements and stress placement did not materialize. But Navarro's last sentence in the quoted paragraph stated that correct pronunciation permits, in general, [s] without a preceding velar element in the words exacto, auxilio and auxiliar.

Even Navarro's statement, of course, does not include, much less account for, the infrequent occurrence of the velar element in excepción; nor does it account for the 5 per cent of exceptions in every other case.

It is quite probable that a usage viewpoint would be profitable here. It could then be stated that the acoustic data indicate that the combination /k/ + /s/ is subject to divided usage in Spanish. In what might be the sequence /k/ + /s/ + consonant, the predominant, and almost exclusive usage is /s/ + consonant. Relative to the sequence /k/ + /s/ + vowel, usage is approximately equally divided, in these data, on inclusion and exclusion of the velar element in the words exacto, auxiliar

and excepción. Elsewhere the velar element is predominantly retained.[1]

Perhaps the most interesting and provocative observation arising from this exploration of pronunciations corresponding to x is that of the occurrence of voicing in these pronunciations.

It is to be observed that there is an average of 2 cs. of voicing on pronunciations corresponding to x in exacto and that there is an average of 5.8 cs. of velar element in eleven occurrences of exacto. Ten occurrences have no velar element in the portion of the word under discussion. It is likewise to be observed that éxito occurs on eighteen occasions with a velar element averaging 4.4 cs. Twice it occurs without a velar element. The average amount of intervocalic voicing in the pronunciations corresponding to x in this word is 3.4 cs.

It is further to be observed that the average amount of voicing in the pronunciations corresponding to x in extraña and excursión is 3 cs. and here there are, with one exception, no velar elements.

Navarro readily observes the voicing in exacto and éxito, etc., when they are pronounced with velar elements, but he apparently does not observe the voicing that would seem to be equally observable (at least to a person whose perception is not compromised by phonemic reaction) in the sibilants of extraña, of excursión, and of all the other x words when they are pronounced without a velar element.

---

[1]Cf. Dwight L. Bolinger, "Evidence on x," Hispania, XXXV (Feb., 1952), 58, 61. The acoustic data used for this thesis show 95 per cent occurrence of some type of velar element in exceso. Ninety-seven per cent of Bolinger's respondents testified that they say [ks] in exceso. These data show evidence of velar sound of some kind in 65 per cent of the occurrences of auxiliar and in 42 per cent of the occurrences of excepción. Forty-four per cent of Bolinger's respondents believe that they use [ks] in auxiliar when used as a verb and 53 per cent believe they use the affricate in excepciones.

C. [gs], [ks] and [s] corresponding to cc.

> §128...El grupo cc se pronuncia ordinariamente [gθ]; la [g] en este caso es débil y relajada, y además bajo la influencia de la [θ] siguiente, suele resultar en parte ensordecida: ...En formas fuertes o enfáticas cc se pronuncia [kθ]: ...el habla vulgar, por el contrario, reduce este grupo a una sola c: ...[1]

Navarro says that the written group cc corresponds, in pronunciations, to [θ], [kθ] and [gθ], in which [g] in the last combination, under the influence of the following voiceless [θ], may become partly voiceless.

There seems to be a bit of vacillation here because this paragraph implies that since [g] is pronounced in a weak and relaxed fashion, it easily comes under the influence of the following voiceless sound and part of the voicing of [g] is lost in the process of regressive assimilation of the voiceless characteristic. But elsewhere in this paragraph, and also in paragraph 129, he states that this combination occurs fully voiceless in strong, emphatic or learned pronunciation. This puts him in the position of saying that the velar element is voiceless in careful pronunciation, voiced in ordinary conversation except that in ordinary conversation it may sometimes be partly voiceless at the end, due this time, not to fairly strong, emphatic or learned pronunciation, but to the following voiceless [θ] or [s].

The evidence in the spectrograms makes it desirable to reword this slightly if it is to apply to these data. It is of course necessary to state that there is divided usage in the words under discussion. Further description may then account for the facts. In the group /k/ + /s/, the velar element may occur as a fully or almost fully voiceless sound.

---

[1] Ibid., 140.

In some pronunciations of the words where the velar element is found in the samples of a few or even most informants it may nevertheless be entirely dropped by others. It may also be present and assimilate voicing from the preceding vowel throughout part or all of its length. This explanation is more consistent with the data than one which says that the pronunciation corresponding to cc is either voiceless (always in emphatic, strong or learned pronunciation), or voiced (always in ordinary conversation), and that, if it is voiced, it may assimilate some voicelessness from the following [θ]. This explanation avoids a correlation of voiced and voiceless with emphatic and non-emphatic since that correlation is not borne out in the data. It also avoids categorizing the informants' language usage as "habla vulgar," "conversación ordinaria," and "formas fuertes" on a basis which must of necessity be questionable in this situation as can be seen, for example, from the fact that all informants were given the same instructions and all are highly educated, yet 14 per cent of the spectrograms showed absolutely no trace of a velar element in inyección.

It seems possible, here, to supplement Navarro's statements. Since he has stated that seseo consists of substituting [s] for [θ], it may be said that, according to him, seseo speakers would pronounce [s], [ks] and [gs] or [gxs] corresponding to written cc.

A survey of the data shows striking similarities between the pronunciations corresponding to cc in inyección and those corresponding to x in the previously described words. Spectrograms correlating with [s], [ks], [gs] and [gxs] may be found for both.

As before, in the similar sequence corresponding to x, there is evidence of less duration of voicing (2.8 cs. per occurrence on the

average) than of velar elements (3.7 cs. each in sixteen of nineteen potential occurrences). These figures are precisely equal to those obtained as an average for intervocalic pronunciations corresponding to x.

The evidence of [s] friction indicated an average duration of 11.9 cs. in the pronunciations corresponding to cc. This is slightly, but by statistical analysis not significantly, higher than the 10.5 average for [s] corresponding to x.

The velar + [s] sequence corresponding to cc does not contrast in the amount of voicing or in the length of the non-sibilant element with the [k] + [s] sequence which corresponds to written x. It does, however, contrast in amount of voicing and in length of non-sibilant element in the sequence represented by pc in excepción. For [p] + [s], there are found in the data averages of 3.3 cs. of voicing and 6.8 cs. of bilabial stop. And the bilabial element is found as a stop, i.e., with at least a small portion of clear gap on the spectrograms, in almost all occurrences of the word excepción. In the two exceptions, a voiceless velar stop and a voiced velar fricative are substituted for the bilabial element, but it is never simply omitted as is the velar element in a potential /k/ + /s/ sequence.

As stated, the average of voicing in all pronunciations corresponding to written x equals 2.8 cs. Exactly the same figure is found to be the average duration of voicing in the pronunciation corresponding to cc in inyección. This contrasts with an average of 4 cs. duration of voicing in the pronunciation corresponding tp pc in excepción.

Navarro indicates that what is, in careful or strong pronunciation, [p] before [s] may in ordinary conversation readily shift to [b̥]. But, he states, this [b̥] may be partly voiceless at the end. The data indi-

cate an average duration of 7.3 cs. for /p/ of which an average of 4 cs. is voiced.

As in the cases of pronunciations corresponding to x and cc, Navarro's statements cover the correlatable spectrographic data remarkably well. In view of this it may seem surprising that he does not mention the relatively greater length and certainty of occurrence of bilabial over velar in their combinations with [s]. From a phonetic point of view these could be expected to be much more noticeable than the individual voiced and voiceless segments which occur within the bilabial and velar segments. The average difference in length is considerable: over 6.6 cs. for /p/ in excepción; under 3 cs. for /k/ in inyección. The velar element in inyección is detectable only seventeen times in twenty-one occurrences of this word whereas the bilabial element is clear in all twenty-one occurrences of excepción. Furthermore the velar element in inyección is more susceptible to variation in the manner of articulation. It occurs as voiceless stop, voiced stop, voiceless fricative, voiced fricative or without spectrographic indications of any of these. The bilabial element in excepción, however, very regularly occurs with substantial indications of voiceless bilabial stop.

D. [z] and variations.

> §94...En la conversación ordinaria, la z débil final de sílaba o de palabra, en contacto con una consonante sonora siguiente, se hace también sonora, resultando un sonido muy semejante al de la th inglesa en their, this. La pronunciación lenta, fuerte o enfática impide parcial o totalmente esta sonorización.[1]

---

[1] Ibid., 94, 95.

> §107...La s sonora aparece únicamente, en nuestra lengua, en posición final de sílaba, precediendo inmediatamente a otra consonante sonora; en cualquier otra posición su presencia es anormal y esporádica. Es siempre, asimismo, una articulación breve y suave; la pronunciación lenta o fuerte impide su sonorización, reapareciendo en su lugar la s sorda...
>
> En el grupo sr (israelitas, los reyes, dos reales), la s se sonoriza como en los casos precedentes; ...y produciendo propiamente, en vez de la z ordinaria y regular, una [ɹ], o sea una [r] fricativa...otras veces, en pronunciación relativamente fuerte, la s se pierde por completo, aumentándose a manera de compensación, las vibraciones de la [r] siguiente.[1]

Although Navarro, in paragraph 94, speaks about a [d̯] similar to that in English they in the material to be considered here, this section as a whole is entitled "[z] and variations" in accord with most of the data used in this study.

If we apply the foregoing quotations on fricative voicing to the occurrence of such voicing in this body of material, of which 90 per cent represents seseo speech, the sibilant type (including seseo forms) that can be expected regularly preceding voiced consonants is the voiced one. Before vowels and voiceless consonants the voiceless variety can be expected: voicing would be considered abnormal and sporadic here.

With reference to seseo, it can be said here that the data show no significant difference in the relative strength and duration of the voicing (in the sibilant or fricative element before a voiced consonant) which shows any regularly recurrent correlation with s and z spellings before d as in luz dorada vs. escollos de. The sibilant, the following sound, or both, are voiceless in 25 per cent of the data corresponding to each of the combinations as found in luz dorada, juzgar and escollos de, without respect to the spelling. The fact that spelling does not

---

[1] Ibid., 108, 109.

influence the pronunciation of the sequence /s/ + voiced consonant is demonstrated by figures 66 through 69. These figures represent pronunciations which correspond to written s + d in the sequence escollos del. It can be observed that the features illustrated are all found in the figures corresponding to the written z + d in luz dorado and z + g in juzgar (figures 50 through 65). Owing to this and to the fact that spectrographic representations corresponding to differences in seseo and non-seseo speech have already been considered, no further reference will be made to this matter when variantly spelled material is selected for comment from the data.

Much of the further comment will be in the form of description of types. The reason for dealing with these data by type and through illustrations of the type will be apparent. Measurement from one analogous transition point to another, highly practicable in the measurements of /č/, is less useful if analogous transition points do not recur often in a body of data.

Some of the following figures illustrate the difficulty of finding analogous transition points. Figure 51 shows a combination of grooved sibilant and ungrooved or flat tongue dental corresponding to zd in luz dorada. There is very light friction above 2500~ in the portion of the spectrogram that corresponds to the voiced sibilant. This is followed by four or five vibrations of the vocal cords with no evidence of friction above 400~. This is a spectrographic indication of what ordinarily corresponds to a physiological voiced stop.

A fairly close estimate can be made concerning the borderline between [z] and the following stop. It is rather more difficult to find a suitable borderline between [z] and the preceding vowel than between



Figure 51. Informant V, sentence 5. It is possible, with relatively minute disagreement, to establish for this sample a line of division between the markings that represent a predominantly sibilant sound and those that represent a predominantly stop sound. It is similarly not difficult to agree upon a dividing line between [d] and [o]. Compare this however with the difficulty of pinpointing a separation between [u] and [z] in some way that might be useful in all, or nearly all, other pronunciations of this sequence.

[z] and the stop. If all spectrograms resembled this one, it would be simple to set up some arbitrary rule which would provide a basis for measuring and obtaining a close estimate of the duration of sibilant voicing.

However, figure 52 illustrates a type in which, while the borderline between [u] and [z], the grooved tongue articulation, presents the same relatively easily solvable problem as before, the borderline problem between [z] and the following flat tongue articulation, apparently midway between a stop and a fricative, is no longer analogous to that in figure 51.

Figure 53 shows a fully fricative transition of mutually assimilated varieties of [z] and [d̪]. The difficulty of establishing borderlines or



Figure 52. Informant IV, sentence 5. Establishing a borderline between [u] and [z] in this sample presents the same problems as it does in figure 51. But the transition from the sibilant to the following dental fricative or stop is a more gradual progression than in the preceding example and consequently more difficult to divide.



Figure 53. Informant I, sentence 5. This illustrates a high degree of mutual assimilation.

starting points for measurement which would be analogous to those used in the preceding figures is a deterrent to the use of the same type of statistical procedures that were used in Chapter III.

Figure 54 indicates a sequence in which the grooved tongue activity is strongly voiced, and, judging by the markings surrounding it, the very

lightly voiced portion (therefore probably auditory "voiceless" by contrast) must be assigned almost wholly to the flat tongue sound. Although the flat tongue sound looks almost like a stop in view of the fact that there are no indications, for about 2 cs., of any energy above 500~, it does not have the features of aspiration which usually accompany stops that are fully voiceless.

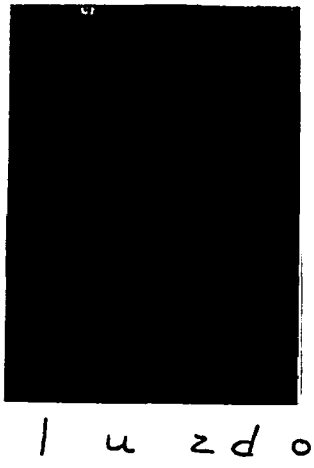


Figure 54. Informant VII, sentence 5. Strong voicing continues beyond the indications of sibilant noise but diminishes abruptly in the area corresponding to the flat tongue pronunciation.

This is perhaps better illustrated by the same linguistic sequence recorded by another informant and shown in figure 55. Here are found three spans of 2 cs. each. The first of these spans shows evidence of some noise only below 200~. The second is completely blank and the third shows aspirated release of the voiceless stop before the onset of vocal cord vibration for the following vowel.

It is important to observe that in a substantial number of these illustrations the direction of assimilation is clearly the reverse of the only assimilation that Navarro mentions, namely, regressive voice assimilation on the part of the sibilant. Since closely analogous borderlines cannot be drawn on all of these data and since therefore

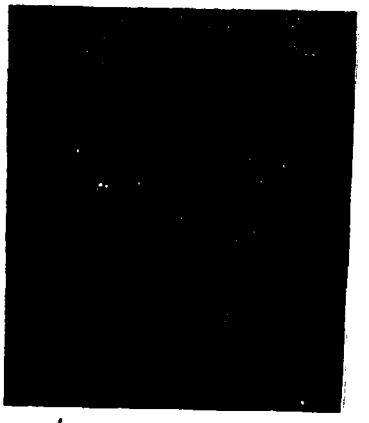


Figure 55. Informant II, sentence 5. The tendencies illustrated in figure 54 are here shown in exaggerated form. In a sample of this type it is easy to establish a borderline between the dental and the following vowel, but that between sibilant and stop remains elusive.

statistical quotations on comparative amounts of regressive and progressive voice assimilation are not likely to be very accurate, it is considered wiser to illustrate the occurrence of this phenomenon with selected samples.

Figure 56 clearly shows a voiceless [s] followed by a dental flat tongue fricative which is equally voiceless for approximately the first half of its duration. If the pronunciations corresponding to written d occur elsewhere with voice, then this is certainly progressive assimilation of voicelessness. It is the opposite of the somewhat, but not overwhelmingly more common, occurrence of regressive assimilation[1] of voicing on the part of the sibilant.

---

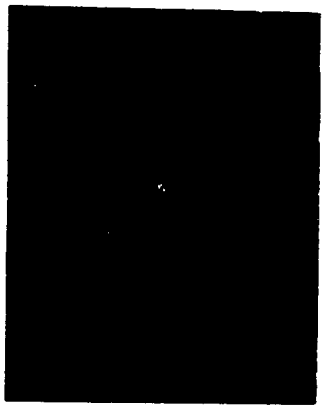
[1] From the evidence of this body of acoustic data it is quite tenable that what is here ostensibly called "regressive assimilation" is itself more likely to refer largely, if not wholly, to progressive assimilation of voicing from preceding vowels. Compare the data described on page 90 and in the remainder of this chapter.



Figure 56. Informant XVI, sentence 5. The sibilant is clearly voiceless. The following flat tongue dental also has characteristics of a voiceless sound. Indications of voicing, though not entirely suppressed, are not as heavy as in sounds which are unquestionably voiced. The spike before the onset of vowel characteristics is reminiscent of the aspiration found in voiceless stops. Since this is regularly a voiced stop or fricative in the data, it seems to be clear evidence of progressive assimilation of voiceless characteristic.

The ultimate in progressive assimilation is shown in figure 57. Here neither voicing nor friction - nor anything - remains of the flat tongue sound. If this were representative of the language as a whole, the flat tongue sound under observation could be said to be historically absorbed by the voiceless grooved fricative through complete progressive assimilation.

A more moderate picture is presented by figure 58. The formants here indicate that the transition from grooved to flat articulator takes place at a time corresponding approximately to the onset of heavy friction which centers around the 3000~ frequency level. It may be said to continue for approximately 4 or 5 cs. before and during the rapidly decreasing frequency level of formant 3. Again, over half of this is voiceless.



l u zd o r a d a

Figure 57. Informant VI, sentence 5. It becomes so difficult to separate grooved and flat tongue sounds in this sample that most phoneticians would probably settle for a description of this as a single segment. Characteristics found in other sibilants or grooved tongue sounds predominate and it would therefore be classified as such. An additional statement might be made to the effect that complete progressive assimilation of these characteristics caused virtual disappearance of evidence of consonantal flat tongue activity.



extraña l u zd o r a d a

Figure 58. Informant III, Sentence 5. The formants indicate that the sibilant occurs largely during a time of strong vocal cord activity. This vocal cord activity could be considered to show progressive assimilation of that characteristic as found in the preceding vowel. The voiceless characteristic, continuing into what the formants show to be the following flat tongue sound, could be considered to show progressive assimilation of that characteristic as found at the end of the preceding sibilant.

The immediately preceding spectrograms illustrate pronunciations corresponding to written z at a word boundary. The succeeding ones, except for figures 60 and 61, will illustrate comparable data within a single word in which pronunciations corresponding to written g follow those corresponding to z.

Figure 59 represents the pronunciation of juzgar by a non-seseo speaker. The pronunciation corresponding to z is accordingly produced by a flat articulator against the teeth. The transition from [d̪] to [g] is well marked and can be established with a very small margin of indeterminacy. Voice assimilation on the part of the dental fricative is complete and covers approximately 7 cs.

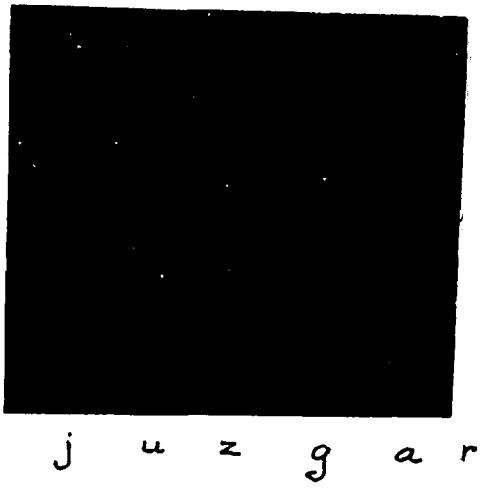


Figure 59. Informant XVII, sentence 2. The boundary between front and back tongue fricatives is here marked by a type of open transition.

Figure 60, by contrast, shows a pronunciation of luz dorada by another non-seseo speaker in which an accurate separation of the two dentals is impossible because of the unbroken gradual transition from the beginning of the flat tongue voiced dental fricative to the onset of the release into the vowel [o].



Figure 60. Informant XVIII, sentence 5. This represents a steady progression of smooth movements from the beginning of [ũ] to the end of the dental stop which, at the end, shows voiceless stop characteristics.

Figure 61 shows that even a non-seseo speaker, the same one who produced figure 59, may have a clear separation in this context. There is a clear break between the pronunciations corresponding to luz and dorada. The initial bit of voicing in [θ], as evidencedon the base line, is obviously not assimilated from, nor produced in anticipation of the following voiced sound. It seems again to be rather progressive assimilation or simply a continuation of vocal cord vibration a bit beyond the transition area of cavity resonance for [u] to local friction resonance for [θ].

Returning now to pronunciations corresponding to z within a word, it is found that these, too, may be wholly, partly or not at all voiced. As at word boundaries, a sizeable number of examples illustrate the progressive assimilation, not mentioned by Navarro, in which the sibilant (for seseo speakers) is voiceless and this voicelessness is carried over into the following velar sound as illustrated in figure 62. There is



Figure 61. Informant XVII, sentence 5. Fricative and stop dentals are here separated by silence which, by reason of its brevity, undoubtedly corresponds to an unreleased [t]. The dental fricative and stop, inseparable in figure 59, are here easily delineated.



Figure 62. Informant III, sentence 2. Due to progressive assimilation the sibilant is mostly voiced. The low frequency friction, which indicates the velar fricative, occurs with only the lightest semblance of vocal cord activity which is, of course, indicative of a physiological voiceless sound.

at least a 2 cs. span of voiceless transition in which dental sibilant friction in the high frequencies, no longer showing the striations produced by vocal cord vibration, is decreasing in intensity while voiceless velar friction rapidly becomes predominant.

Figure 63 again shows progressive assimilation of the voiceless characteristic of [s] into the velar element, but it also shows somewhat more clearly the progressive assimilation of voicing from the vowel into the following sibilant element which may be said to begin approximately in the area of the vertical spike which seems to represent a click as

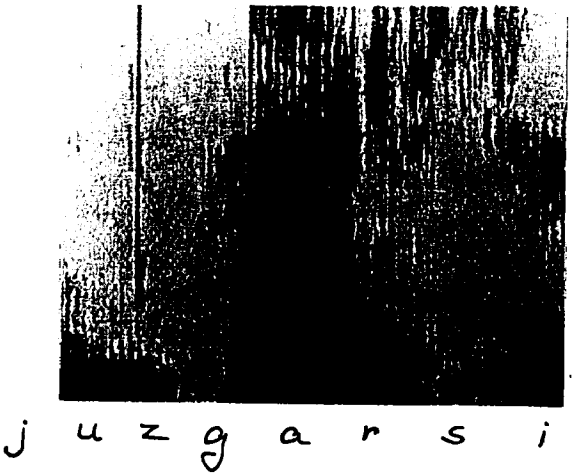


Figure 63. Informant VIII, sentence 2. Progressive assimilation of voicing into the consonant area is especially obvious here because the spike serves as an end marker for the vowel. 37 per cent of the data for juzgar and 29 per cent for luz dorada show double progressive assimilation of this type.

produced by closing the teeth upon each other. There is a voiceless velar stop of 3 cs. duration followed by 4 cs. of voiceless velar friction. Regressive assimilation of the type named by Navarro is entirely substituted by progressive assimilation both of voicing into the sibilant and voicelessness into the velar.

Figure 64 serves as an excellent illustration of what Navarro would call strong or emphatic pronunciation. Since the data were not gathered from unsolicited "ordinary conversation," the "strong or emphatic" label cannot be categorically denied. But it may be said that this occurred in a larger context which as a whole does not seem to be more emphatic than the other data provided by this or other informants.

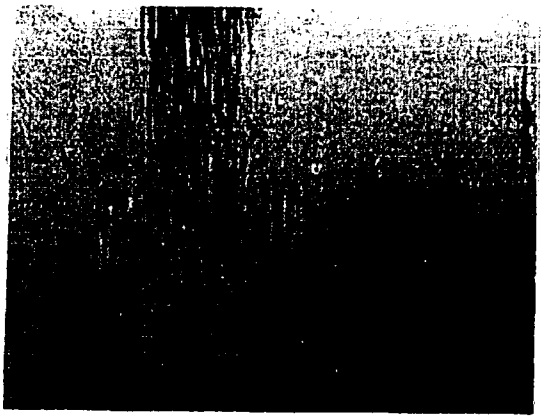


Figure 64. Informant XVI, sentence 2. This is an exceptional example of minimum assimilation relative to the allotments of time shown with homogeneous characteristics. The transitional friction is voiceless velar. It cannot ultimately be determined whether this is regressive assimilation of the point of articulation or progressive assimilation of a voiceless characteristic without some arbitrary rule for setting up boundary lines at fuzzy edges of this type.

The onset of sibilant friction is well delineated and it can be seen that some voicing continues for about 2 cs. The transition from sibilant to flat velar friction is again relatively sharply delineated and again it may be said that the voiceless characteristic continues for some 2 cs. after velar friction becomes clearly predominant. As wholes, of course, the sibilant would be said to constitute a voiceless sound while the velar would be called a voiced sound.

Figure 65 is especially interesting in that this phenomenon, when described, shows broad application and raises formerly unconsidered problems in the area of assimilation.

It is observed here that there is double assimilation. There is no evidence of friction that might be caused by a front sibilant. There is considerable evidence of velar friction with a flat articulator.

There is, in fact, a remarkable symmetry of the assimilated velar friction following [u] with the known velar friction preceding the [u]. This figure represents regressive assimilation of the type of fricative produced but it must equally be considered to represent progressive assimilation of a voiceless characteristic.

This mutual assimilation, producing what is customarily called one sound, raises the question of which sound assimilated something. It seems, in fact, that it cannot be said here that either one sound or the other assimilated anything but rather that a person's total speaking activity was reduced by one vocal gesture due to mutual complete assimilation of certain component parts.

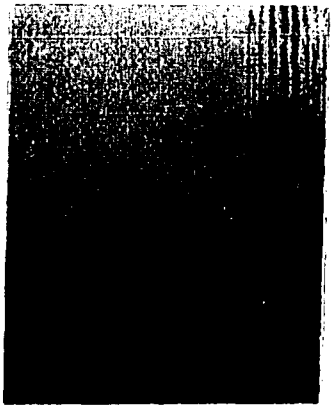


Figure 65. Informant X, sentence 2. This is related to figure 57 which illustrates complete assimilation. Here, again, two historic or potential segments are fused into one. In this example, however, the point of articulation is regressively assimilated while the voiceless characteristic is progressively assimilated with the result that we have a single phonetic voiceless velar segment for a potential voiceless sibilant and a voiced velar fricative.

The problem raised may be extended to borders between the peaks or centers of vocal gestures. Figure 64, for example, was described as having a relatively clear delineation of borders between sibilant friction

and velar friction. It was stated that there was progressive assimilation of a voiceless characteristic into the velar element. But if one began by measuring the total span of voiceless speech here he could justifiably say that the succeeding voiced sound which has only very light velar friction has undergone partial assimilation by the voiceless sound since the voiceless sound shows a portion of velar friction the intensity and voicelessness of which are not regularly characteristic of the sound corresponding to written g. Thus even here at least 3 cs. of transition can be equally well assigned to [s] or to [g].

The next five figures show phenomena of the general type described above but occurring in pronunciations that correspond to a different spelling for the sibilant.

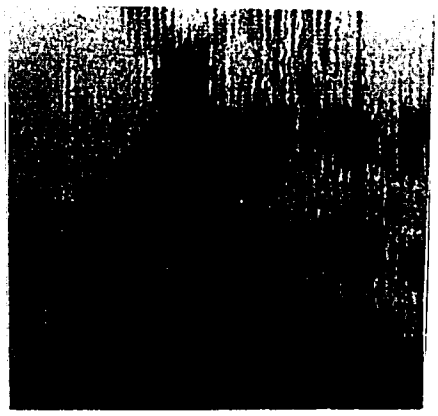


Figure 66. Informant III, sentence 15. Areas corresponding to two segments - one more [z]-like and the other more [d]-like - might be separated here but it would be difficult to do so on a basis that could be closely paralleled in all other comparable examples as they are illustrated by the figures in this section.

Figure 66 shows a completely voiced sequence with borderlines that are extremely difficult to fix. If an attempt were to be made, it might be maintained that the grooved articulator is quite prominent in the area where three vertical striations between formants one and two appear

dark and unbroken. It might further be said that the voiced dental fricative with flat articulator rapidly loses its prominence in the area of the appearance of the dark vertical striation which spans the entire frequency scale just before the final dip in formant two. Between these ends [z] and [d̪] are mutually assimilating vocal gestures and it is not clear whether we still have two segments or one.

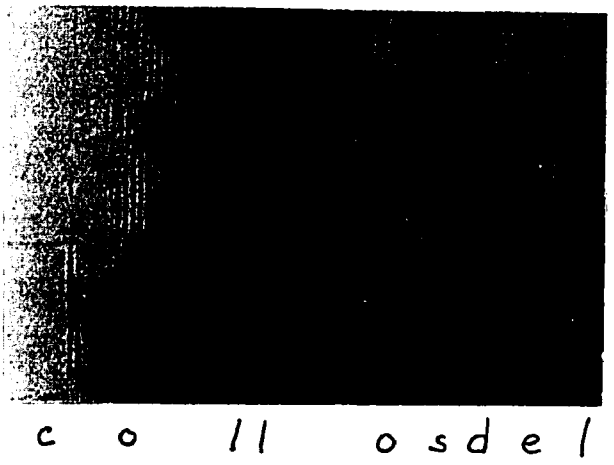


Figure 67. Informant VII, sentence 15. The dental produced with a flat tongue as articulator seldom gives so clear an effect of a complete stop within a breath group unless preceded by another vocal gesture which requires tongue front closure. In such cases it may be said that preceding closure is progressively assimilated. In this case the continuant nature, or the opening at the tongue front position, of the preceding sibilant is not progressively assimilated as it usually is.

This spectrogram also provides an interesting commentary on transition or border phenomena. The palatalized lateral continues its high front closure, i.e., high front closure is progressively assimilated, while the high tongue back position of the [o] begins or is regressively assimilated. The phonetic, non-phonemic, phenomenon resulting from these indeterminate borders is very much like a [d] or [d̪].

Figure 67 shows a sibilant of very short duration followed by a stop. In this case, the stop has not assimilated a continuant characteristic. Ordinarily a non-initial sound of this type fails to assimilate a continuant characteristic only when that characteristic cannot be as-

similated, at the physiological level, because of the nature of the preceding sound. [n] and [l] provide the tongue tip position for a dental stop. Therefore a lowered tongue tip position, necessary for the continuant characteristic in [ð], cannot be assimilated from these sounds. After [s] the continuant characteristic may be assimilated, though, as illustrated, it is not always accomplished.

In figure 68, a midpoint between voicing and voicelessness is observed. This would probably be most often interpreted, by ear, as a voiceless segment because the dental stop is followed by aspiration. This is regularly true in the case of clearly voiceless dental stops (figure 69), but regularly not true in the situations where the dental stop is more strongly voiced (figure 67).



Figure 68. Informant VIII, sentence 15. Considerable vocal cord activity is present in this grooved and flat tongue combination but the aspiration spike together with the relatively weak nature of the base markings indicate that it is undoubtedly an auditory voiceless segment.

One further illustration shows that partial voicing, here held to be usually progressively assimilated, is not dependent upon the occurrence of a following voiced sound. Figure 70 represents the sequence corresponding to exceso hasta enco... The only sibilant, of the three

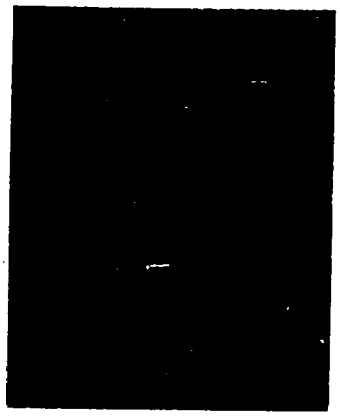


Figure 69. Informant XIII, sentence 15. This illustrates the same linguistic material found in figure 68 and shows that the dental may be unambiguously voiceless and yet have a lighter aspiration spike than is found in that figure.

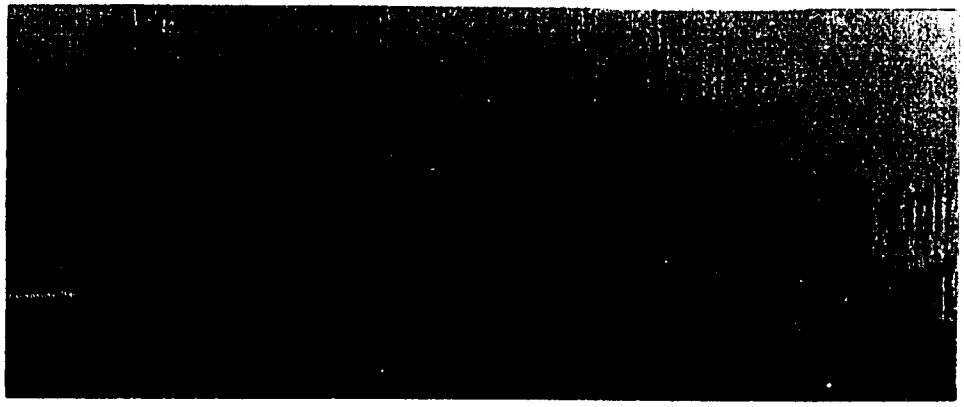


Figure 70. Informant XVI, Sentence 15. Three sibilants are illustrated. Two are followed by voiced sounds and one by a voiceless stop. Yet it happens that the only one which is strongly voiced throughout over half of its duration, and which contains substantial vocal cord activity throughout almost all of its duration, is the one followed by the voiceless stop. This could not be regressive assimilation.

shown here, which is strongly voiced for half its duration is the one preceding [t]. In addition the weaker striations continuing almost to the onset of [t] show the presence of regularly recurrent vocal cord

activity of the type which, when stronger, causes interpretations or sensations of voicing.

The data for istmo and mismo show that there is a voiceless element for istmo in 81 per cent of the occurrences. In the remaining 19 per cent the word is clearly voiced throughout. For mismo however, the data showed just the reverse situation. Only 19 per cent of all pronunciations contained a voiceless element and 81 per cent were clearly and strongly voiced throughout.

In accord with Navarro's treatment, voiced sibilants in both words will be illustrated first. Figure 71 illustrates istmo with no interval of voiceless vocal gesture. The sibilant is clearly voiced in this fashion is only 19 per cent of the samples. Figure 72 shows an occurrence of mismo in which sibilant friction is clearly present but voicing

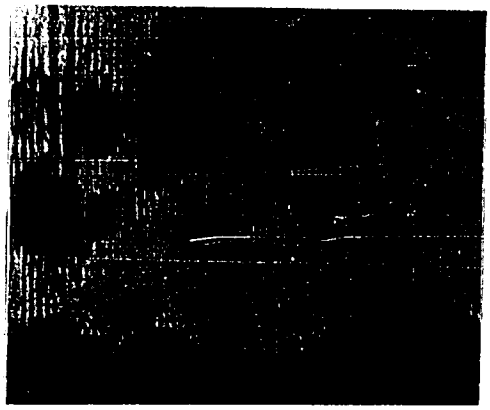


Figure 71. Informant IX, sentence 5. The voiced characteristic of the sibilant in this linguistic environment is representative of only 19 per cent of the data.

continues also as it does in 81 per cent of the samples. The same phenomenon appears in figure 73 except that the sibilant friction has a shorter duration and the characteristic rapid fall of the formants toward the position they occupy for the word final [o] is more pronounced.

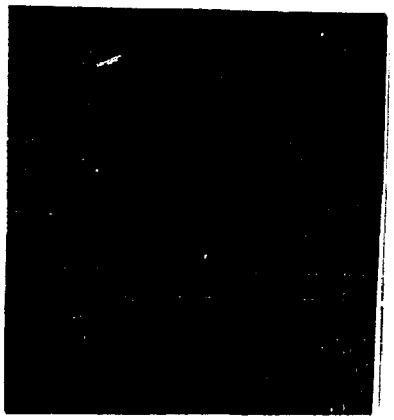


Figure 72. Informant VI, sentence 8. This sample illustrates the occurrence of the word mismo with clear sibilant noise. As in 81 per cent of the data, there is no voiceless segment.

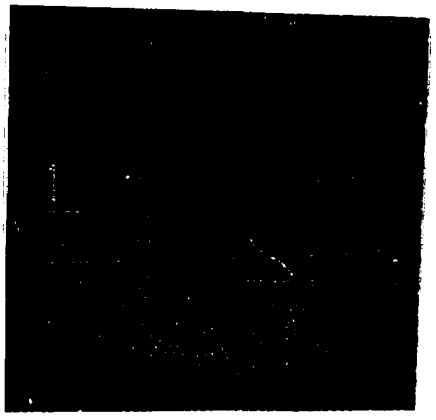


Figure 73. Informant XVIII, sentence 8. Sibilant noise occurs here, but it is much less prominent than in figure 72.

In 14 per cent of the samples this shift in formant position begins where evidence of sibilant friction is lacking as in figure 74. It may be observed that formants 2 and 3 are lower at the onset of the vowel than at the end and this may be an indication that the total sequence, between bilabial nasals, is not simply [mim]. The [z] which has a potential for occurring in this sequence may thus leave its mark, but there

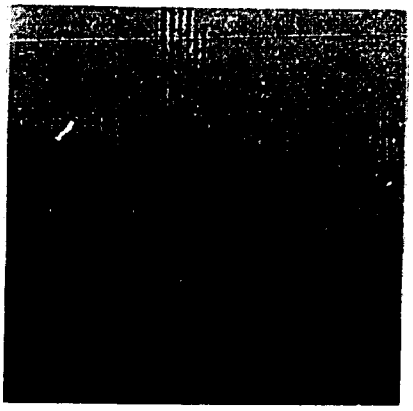


Figure 74. Informant VIII, sentence 8. The word mismo occurs without a voiceless segment in 81 per cent of the data. This figure is representative of 14 per cent, included in the previous figure, which fail to show clear evidence of sibilant noise.

is no unambiguous indication of a sibilant vocal gesture in these samples which constitute some of the 81 per cent considered to be relatively uninterrupted by a voiceless segment.

The 81 per cent occurrence of voiceless segments in istmo is illustrated by the next three figures. Figure 75 shows 5 cs. of voiceless sibilant followed by very sparse and light indications of voiceless friction indicative of a 4 cs. voiceless bilabial nasal segment. A similar voiceless nasal segment appears in 24 per cent of occurrences of istmo. It may become as prominent as the voiced nasal segment in figure 76. The onset of the voiceless nasal may also be accompanied by the vertical striation indicative of the percussion of lip closure illustrated in figure 77.

The data also showed 24 per cent occurrence of a voiceless dental stop in istmo as exemplified in figure 78. However one of the five in-

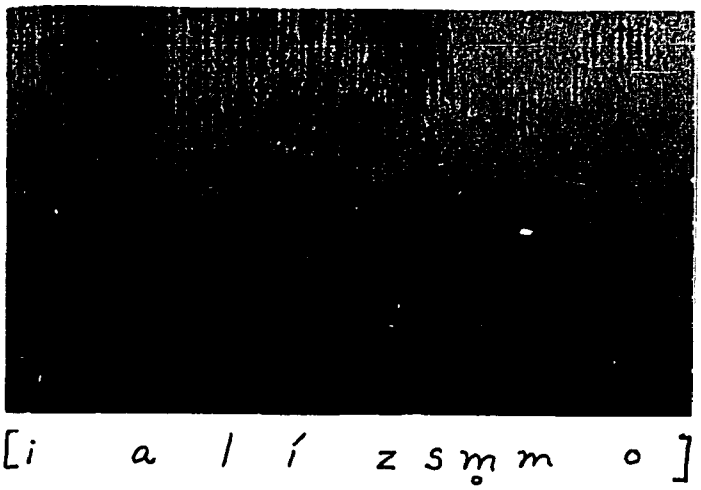


Figure 75. Informant V, sentence 5. A voiceless segment in the word istmo is characteristic of 81 per cent of the data. This spectrogram shows a sibilant partly voiced and partly voiceless. The first half of the following bilabial nasal is also voiceless and the last half is voiced.

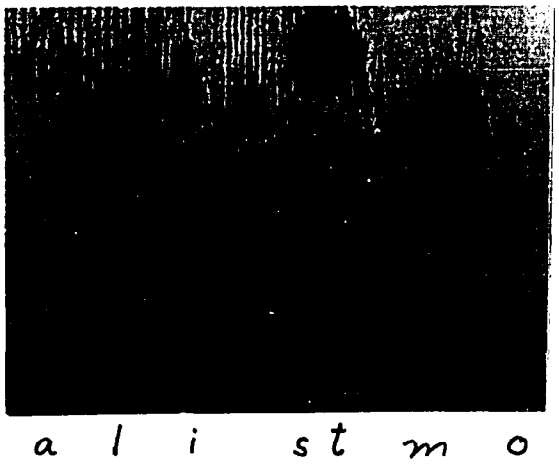


Figure 76. Informant XVII, sentence 5. Some friction in the relatively voiceless portion of the bilabial nasal is apparent in this spectrogram.

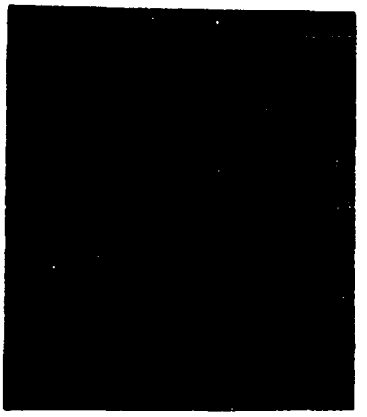


Figure 77. Informant XX, sentence 5. A relatively gradual onset of voicing is illustrated here after the beginning of bilabial closure in the area of the fuzzy spike. This demonstrates a possibility of separating segments which produce almost no identifiable fricative noise energy.

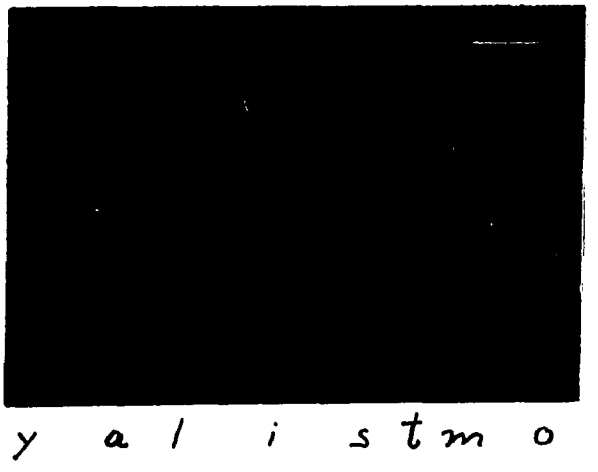


Figure 78. Informant VII, sentence 5. A voiceless dental stop, found in 24 per cent of the data, is clearly indicated in the sample.

formants producing a dental stop in <u>istmo</u>, consistently reversed the order of the sequence of sibilant and stop. His pronunciation is spectrographically illustrated in figure 79.

The relatively rare occurrence of voiceless segments as they occur in <u>mismo</u> is illustrated by figures 80 and 81. The former shows the spec-

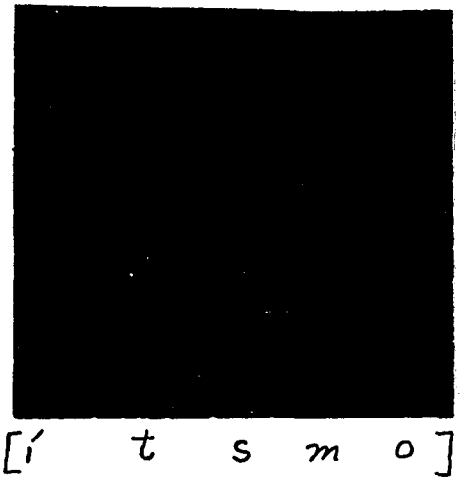


Figure 79. Informant XVIII, sentence 5. This spectrogram was produced from the records of the single informant who reversed the sequence of sibilant and dental stop.

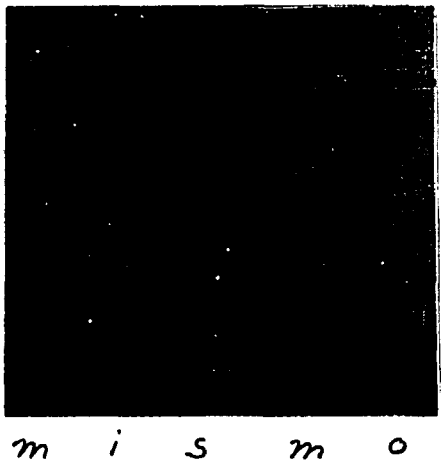


Figure 80. Informant XVI, sentence 8. The spectrographic response to a clear perceptual [s] in mismo is illustrated by this sample. A voiceless segment typifies 19 per cent of the data for this word.

trographic response to a clear physiological voiceless sibilant. The latter illustrates an alveolar stop segment in which the air movements due to vocal cord vibrations are rapidly diminishing to the point where, by contrast, they would be labeled voiceless. This voiceless characteristic is carried throughout the nasal portion of the sequence.

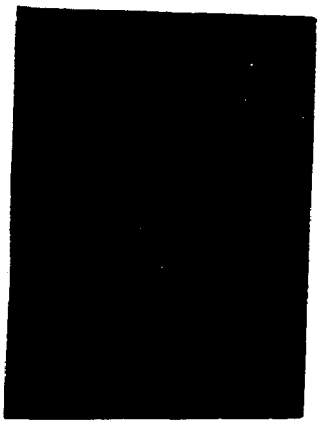


Figure 81. Informant X, sentence 8. Another type of voiceless segment in the word mismo is illustrated here. Sibilant friction is lacking as the tongue apparently failed to form a groove thereby producing instead a stop of the type heard in substandard English [ídn̥ət] for isn't it. This is followed by a nasal clearly lacking strong vocal cord activity throughout. The spike remains unexplained. It might correspond to release of bilabial closure followed by a two centisecond voiceless segment.

Navarro lists mismo with other words (in paragraph 107) of which he says the sibilants are voiced, except in slow or emphasized pronunciations. Of istmo, he says, "no se pronuncia la t en istmo [ɨ̟smo] o [ɨ̟zmo]."[1]

It is conceivable that the presence of microphones and the reading of the data caused some informants to substitute pronunciations that are otherwise characteristic of slow or emphasized speech and that therefore we have four informants who used a voiceless sibilant in mismo. However the same number of informants failed to use a voiceless sibilant in istmo. And there was no individual inconsistency here. The four informants who effected some voiceless articulation in mismo did the same in istmo.

---

[1]Navarro Tomás, Op. cit., 97.

Navarro lists istmo with alternant pronunciations, presumably equal in their acceptability, probability and distribution with [mɪ́zmo] and [mɪ́smo]. As stated, 81 per cent of the data on mismo shows wholly voiced pronunciations and an equal percentage shows a voiceless segment in istmo. In addition, the data show that 62 per cent of all informants pronounced mismo entirely voiced yet pronounced istmo with a voiceless segment. Interestingly enough, there is not one example of an informant who pronounced istmo fully voiced while using a voiceless segment in mismo. We should have expected a few situations of this type if the microphones and mimeographed copy were to be presumed responsible for artificial pronunciations.

The fact seems to be that there is divided usage in the pronunciations of mismo and istmo. Istmo occurs more often with a voiceless sibilant and, in this data, 24 per cent of the time with a [t]. Mismo occurs more often with a voiced sibilant.

This might seem to indicate that [s] and [z] are separate phonemes in Spanish but that conclusion would not be warranted. It is possible to argue that istmo is a relatively learned word and therefore more likely to occur with a pronunciation that is characteristic of parallel or similar words as pronounced in slow or stressed speech. But there is another argument. All five informants who pronounced a /t/ in istmo are included in the group of thirteen who had a voiceless sibilant in istmo but not in mismo. Probably the difference of environment is connected with the difference in sibilant voicing. We expect more frequent occurrence of a voiceless characteristic in a sibilant before a voiceless stop than in a sibilant between voiced sounds. Concerning the eight informants whose data so far remain unexplained, it is clear from other

instances that the potential occurrence of a phoneme can affect the pronunciation of segments in a word. This might be more acceptably stated by saying that there is divided usage of sibilant allophones in istmo when /t/ does not occur. However, the voiceless variety seems to be favored (by a ratio of three to one in the data) perhaps because istmo is so often heard with /t/ (24 per cent of the time in the data) and perhaps because the speakers are conscious of a potential, fully stopped [t] of which only the voiceless characteristic is here actually retained or assimilated. Single assimilated characteristics of potential sound segments have also been observed and reported elsewhere in this chapter (cf. figure 57). In this instance the omission of the tongue position element and the retention of the voiceless characteristic means that the latter must necessarily be assimilated into the preceding sibilant, the following nasal or both. The fact is clear enough that istmo, even when no full dental stop is actualized, occurs more often with a voiceless segment than does mismo.

Concerning the occurrence of voicing characteristics in sibilants, Navarro says simply that the voiced variety of the sibilant occurs only in syllable final position before a syllable initial voiced consonant and that elsewhere it is sporadic and abnormal.

The data used for this study show a good many voiceless intervocalic sibilants of the type illustrated in figure 82. Figure 83 shows an even more strongly voiced intervocalic sibilant at word boundaries. Navarro would call these sounds intervocalic syllable initial sibilants. Since they are voiced he would also call them sporadic and abnormal.

The data, recorded by well educated to highly educated individuals, show also examples of linguistic sequences, parallel to those found in



Figure 82. Informant VII, sentence 15. The second sibilant segment occurs simultaneously with considerable voicing. Since it is intervocalic it falls into Navarro's classification of sporadic and abnormal.

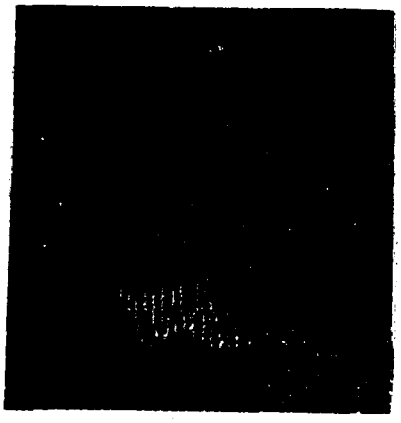


Figure 83. Informant XIX, sentence 15. The intervocalic sibilant here is more clearly voiced than that in figure 82. Voicing occurs quite commonly in sibilants in unstressed intervocalic word boundary environments like this one.

figures 82 and 83, in which voicing is very apparent but in which there is no evidence at all of a sibilant, much less of a voiceless sibilant segment. (Figure 84).



Figure 84. Informant XII, sentence 15. This represents the same linguistic material found in figure 83. It shows a type in which the formants shift position smoothly from [o] to [e]. Not only is it devoid of a voiceless segment but there is no evidence of sibilant friction either.

It will be useful, in reference to Navarro's statements above, to observe the statistics on the occurrence of voiced and voiceless portions in the sequence las islas y, as it occurs in context.

The first intervocalic sibilant occurs voiceless 94 per cent of the time. This is what might be expected and indeed seems to be roughly predicted by Navarro.

The syllable final sibilant preceding a voiced consonant, however, is voiceless 32 per cent of the time. Navarro equates the occurrence of voiceless sibilants in this position with the occurrence of slow emphasized speech.

The closely following intervocalic sibilant occurs voiceless 74 per cent of the time. Significantly, it occurs voiced 26 per cent of the time.

If the data is to be interpreted in the light of Navarro's statements it must be said that it shows 32 per cent slow and emphasized speech on the first syllable of islas but sporadic abnormal pronunciations 26 per cent of the time on the second syllable.

If we discount the data of the informants who gave either "sporadic abnormal" pronunciations or "slow, emphasized" pronunciations, we are left with only 42 per cent "normal" pronunciations of the total data on "las islas y...".

It should be observed that Navarro's generalizations do not adequately cover in a phonemically unprejudiced manner, the data found here for the sequence ...islas y... Over 26 per cent of these data show a characteristic that Navarro calls "sporadic abnormal." 25 per cent of the data for the sequence ...aliarnos a... also clearly showed continuous voicing throughout.

The common environment of the sibilant in the sequences named in the preceding paragraph is between unstressed vowels. Voicing under this condition is not "sporadic abnormal" in this body of material though, of course, it is not as common as is the inclusion of a voiceless segment.

Figure 85 illustrates a sequence with three voiceless sibilants in the varying environments mentioned: intervocalic stressed; before a voiced consonant; and intervocalic unstressed. Similar voicelessness is characteristic of sibilants in these positions to varying degrees. The first is representative of 94 per cent of the data as found in this environment. The second is representative of 32 per cent of the data as found here. The last represents 74 per cent of the data in this position.

Figure 86 shows, well before the onset of the [l], an articulation which was more like an ungrooved dental sound between [d̪] and [x̟] than



Figure 85. Informant VI, sentence 5. Three voiceless sibilants in varying linguistic environments, are shown as produced by one informant. We should rarely expect the second sibilant segment in this sequence to be voiceless in accord with Navarro's impressionistic findings.

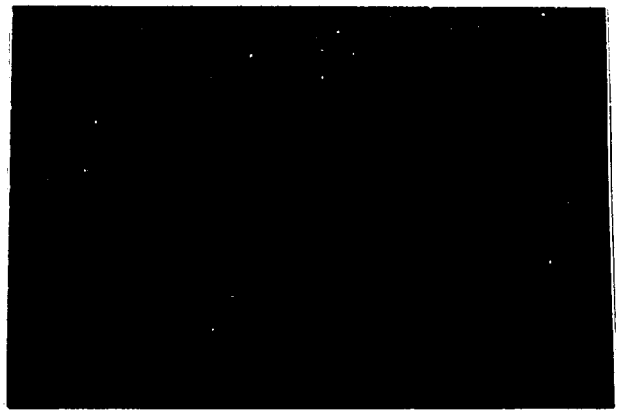


Figure 86. Informant IV, sentence 5. This spectrogram illustrates Navarro's general statement concerning the voiced characteristic of sibilants in syllable final position before a voiced consonant and the voiceless characteristic of those segments elsewhere. The sibilant spot, however, does not have much sibilant friction. Its acoustic qualities are related to those found in figure 81.

like a grooved fricative. The voicing here is characteristic of 68 per cent of the data.

Figure 87 shows an intervocalic unstressed sibilant with voicing characteristics even more prominent than those of the sibilant before [l]. This intervocalic voicing is representative of 26 per cent of the data.

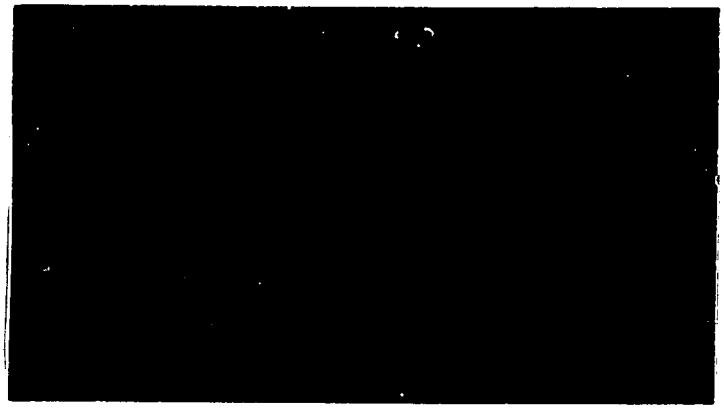


Figure 87. Informant XI, sentence 5. The intervocalic sibilant initial in a stressed syllable here remains voiceless as it does in 94 per cent of the data. It is seen, however, that in the unstressed intervocalic sibilant the indications of voicing are here even stronger than the expected indications of voicing in the sibilant preceding [l].

Navarro says that in the sequence sr in israelita, dos reales, etc., the pronunciation corresponding to s consists of a voiced sound.

Very consistently, the data show once more that approximately 25 per cent of pronunciations of sequences of this type will contain a very considerable voiceless segment. Figure 88 exemplifies the ten voiceless segments that were recorded in forty-two pronunciations corresponding to the written sequence sr in israelita and dos reales.

Navarro further says that these pronunciations consist of a fricative [ɹ] plus a trilled [r̥]. The data show clearly in twenty-six of forty-two samples the presence of high frequency friction of the type found regularly in [s]. However trills occur in only three of twenty-

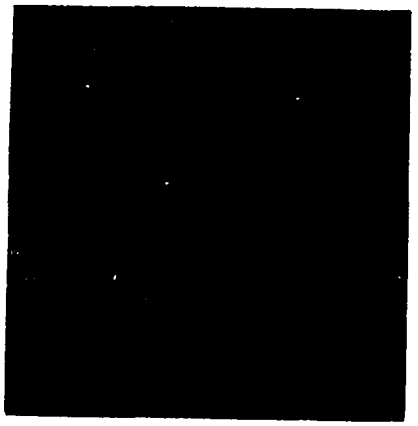


Figure 88. Informant II, sentence 11. The one-fourth of the data that show a voiceless segment in this context are represented here.

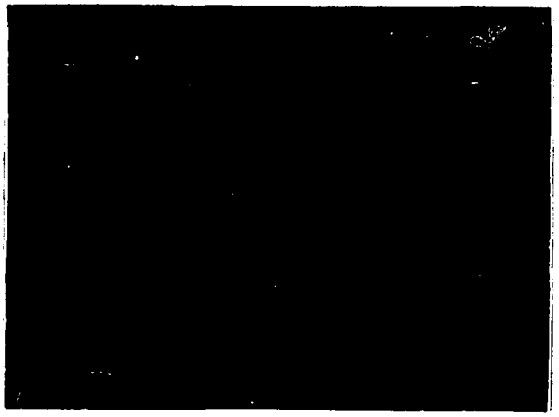


Figure 89. Informant XV, sentence 11. There is continuous voicing throughout the pronunciations corresponding to sr in 75 per cent of the data. As in figure 88, high frequency friction noise is clearly represented but, as in 62 per cent of the data, there is no evidence of a single flap, much less of a trill. Evidence points to progressive, or perhaps mutual, assimilation rather than regressive assimilation consisting of an increase in the number of flaps making up the trill referred to by Navarro.

one pronunciations of israelita and in eight of twenty-one pronunciations of dos reales. Single flaps occur in four pronunciations of israelita and in one pronunciation of dos reales. Twenty-six samples show no oc-

currence of flap or trill. Figure 89 is representative of those spectrograms that show neither a voiceless segment nor any trill. High frequency sibilant type friction was obviously intense.

Navarro says that sometimes pronunciations corresponding to this sequence occur without any sibilant. One form of this is illustrated by figure 90 in which there is shown a gradual shift from [i] to a fricative which strongly resembles a very brief and weakly articulated or lenis stop with a flat articulator. This is related to the type of assimilation causing the transition stop between palatalized [l] and [θ] in figure 67.



Figure 90. Informant I, sentence 7. This pronunciation corresponding to the sr sequence shows mutual assimilation in which the acoustic characteristics of the [r] seem dominant. Compare the dominance of the sibilant in figures 89 and 57. The end product in this instance, however, is a single vocal gesture more [r]-like than [s]- or [z]-like. For comparison with a quite typical [r], see figure 92.

A sample of a trill with no evidence of a preceding voiced sibilant except for the progressively assimilated high frequency friction is spectrographically illustrated in figure 91. It will be observed that, in general, the trills produce a series of relatively blank vertical striations in which even the base is very light. This is in contrast



Figure 91. Informant VIII, sentence 7. The type of trill that occurs in only 24 per cent of the data is illustrated here. A single occurrence of three flaps, constituting 2 per cent of the data, was encountered as doubtful substantiation for a theory of regressive assimilation of flaps.

to the voiced dental stops and fricatives. (See, for example, figure 10).

Navarro says that when the sibilant is lost the vibrations of the [r̃] are increased by way of compensation. But a number of spectrograms which show no clear evidence of a sibilant also lack evidence of a trill (cf. figure 80). In fact there are only eleven trills corresponding to the sr sequence and only one of these consists of three flaps. At best, therefore, the spectrographic data used here could be interpreted as showing only one case, or two per cent, of the data in which the vibrations of [r̃] might have been increased to compensate for the partial loss of a sibilant. However, even here it is very doubtful that the reason given is valid since [r̃] occurs elsewhere also with two or, less commonly but apparently in free variation, with three vibrations.

Navarro's expression, "a manera de compensación," is probably somewhat misleading here. He certainly does not believe that this is being

consciously or subconsciously done with a purpose on the part of the language or on the part of the individuals who speak the language. If he had said that there is mutual assimilation of sibilant and trill it would have been less striking but more adequate for the acoustic data. It seems slightly incongruous, in any event, that a grooved tongue movement will, by assimilation with [r̃], increase the number of vibrations of the tongue.

The whole matter should, of course, be compared with other observations of word initial [r̃] and word medial [r̃].

E. Assimilation and defects of pronunciation relative to /s/.

§109 ...La asimilación orgánica de la s final de sílaba respecto a cualquier consonante que no sea la [r̃] en las circunstancias citadas en el párrafo anterior, es un fenómeno considerado como un vulgarismo inaceptable en la lengua culta. Debe evitarse, por consiguiente, la compenetración entre la s y una consonante sonora siguiente, con pérdida de la sonoridad de dicha consonante, fenómeno frecuente en el habla vulgar de una gran parte del Sur de España: [laḅótɐ'], por las botas - [laz bótɐs]; [laxayínɐ'], por las gallinas - [laz galínɐs]; [mímmo], por mismo - [mízmo]; [áṇno] por asno - [ázno]; [íḷlɐ], por isla - [ízla].

Recházase igualmente como vulgarismo la pronunciación de la s final como una simple aspiración, y asimismo su eliminación total en determinadas circunstancias, hechos corrientes, según es sabido, en el lenguaje popular de varias regiones de España y América: [páhtɐ], por pasta - [páṣtɐ];...

Al corregir este uso se necesita cuidar de no caer en el defecto contrario, que consiste en pronunciar la s final prolongándola, o, como suele decirse, 'arrastrándola': lass passtass. Otro defecto que es preciso evitar es la transformación vulgar de la s en r en formas como [aɹθénso], por ascenso - [aṣθénso];...[1]

§106 ...Es defecto general entre los extranjeros hacer la s final española demasiado tensa y larga. Ante [f], lo mismo que ante [θ], §105, la s final es absorbida, en parte, por estos sonidos, resultando más débil y menos perceptible que ante otras consonantes: esfera - [e'férɐ], desfile - [de'fílə], etc. En la pronunciación de algunas regiones españolas, y principalmente en pronunciación andaluza, toda s final ante cualquier consonante, o ante pausa se reduce a una simple aspiración, generalmente sorda;

---

[1] Ibid., 110, 111.

> la pronunciación correcta española, aun en su forma menos culta, rechaza esta transformación...[1]

Mutual assimilation between sibilants and following consonants has been obvious throughout much of the preceding discussion in this chapter. At times characteristics commonly associated with Spanish sibilants have seemed to become dominant in these clusters and at times characteristics usually associated with the following consonants have predominated in the spectrographic representations of these sequences.

Usage among educated speakers of Spanish does not indicate that assimilation is restricted to one very specific type in each specific sequence. Assimilation is sometimes progressive and sometimes regressive. Sometimes it consists of assimilation or spreading of a voiced characteristic, and at other times of the spreading of a voiceless characteristic. And whether the sequence consists of pronunciations corresponding to s + r or s + d or s + t, the phenomenon of assimilation goes on unabated.

Navarro recognizes and condones the "compenetración" of pronunciations corresponding to s + r. In other combinations he says prescriptively that this is to be avoided. The reason he gives is that it is considered vulgar and unacceptable in the standard language.

It would be enlightening if Navarro had included here a statement concerning the basis for this reason. It does not descriptively cover the data used in this study which show, for example, progressive assimilation of voicelessness in approximately 25 per cent of samples regardless of the type of voiced consonant following a sibilant.

Figure 92 shows that in one short sequence we may find spectrographic representations of various such influences. There is a voiced sibilant

---

[1] Ibid., 106.



Figure 92. Informant VII, sentence 2. The difficulty in predicting the occurrence of non-phonemic characteristics in a vocal gesture is illustrated in this spectrogram. Three sibilants are shown as they occurred in the speech of one informant within a span of seven-tenths of a second. According to Navarro's description of normal, we should expect a fully voiced sequence corresponding to zg and a voiceless segment simultaneous with the second sibilant. The reverse is shown here. The voiceless characteristic of the third sibilant is "normal" according to Navarro's earlier description. In this data, however, it is representative of only three-fourths of the data. It is manifestly undesirable to call the occurrence of one woman for every three men "abnormal and sporadic" in a university where that is the normal situation. With reference to this data, the voiceless characteristic in the third sibilant is not representative of all normal pronunciations but is rather simply representative of the majority of pronunciations.

followed by a pronunciation, corresponding to g, which is largely voiceless and in which we see an aspirated or affricated release regularly found as part of the spectrographic representation of a voiceless velar stop. The sibilant itself has become voiced, but its voiceless characteristic has carried over into a following velar stop which is more commonly voiced. There is also an rs sequence in which the voiceless characteristic has completely disappeared as the sibilant has assimilated a voice characteristic. Finally there is intervocalic sibilant friction which has a voiceless characteristic.

In his samples of "compenetración" Navarro lists [laβóteˢ] and [mímmo]. He apparently considers these to be parallel samples. It appears however that the latter is not "compenetración" at all for it may be said that the /m/ retains all of its regular features and has none added while the preceding /s/ substitutes a new point and manner of articulation for its regular features but retains its voiceless quality.

The data show,too, that we may regularly expect, among speakers of standard Spanish, [mízmo], [mísmo], [mímmo] or [mímo]. Only the last would be a sample of "compenetración" parallel to [laβóteˢ].

Significantly, however, istmo does not occur as *[ímo]. Modern Spanish still allows for a potential [s] + [t] + [m] and the "compenetración," in the data used here at least, does not reach the point of having one gesture substitute for three.

Navarro says that elimination of /s/ or its pronunciation as simple aspiration is rejected. The data show many samples in which the evidence for /s/ varies from a blank segment to representations of extremely intense friction and strong vocal cord vibration with infinite gradations between. A goodly number of the pronunciations undoubtedly contain or correspond to "aspiration" and "elimination" but since these terms in Navarro are highly subjective, an accurate measurement of energy present in the various pronunciations would be to no avail for purposes of close comparison with Navarro's statements.

The same is true in regard to Navarro's statements concerning what he calls the defect of pronouncing prolonged sibilants. Length can be very accurately measured on the spectrograms and the representations for /s/ vary with a high frequency of occurrence for lengths between zero and 16 cs. An accurate compilation is useless here because of the indeterminate length of Navarro's "defectively prolonged" /s/.

This chapter began with a discussion related to Navarro's statements concerning the fronting of /s/ before [t] and [d]. It was seen that the collected data, and the treatment of that data, as done for this study, was not sufficient to present a structured picture which would show any relationship to Navarro's statements. Navarro also said that /s/ is fronted to a greater degree before [θ], and that it is, in part, absorbed by the friction of the latter sound. Since efforts toward finding acoustic relationships corresponding closely to Navarro's "fronting" did not prove productive, emphasis here was shifted to an analysis of /s/ plus the following seseo form, which was predominant in the data. As supplementary to Navarro, it was found in section A, of this chapter, that /s/ + a seseo form gives a physiological sibilant of no greater length than /s/ without a seseo form.

On the phonemic distribution level, additional strong support for the validity of the preceding conclusion was found in sections B and C. Here it was seen that, among seseo speakers, potential velar elements occur, in pronunciations corresponding to cc, to the same extent (speaking in terms of averages) that they occur in pronunciations corresponding to intervocalic x together with, significantly, intervocalic xc + e. The data which showed pronunciations corresponding to x followed by a consonant could not be included, for in these there was a very sharp drop in the average extent to which velar elements were found to occur before /s/. If the seseo form were serving as a separate consonant, then the pronunciations corresponding to x might reasonably have been expected to pattern more like pronunciations corresponding to x before other consonants.

Also of importance in section B was the material concerning voicing. Navarro had said that pronunciations corresponding to x showed voicing at the beginning if there were a velar element, but that the sibilant was voiceless in all cases. The acoustic data show the sibilant to be wholly voiceless, in terms of averages, only in those cases where the pronunciation corresponding to x began with a velar element. Where no velar element occurred the amount of progressively assimilated voicing of the sibilant was equal to the amount found in the velar elements.

In section C, it was found that pronunciations corresponding to cc closely paralleled those corresponding to intervocalic x and xc + e. Exception was taken to Navarro's explanation of the partially voiceless and partially voiced nature of pronunciations corresponding to cc. Navarro says that, in strong or emphatic pronunciation, a clear voiceless velar stop, [k], is the first element in pronunciations corresponding to cc. He implies further that, in ordinary, relaxed speech, this [k] gives way to a voiced fricative sound [g]. This, in turn, being weak and relaxed, comes under the influence of the following voiceless fricative and becomes partially voiceless itself. Since those pronunciations that wholly omit the velar element are found to contain progressively assimilated voicing to an extent equal to that of pronunciations which contain the velar element, it seems more likely that regressive assimilation is not responsible for the appearance of some pronunciations which include a velar element that is voiced initially but also has a voiceless portion. It seems rather to be a mere limitation of the amount of progressive voice assimilation.

The findings of section C are, in turn, borne out by the data of section D. Navarro says pronunciations corresponding to x before a voiced

consonant are themselves voiced, except that, in stressed speech, this voicing is partially or wholly arrested. He is saying, in effect, that, in unstressed speech, corresponding to z, we will get [ð] in Castillian, and [z] in seseo. In stressed speech, we will expect [θð] or [θ] in Castillian, and [sz] or [s] in seseo. The acoustic data show that assimilation, in these cases, proceeded in the opposite direction. The findings indicate, for example, that rather than [sz], the more appropriate transcription would be [zs]. Also, a voiceless characteristic is frequently assimilated progressively by the following "voiced" consonant.

Navarro says that the voiced variety of /s/ appears only in syllable-final position before voiced consonants. Spectrographic evidence shows the presence of strong vocal cord vibration in intervocalic position as well.

Navarro says that, before [r], the sibilant may be completely lost - the vibrations of the following [r] being increased by way of compensation. The limited acoustic evidence could not substantiate this. The tendencies seemed best described by stating that there was mutual assimilation, frequently leaving what seemed like one indivisible sound unit whose [r] vibrations were not increased but whose characteristics seemed rather to be a blending of those found in both /s/ and /r/ with accompanying loss of some features of both.

Section E quotes Navarro's judgments concerning variations in pronunciations of /s/. The acoustic data used as a base for this study did not include material that could test these statements.

CHAPTER V

CONCLUSIONS AND IMPLICATIONS

Our purpose was to compare Spanish physiological phonetics as represented by Navarro Tomás with a body of Spanish acoustic data. The body of Spanish acoustic data is equated with a series of spectrograms which are produced by the multiple interactions of a complex series of elements, devices or events: e.g., brain and neural activity, movements of the vocal mechanism, air, microphone, recording tapes, electrical circuits and charges, spark and sensitized paper to mention a few. Spanish physiological phonetics, for the purposes of this thesis, consists of the written observations of Navarro Tomás on the reactions of his eyes and ears to light and sound waves coming from the dynamics of the vocal mechanism directly or indirectly through x-ray pictures, palatograms, kymograph, etc. The bulk of Navarro's information on speech phenomena was tapped along the area of vocal mechanism movement whereas all of the spectrographic data are an index to a portion of the air movement present in speech communication.

No theory demands that we find a pure and simple one to one relationship between the acoustic and the physiological data. And we may reasonably expect less frequently to find this simple isomorphism between a native's reaction to the physiological activity of his language and a non-native's reaction to a spectrographic record of the acoustic phenomena in the speech of a different set of speakers of that language.

In view of this it is surprising that Navarro's statements correlate with the spectrographic observations to the very great extent seen in preceding chapters. If we permit ourselves to follow an unsafe tendency toward speaking in terms of physiological data, gathered im-

pressionistically, which can be corroborated or refuted by acoustic data gathered mechanically, then it may be said that Navarro's work has been largely substantiated. Conversely it can be said that Navarro's statements cover a substantial percentage of the total spectrographic data.

Yet, in spite of a large area of overlap, there is some disparity of observation and report, and it is the nature of this disparity which we feel compelled to examine since isomorphism or near isomorphism in reports is usually considered readily explainable if not self explanatory in most situations. Some possible causes for the disparity in the two reports are readily discernible. One of these is the difference in the selection of data; another is the demonstrable lack of simple isomorphism along the chain of events constituting the speech act; and a third consists of variant approaches to the relationship of phonetics and phonemics.

The first of these sources - difference in the selection of data - is quite obvious and need not be labored beyond a summary statement concerning the nature of these differences and their effect. Navarro undoubtedly observed the physiological aspects of the speech act as manifested in a large number of individuals with a wide variety of dialect histories. He does not describe these aspects of his procedures clearly but it is probable that a relatively large percentage of his informants were from Castille. It is also important to note that his observations were made more than a quarter of a century ago. The statements concerning acoustic data, on the other hand, were derived from spectrograms of the speech acts of a different set of informants differently chosen at a different time. This set is perhaps more equally

representative of the dialect variation which occurs throughout the Spanish speaking world and there is proportionately less representation of the geographical dialect of Castille in the acoustic data. Since Navarro also drew from a wide range of dialect backgrounds, however, his statements might still be expected to cover the comparable acoustic data. The first source, then, might be expected to give rise to variation primarily in statements concerning frequency and distribution of forms.

The second source of unlike statements rests on the fact that Navarro's reports are based on information siphoned from the area of the physiological dynamics of the vocal mechanism whereas the acoustic data were gathered from another area - air vibrations - along the chain of reactions in the speech act. The lack of simple isomorphism between the different areas is well attested. Peterson[1] has shown that a single pulse of the vocal chords - in the physiological area along the chain of speech events - produces not one but a series of damped oscillations in the acoustic area. Joos[2] describes clearly the function of reinforcers, resonators and filters, all of which disturb the proportion of sequent events as they appear first in one area and then in another. He further warns us "not to assume that all the acoustic details visible on a spectrogram are auditorily perceptible as being what we can see they are: as to timing and duration, the spectrograph is five times as sensitive as the ear and brain, and the auditory perception of explosions in particular is nothing like the spectrogram description."[3]

---

[1] G. E. Peterson in an unpublished report substantiates the theory with spectrographic evidence.

[2] Martin Joos, Op.cit., 32-35.

[3] Ibid., 92.

Joos also says, "one millimeter of tongue shift in the high front region might make twice as much color difference as one millimeter in the low-back region."[1] He goes on to give additional examples of acoustic changes which are disproportionate to articulatory changes.

All of this is, of course, not an attempt to nullify cause and effect theories since it is readily conceded that, with enough information, laws of correspondences could be formulated. But the important point is that the relationships commonly do not demonstrate what has here been termed simple isomorphism and therefore the resultant descriptive statements of articulatory and acoustic phoneticians must be expected to reflect that absence of simple isomorphism by showing, instead, the real shift of proportion among sequent events in the speech act.

The third source of difference in physiological and acoustic descriptive statements found in this dissertation is a very fertile source indeed, and consideration of it will permeate the remainder of this chapter. It consists basically of one's assumptions concerning the structure of language.

Mindful of his fundamental assumption[2] that some utterances are alike in form and meaning, i.e., that portions of linguistic behavior do recur, Bloomfield says:

> "....The phonetician's equipment is personal and accidental; he hears those acoustic features which are discriminated in the languages he has observed. Even his most 'exact' record is bound to ignore innumerable non-distinctive features of sound; the ones that appear in it are selected by accidental and personal factors. There is no objection to a linguist's describing all the acoustic features he can hear, provided he does

---

[1] Ibid., 117.

[2] Leonard Bloomfield, Language. (New York: Henry Holt and Co., 1933), 78.

not confuse these with the phonemic features. He should remember that his hearing of non-distinctive features depends upon the accident of his personal equipment, and that his most elaborate account cannot remotely approach the value of a mechanical record.

Only two kinds of linguistic records are scientifically relevant. One is a mechanical record of the gross acoustic features, such as is produced in the phonetics laboratory. The other is a record in terms of phonemes, ignoring all features that are not distinctive in the language. Until our knowledge of acoustics has progressed far beyond its present state, only the latter kind of record can be used for any study that takes into consideration the meaning of what is spoken."[1]

Navarro does not make it clear whether his purpose is to approximate a good mechanical recording of all sound features or to make a description based on phonemic analysis. One is almost led to believe that he is aiming at a complete recording of all Spanish sound variations, at least in their physiological phase, when he reads, "El alfabeto fonético tiene por objeto representar lo más exactamente posible, por medio de la escritura, los sonidos del lenguaje."[2] This impression might be strengthened by the observation that Navarro lists and describes, for example, five high front vocoids, four laterals and five sibilant sounds. But it is difficult to resolve this with the impressions gained from, for example, his statements concerning the pronunciation of the fricative [ð]: "Diferentes circunstancias hacen que la articulación de la [ð] vacile entre la posición dentointerdental claramente fricativa y la posición dental más o menos oclusiva ....pero tanto en uno como en otro sentido lo más frecuente no son en realidad las articulaciones extremas, sino diferentes matices intermedios difíciles de precisar y describir;"[3] In reference to the d in regular

---

[1] Ibid., 84,85.

[2] Navarro Tomás, Op.cit., 31.

[3] Ibid., 99.

past participles he speaks of the "uso predominante de una u otra variante en la pronunciación de este sonido."[1] And in his discussion of s, the matter is confused. Here he describes the physiological properties of one type of voiceless sibilant which occurs before [t]. He then describes another type, different in manner and point of articulation, occurring only before [θ] and says, without giving a reason, "la representamos también en este caso con el mismo signo [ş]."[2]

It is probable that Navarro intended to describe and symbolize all differences that he could perceive and to which he reacted subjectively as "considerable differences." It is further probable that he intended to describe what he could of the other perceptions to which he reacted as "relatively slight differences." This hypothesized picture places Navarro in somewhat the same position in which Henry Sweet found himself when, for practical reasons he felt constrained to limit his Narrow Romic and to create Broad Romic while continuing to believe that the former was somehow more scientific.

Strong justification for Navarro may arise from the fact that he sets forth, as one of the aims in his text, the idea of a brief and simple phonetic description "tendiendo, sobre todo, a facilitar la enseñanza práctica de nuestra lengua en este aspecto poco conocido de su naturaleza."[3]

However, this justification, too, is open to question for he places no restriction on the background of those for whom the "practical teach-

---

[1] Ibid., 104.

[2] Ibid., 104.

[3] Ibid., 5

ing" may be designed. In this situation, with no more specific goal, each sound distinction which he can discriminate should be reported and recorded with the same degree of emphasis as every other.

Navarro's approach is, of course, not primarily phonetic but phonemic. If it were strictly phonetic he might describe one range of positions of the vocal organs as, for example, that used in the production of palatalized laterals, and then proceed to show what, in Spanish, was pronounced in this way. Instead he takes a "sound" or spelling and follows through with descriptions of corresponding pronunciations. These pronunciations of one "sound" may include personal and dialect variations which may range, for example, from voiced palatalized lateral, through voiced or voiceless retroflex sibilant with heavy friction, to a voiced sibilant affricate, or to a high front close vowel-like sound.[1] The fact that Navarro's attack is "phoneme-oriented" seems to be partly due to a circumstantial accident: the nearly phonemic spelling of Spanish. It is also largely due to the fact that Navarro reacts to Spanish as one who understands it in its function as a communicative agent.

It is largely insofar as Navarro fails to take into account more fully than he does the implications of the structured nature of language that this lack of distinction between phonetic and phonemic serves as the basis of the third source of non-corroborative or "non-sympathetic" statements found in the description of the acoustic data. To this shortcoming may be attributed Navarro's failure to take into account the fact that his own observation will be shaped by his history. In other words, it is perfectly legitimate, in one sense, to state that

---

[1] Ibid., 133-136.

one's perception will not deceive him, since an accurate descriptive statement of a perception remains forever good as a description of that perception; but it is deceptive to believe that the description will necessarily hold for the perceptions of other observers. Navarro's perception, like that of any other observer, may be influenced by psychological set; by the differential frequencies of occurrence, or probability of occurrence, of the speech activities in question; and by perceptual defense and vigilance.

The way in which these forces may have operated and the reasons supporting a belief that they did operate here can be shown by way of examples. The reported perceptions of the occurrence of length, voicing and phonetic variation, especially border phenomena, are chosen as the vehicles by which to discuss these forces as they operate on the person who reports his impressions on phonemic and non-phonemic linguistic behavior.

For the purposes of this argument, an event is assumed to occur if it is perceived. This holds for a span of time in which a composite of various sound features occurs and also for specific sound features occurring simultaneously with others.

Thus Navarro observes the recurrence of some words in Spanish from which he can usually abstract sibilant portions which may vary from long and heavy sibilant sounds to very light aspiration. However, at times, he also observes the absence of these physical characteristics in the relative spot where he usually perceives a sibilant or spirant. A description of the spectrograms echoes these statements.

On the other hand Navarro also observes length relative to the voiceless affricate and stops in Spanish. He says it is the same for

all. Of course, we partially discount this categorization at once because we know he means that there is no regularly recurrent or patterned difference in length between the stops on the one hand and the affricate on the other. We should not be inclined to dispute his statement if we find merely that the tongue retains contact with the alveolar ridge longer in one particular phonetic environment, stressed syllables for example, than in another; nor if we discover that one person's /č/ is longer than another's /t/. But confronted with the facts and statistics in the section devoted to a discussion of "length" in chapter III, we are faced with a choice of proclaiming that Navarro's impressions were false or of explaining what must be understood by Navarro's use of the term "same."

The argument of Chapter III was so definitive that any escape on the grounds of difference of procedure alone was not considered in the preceding paragraph. It must be understood now, however, that the three previously given sources of variant statements, which were separated for purposes of analysis and discussion in this chapter, do not operate in a vacuum but are interacting forces, the borders of which are not clearly definable. Thus even the differences of the procedures used in Navarro's work and in the acoustic portion of this study could and probably did subsume different outlooks on the nature of language.

Similarly, lack of isomorphism along the chain of events in the speech act is not sufficient by itself to explain fully the contradictory descriptions implied by the data presented in the earlier chapter, yet in a sense it lies at the very heart of the best explanation that can be given at present. To state it as a hypothetical equation: that which may consist of x number of units in the brain may result in y

activities of the vocal mechanism which in turn may stimulate z activities in the air.

Finally, variant assumptions on the nature of language underlie different approaches to the phonetics and the phonemics of a language. The different attitudes toward the structure of Spanish and the consequent difference in procedure, especially insofar as it relates to description of the perceptions of the informant activities, must be held as the biggest single reason why Navarro says "same" while the acoustic report shows "different."

In this study the linguistic data on length are considered as "same" or "different" on two distinct levels. On the phonemic level they are all "same." That is to say, in essence, that I cannot alter "árbol," for example, from the meaning "tree" by maneuvering the length of any of the vowels or consonants. On the phonetic level, the data are considered "same," first, insofar as measurements show absolutely same number values, and second, insofar as lengths vary without pattern. In reference to the affricate and voiceless stops, there is no absolute sameness of length. There is a definite pattern of length variation. The velar stop is the shortest. Bilabial and dental stops and the affricate are, in that order, progressively longer. There is left no recourse but to say that, phonetically, the acoustic data show length to vary in a pattern and that probably physiological data, if gathered with an accurate timing device, would show similar results. Phonemically, however, length is invariably "same": the conspicuous lack of isomorphism in this instance occurs somewhere between the brain area of the speech act and the other areas.

Navarro, in his statement concerning sameness of length, was probably reacting phonemically, as he does at least partially whenever he assumes with Bloomfield that certain utterances or portions of them may recur, i.e., be considered "same." The task of the student consists of discriminating Navarro's phonemic statements from those which are phonetic and those which lie between.

Navarro likewise probably reacts subconsciously to voicing, or the lack of it, on a largely phonemic basis. As before, he is attempting to react to phonetic data without regard to meaning. He succeeds to an extent but, as must be expected, there is evidence that indicates a compromise of his apparent purpose with his phonemic set.

The sibilant /s/ is one meaningful or phonemic unit with various phonetic manifestations in the structure of Spanish. It may, for example, occur voiced or voiceless. Navarro's reaction is, in substance, that it is voiceless sibilant noise in careful or stressed speech and that it retains this voiceless quality between vowels even in rapid conversation, while voicing occurs simultaneously if it is immediately followed by a voiced consonant. He implies that anything else is a sporadic anomaly. The acoustic data, assumed to represent rapid speech in standard Spanish, show frequent occurrence of both voiced and voiceless sibilants in the two environments. The frequency of each is far too great to allow dismissal on the grounds that these are sporadic anomalies in the phonetic make-up of Spanish.

It is desirable to explain why, if voicing is not phonemic here, Navarro still reacts so sharply to what he considers to be the presence or absence of voicing and why this reaction does not fully correspond to that which is found in the acoustic data. The full explanation will

undoubtedly have to await further developments in the measurement of human behavior; but there seem to be some indications now of a direction this explanation may assume.

Of primary importance is the fact that voicing is a phonemic characteristic of Spanish speech behavior in the stops, though not in contrast with voiceless forms of /s/. A corollary to this is the fact that even elsewhere the presence or absence of voicing never occurs in more than a single contrast. Phonemically or meaningfully, there exists in Spanish only definitive vocal cord vibration contrasted with light vocal cord activity varying to apparent lack of movement. There are not three or four meaningful gradations of intensity of activity as there might be, and as there are indeed in reference to frequency and duration of vocal cord vibration in some languages.

In light of this phonemic frame which certainly facilitates (if indeed it does not impose) interpretations of "either-or" in relation to presence and absence of voicing in any given situation, why does Navarro apparently react to a significant number of occurrences of acoustic "voicing" as though no voicing had occurred?

As before, differences in procedure must be considered, though the form of the question above may betray the assumption that these differences are not uppermost in this instance. Differences in data are also important, but it seems highly improbable that, of the many informants that Navarro undoubtedly picked for observation, he could have failed, even in a different era and location, to pick any who would have produced vocal cord phenomena similar in distribution to that produced by such a significant number in such significant quantity for the acoustic data. Similarly, Navarro's procedure of reporting phenomena is not

likely to be at fault here, for, though he does display what might be termed preferential reporting, he does not display selective or exclusive reporting: he frequently labels one form as "correct," "best," "cultured" or "non-pretentious" while describing another as "dialectal," "vulgar" or "to be rejected," but if they occur in some pattern - phonetic, social or other - he does not say that they are sporadic or anomalous. The lack of simple isomorphic activity between physiological and acoustic areas of the speech chain is a more likely cause for the differences in reports.

However, Navarro speaks of voicing rather than strength of vocal cord vibration; and it may be assumed that he did not visually see the relative strength nor even the presence and absence of vibration but rather relied on his auditory perceptive mechanisms to warn him of voicing. It is not likely that strongly marked pulses in the acoustic area are in any case indicative of weaker or inherently less perceptible vocal cord vibration, in the physiological area, than are relatively more weakly marked pulses. Any gross lack of isomorphism must then be found between the acoustic report and Navarro's report rather than between the acoustic report and the physiological activity.

Thus, the inquiry into outstanding causes for the differences of acoustic and physiological report is shifted, in this instance, from (1) differences in procedure and (2) lack of isomorphism between physiological and acoustic areas to (3) differences in the interpretation of phonetic data and phoneme substitution.

In reporting the acoustic data it was recognized that phonetic variation is absolute while phonemic variation is highly relative and therefore not determinable on an absolute scale. Navarro did not make



a sharp distinction between his phonetic and phonemic reports. Stated slightly differently, and perhaps more accurately, it may be said that he did not maintain a clear distinction between his reports on linguistic activity as it occurred in different areas along the chain of events in the speech act. He assumed that the isomorphism between the movements of the vocal mechanism, the perceptions of the ear and the interpretations of the brain was far simpler than phonemic theory and the acoustic data warrant.

Experience has shown, for example, that two native speakers of a language may agree that several pronunciations of a word are exactly the same, though a novice phonemicist, in trying to analyze the language, may complain bitterly that his informant is constantly altering his pronunciations. The novice may be listening to voiced and voiceless variations, whereas the language he is attempting to describe may operate phonemically with a fortis-lenis contrast. For the native, the "same" response is elicited by recurrences of a lenis characteristic while a "different" response is elicited by a lenis-fortis contrast. For the novice, "same" may be elicited by recurrences of a voicing characteristic and "different" is then elicited by a phonetic, non-meaningful occurrence of a voiceless segment. In such a situation different acoustic-perceptual relationships exist side by side, with the result that one set of acoustic properties may elicit variant perceptions: perceptions, or responses may be related to different portions of the total acoustic signal.

In most cases this subconscious psychological "set" to perceive certain portions of a total linguistic activity is determined by the structure of the observer's learned behavior which in this case is his



native language. But the learned behavior may also include other linguistic behavior or experience of the observer. Though acoustic phenomena may remain constant, a bilingual person, a trained phonetician, or a literate observer may have different perceptions from those of the observer who simply speaks one language; i.e., a single set of acoustic data may fire different neural combinations, or neural combinations which are differently perceived, in different observers. In a more generalized statement, it may further be postulated that these variant perceptions correlate with the intensity of perceived occurrence in the observer's history. This theory receives strong support from the studies on the correlation of frequency of occurrence and ease of perception in the field of psychology.[1] Solomon, Howes and Postman have shown that ease of word perception in certain tests show a direct correlation with the position of these words in the Thorndike-Lorge word frequency count. Related to these studies is Tanner's report[2] on experiments showing that the perceptual threshold itself is variable, dependent upon the observer's feeling concerning the probability of occurrence of a given phenomenon.

---

[1]R. L. Solomon and D. H. Howes, "Word frequency, personal values and visual duration thresholds," Psychological Review, LVIII (1951), 256-270.

R. L. Solomon and L. Postman, "Frequency of usage as a determinant of recognition thresholds for words," Journal of Experimental Psychology, XLIII (1952), 195-201.

D. H. Howes and R. L. Solomon, "Visual duration threshold as a function of word probability," Journal of Experimental Psychology, XLI (1951), 401-410.

R. S. Lazarus and R. A. McCleary, "Autonomic discrimination without awareness: a study in subception," Psychological Review, LVIII (1951), 113-123.

[2]W. P. Tanner, Jr., Psychophysical Application of the Theory of Signal Detectability, (Ann Arbor: Engineering Research Institute, 1954).

If the findings of these studies are accepted, some interesting explanations suggest themselves in reference to Navarro's report and this thesis. Navarro reported observation of initial voicing in pronunciations corresponding to written x, for example, only when that pronunciation contained a velar element with a duration equal to or exceeding the extent of the voicing. The acoustic data showed that progressive assimilation of voicing from the preceding vowel into the speech segment corresponding to written x remained statistically constant whether that segment contained a velar element or not. On the basis of the phonemic structure of Spanish and on the basis of the principles of perception mentioned above it becomes evident that Navarro, though he is a highly trained and an expert phonetician, will necessarily be less likely to observe voicing in sibilants, for Spanish offers no incentive to perceive contrast here, than he will be to observe an equal amount of voicing in velars.

There are perhaps many other factors relevant to perception and the variability of thresholds for perception. Among these, one example whose relevance to the perception of acoustic phenomena is more tenuous but still plausible, is that of perceptual defense and vigilance.[1] Dulany[2] shows that potential punishment of responses to perceptions may, wholly subconsciously, cause those perceptions themselves to be less likely of occurrence. A great deal of training, and of study of the observations of other scholars, as well as the publication of one's own findings, usually constitute a vested interest in a given inter-

---

[1] Cf. L. Postman, W. C. Bronson and G. L. Gropper, "Is there a mechanism of perceptual defense?" Journal of abnormal Social Psychology, XLVIII (1953), 215-225.

[2] D. Dulany, Avoidance Learning in Perceptual Defense and Vigilance, Thesis (1955), 39.

pretation of the scheme of things. If, in certain cases and to a limited extent, loss of vested interests may be equated with punishment, it is possible that perceptual defense and vigilance are relevant to the perception of acoustic speech phenomena.

Credibility is given this theory as an outgrowth of the common tendency of people to "jump to conclusions." In phonemic matters these conclusions can be and are more readily revised by significant contrasts in the data but in phonetic matters such revision is much less strongly called for and each conclusion is consequently much more likely to develop into a vested interest.

Something of this nature is a plausible partial explanation for Navarro's failure to observe voicing in unstressed intervocalic /s/ when this phenomenon is equally frequent with that of the readily observed variation of voiced /s/ to a voiceless variety before a voiced consonant in stressed speech.[1] It may also serve as a supplementary explanation for Navarro's failure to observe, or to report, the uniformity of the voicing characteristic, in the tested pronunciations corresponding to written x, regardless of the presence or absence of a velar element in the individual pronunciation. His failure to mention even as "anomolous and sporádic" the common occurrence of voicing throughout pronunciations corresponding to intervocalic p, t and k, or the less common occurrence of a voiceless segment in the span corresponding to b, d and g, may be partially, if not largely, due to this perceptual factor. Early accretions in one's phonetic training plus the implementations of spelling and "slow pronunciations" are not easily jettisoned.

---

[1] Navarro Tomás, Op. cit., 108.

In connection with the same problem, it is common experience to learn a new fact or word which has not seemed to appear before but comes to the attention often thereafter. Once consciousness has been directed, perception occurs more readily in everyday experience.

Matters of length and voicing constituted a considerable portion of the subject material for variant statements. In discussing these, assimilation was mentioned since it is not strictly separate from them. If, in turn, we now look with primary concern into the physiological causes of, the acoustic manifestations of, and the perceptual variations of phoneme border phenomena, which is a somewhat broader term than "assimilation" as ordinarily used, we shall have completed a survey of the bulk of the subject material on which variant statements were made.

Border phenomena, or assimilation in this larger sense, means the projection of any activity identified with the sound unit in question into a time area that is not coterminous with other features of the sound. Assimilation defined in this way includes regressive and progressive assimilation, epenthesis and metathesis.

Non-coterminousness of features of sound which give rise to regressive and progressive assimilation has been thoroughly illustrated in chapters III and IV.

Figure 67, showing a stoplike segment in escollo, serves to illustrate a type of physiological assimilation - the back of the tongue is raised for [o] before the sides and front are released from the [l] position. This is commonly called epenthesis because it is felt that the full quality of both the lateral and the vocoid are present and that there is another sound segment in addition. The physiological process, however, is not essentially different from that which occurs in any other assimilation.

Matathesis consists of the same process carried to the point where observer reaction, due to sufficient shifting of enough features, is that contiguous or non-contiguous sounds have reversed their regular order of appearance. In cases where metathesis is complete rather than partial, and where the metathesized sounds are not contiguous, the theory of activity projection would of necessity be applied on the neural or brain level if it were to be applied at all.

If we look at speech as a sequence of phonemes, then assimilation, including its extended forms, epenthesis and metathesis, is a natural process which occurs as transition from phoneme to phoneme. If we overlook the classification of speech behavior into groups of "same" units - "same" on the basis of purpose of the speaker and reaction of the listener - and assume only the phonetic reality, then speech may only be represented as a complex arrangement of physiological (or other) activity waves which intersect indiscriminately.

If we assume that people react phonemically to speech, we should expect to find, recorded by the spectrograph, great quantities of acoustic phenomena which develop from the indiscriminate intersecting of physiological activity waves, and which nevertheless go completely unobserved by native speakers and phoneticians. Bloomfield, in the statements quoted earlier, directs attention sharply to the fact that this is precisely what is found.

It has further been borne out that transition, or non-simultaneity of activity-wave intersection, is a fruitful area for the development of non-significant variation. This non-significant variation, too, is mostly patterned. But, patterned or non-patterned, the divided usage that occurs in this area of speech, forces one to come to grips with,

and to recognize the idea that change is natural: change occurs constantly and it seems at present to be impossible to alter the situation.

A realization of the processes at work in language would seem to indicate the desirability of attitudes of acceptance and study in respect to divided usage of the phonetic content of phonemes. This realization also engenders an awareness of the relevance of some frequently undervalued factors which are of importance in assessing the contributions of a work such as that of Navarro Tomás. Among these are the observer's background and the area, along the sequent chain of events, with which his study is concerned.

APPENDIX

The following list includes all informants who supplied linguistic material for the thesis. The roman numerals correspond to those used throughout the report in reference to the informants. The information following each name is pertinent to the linguistic history of that informant. This information is outlined under the letters (a) through (h) for the sake of brevity and for easy reference. In the outline, information listed under (a) corresponds to the date of recording; all material was recorded in the time between 1:40 and 3:15 p.m. on the dates indicated. Information under (b) corresponds to the informant's age; (c), his place of birth; (d), area of travel; (e), place of residence; (f), number of years spent in school; (g), occupation; (h), geographical history of his parents.

I. Wilson Alva (a) August 12, 1953 (b) twenty-seven (c) Contumaza, Peru (d) Peru (e) Ica, Peru (f) Fifteen years (g) Dentist (h) Contumaza, Peru - one year; Trujillo, Peru - nineteen years; Lima, Peru - five years; Ica, Peru - two years.

II. Alejandro Chapa (a) August 12, 1953 (b) twenty-four (c) Monterrey, Mexico (d) Spain - three months; England, Portugal, Morocco, Cuba, Sto. Domingo, Italy, France, Switzerland - three months; United States - one month (e) Monterrey, Mexico (f) Fifteen years (g) Lawyer (h) Monterrey, Mexico.

III. Hugo García (a) August 20, 1953 (b) twenty-nine (c) Los Andes, Chile (d) Argentina, Uruguay, United States - one month (e) Santiago, Chile (f) Eighteen years (g) Medical Doctor (h) Los Andes, Chile.

IV. Luis Saenz (a) August 20, 1953 (b) Twenty-four (c) Montevideo, Uruguay (d) Europe; Spain - one month (e) Montevideo (f) Sixteen years (g) Lawyer (h) Montevideo.

V. Hernando Echeverri (a) August 20, 1953 (b) Thirty-two (c) Medellín, Colombia (d) Panama; Cuba (e) Medellín, Colombia (f) Nineteen years (g) Surgeon (h) Rionegro, Colombia.

VI. Cornelius Tejada (a) August 20, 1953 (b) Forty-one (c) Salcedo, Dom. Rep. (d) Cuba; Mexico (e) Lajar, Puerto Rico - eight years (f) Eighteen years (g) Medical Doctor (h) Salcedo, Dom. Rep.

VII. José Ricardo Martinez, Jr. (a) August 20, 1953 (b) Seventeen (c) San Miguel, El Salvador (d) Guatemala - one month; Mexico - one month; Brazil - one month (e) San Salvador, El Salvador (f) Eleven years (g) Medical Student (h) San Miguel - three years; San Salvador - sixteen years.

VIII. José Pío Romero Bosque (a) August 20, 1953 (b) Twenty (c) San Salvador, El Salvador (d) El Salvador (e) San Salvador (f) Eleven years (g) Student of Diplomacy (h) San Salvador, twenty years.

IX. Jesús Leal (a) August 21, 1953 (b) Eighteen (c) Chihuahua, Mexico (d) Mexico (e) Chihuahua (f) Twelve (g) Student (h) Chihuahua.

X. José Carrillo (a) August 21, 1953 (b) Twenty-eight (c) Trujillo, Venezuela (d) Venezuela (e) Caracas, Venezuela - twenty years (f) Sixteen years (g) Civil Engineer (h) Trujillo.

XI. Susana Icaza (a) August 21, 1953 (b) Thirty (c) Las Tablas, Panama - one year (d) Argentina - four years; Guatemala - two years (e) Panama (f) Sixteen (g) Nutritionist (h) Panama City.

XII. Carlos Espinosa (a) August 21, 1953 (b) Twenty-nine (c) Cereté, Colombia (d) Venezuela - two years (e) Cartagena, Colombia (f) Fifteen years (g) Physician (h) Cereté.

XIII. Guillermo Lozano (a) August 21, 1953 (b) Twenty (c) Bogotá, Colombia (d) none (e) Bogotá (f) Fourteen (e) Engineering Student (h) Bogotá.

XIV. Lily Luna Andrade (a) August 21, 1953 (b) Thirty-one (c) Chimaltenango, Guatemala (d) United States (e) Ann Arbor (f) Sixteen years (g) Teacher, Student of Economics, Office (h) Guatemala.

XV. Lucía Barrera (a) August 21, 1953 (b) Twenty-five (c) Bogotá, Colombia (d) Colombia (e) United States (f) Ten years (g) Office (h) Colombia.

XVI. Sixto Duré (a) August 27, 1953 (b) Thirty-two (c) Asunción, Paraguay (d) Paraguay (e) Asunción (f) Ten years (g) Army Captain (h) Asunción.

XVII. Emiliano Gallo-Ruiz (a) August 27, 1953 (b) Forty-four (c) Burgos, Spain (d) Europe, Cuba, United States (e) Ann Arbor (f) Over twenty years (g) Teacher of Spanish (h) Burgos.

XVIII. Vicente Rojo (a) August 27, 1953 (b) Thirty-two (c) Madrid, Spain (d) Spain; France; Brazil; Uruguay; Argentina; Bolivia; Peru; United States (e) Cochabamba, Bolivia (f) Over twenty-five (g) Medical Doctor (h) Both parents are Spaniards: father from Valencia; mother from Madrid.

XIX. Eduardo Pérez (a) August 27, 1953 (b) Nineteen (c) Quito, Ecuador (d) United States (e) Quito (f) Twelve (g) Student (h) Quito.

XX. Humberto B. Capriata (a) August 27, 1953 (b) Twenty-one (c) Argentina (d) Peru (e) Lima, Peru (f) Twelve years (g) Student (h) Argentina.

XXI. Jaime Sabal (a) August 27, 1953 (b) Twenty-seven (c) Seville, Spain; France; Morocco; North America; Hawaii (e) Ann Arbor since 1939 (f) Over twenty years (g) Teaching fellow in Zoology, Univ. of Mich. (h) Spain.

XXIII. Marisa Reguzzoni (a) March 25, 1954 (e) Milan, Italy.

XXIV. Rud S. Meyerstein (a) March 25, 1954 (e) Zurich, Switzerland.

XXV. Charles E. Koella (a) March 25, 1954 (e) France, Switzerland.

XXVI. Gerald Dykstra (a) March 25, 1954 (e) Ann Arbor, Michigan.

The following sentences constitute the Spanish data that was recorded on the Ampex tape recording machine by informants I through XXI. The numbering of the sentences corresponds to that used for reference in the descriptions of the figures found in this thesis.

1. La madre del zurdo ha cocido cinco pollos ayer para el muchacho que se halla en la silla.

2. El abogado debe juzgar sin excepción al soldado que haya robado tu dinero.

3. El otro perro se comió la carne y yo no la puedo hallar con el salchichón.

4. La comida para el huésped está servida; llamadlo y tomad su yegua para que coma hierba.

5. El sol dió a las islas y al istmo una extraña luz dorada.

6. El pastor camina despacio porque las botas le maltratan los huesos de los dedos.

7. Desde la excursión israelita Vd. y yo somos como hermano y hermana.

8. La luz del día es lo mismo que en la edad dorada de la literatura española.

9. El llavero que llevo en la bolsa es exacto al que tiene el prieto.

10. Cristo murió en la cruz para el perdón de los pecadores en todo la esfera terrestre.

11. Para adquirir dos docenas de fósforos necesito dos reales.

12. Espero la orden del médico para ponerle la inyección a mi cónyuge.

13. El éxito conyugal del yerno depende del trato a la suegra.

14. Recuerdo que en mi primer examen saqué cero por auxiliar a mi hermana.

15. Para ascender al cerro tuvimos que aliarnos a un grupo y caminar en exceso hasta encontrar los escollos del avión.

16. Escribidnos una canción con ritmo español.

17. El acusado parado en el poyo exhaló ayer su último escolio.

18. El Senor Smith, adscrito a la Embajada Americana se comió un pollo con hiel.

BIBLIOGRAPHY

Alarcos Llorach, Emilio. Fonología Española. Madrid: Editorial Gredos, 1950. Pp. 160.

Bell Telephone System Technical Publications, Monograph B-1415. Technical Aspects of Visible Speech. Pp. 89.

Bloomfield, Leonard. "A Set of Postulates for the Science of Language," Language. Vol. II, 1926. Pp. 153-164.

Bloomfield, Leonard. Language. New York: Henry Holt and Company, 1933. Pp. ix - 564.

Bolinger, Dwight L. "Evidence on x," Hispania. Vol. XXXV, February, 1952. Pp. 49-63.

Cooper, F. S. "Spectrum Analysis," The Journal of the Acoustical Society of America. Vol. XXII, 1950. Pp. 761-762.

Cooper, F. S., Delattre, P. C., Liberman, A. M., Borst, J. M., and Gerstman, L. J. "Some Experiments on the Perception of Synthetic Speech Sounds," The Journal of the Acoustical Society of America. Vol. XXIV, 1952. Pp. 597-606.

Cooper, F. S., Liberman, A. M., and Borst, J. M. "The Interconversion of Audible and Visible Patterns as a Basis for Research in the Perception of Speech," Proceedings of the National Academy of Sciences. Vol. XXXVII, 1951. Pp. 318-325.

Crandall, I. B., and MacKenzie, D. "Analysis of the Energy Distribution in Speech," The Bell System Technical Journal. Vols. I-II, 1922-23. Pp. 116-128.

Delattre, Pierre. "The Physiological Interpretation of Sound Spectrograms," PMLA. Vol. LXVI, 1951. Pp. 864-875.

Dulany, D. Avoidance Learning in Perceptual Defense and Vigilance. Thesis, 1955. Pp. 85.

Guilford, J. P. Fundamental Statistics in Psychology and Education. 1st. ed.; New York and London: McGraw-Hill Book Co., 1942. Pp. xi - 333.

Henríquez Ureña, Pedro. "Observaciones sobre el español en América," Revista de Filología Española. Vol. VIII, 1921. Pp. 357-390.

Howes, D. H., and Solomon, R. L. "Visual Duration Threshold as a Function of Word Probability," Journal of Experimental Psychology. Vol. XLI, 1951. Pp. 401-410.

Jakobson, R., Fant, C. G. M., and Halle, M. Preliminaries to Speech Analysis. 2d. print. Cambridge Acoustics Laboratory, Massachusetts Institute of Technology, 1952. Pp. viii - 58.

Jespersen, O. Phonetische Grundfragen. Leipsig: B. G. Teubner, 1904. Pp. 185.

Joos, Martin. Acoustic Phonetics. Language Monograph No. 23. Baltimore: Waverly Press, Inc., 1948. Pp. 136.

Joos, Martin. "Description of Language Design," The Journal of the Acoustical Society of America. Vol. XXII, 1950. Pp. 701-708.

Kantor, J. R. The Logic of Modern Science. Bloomington: Principia Press, 1953. Pp. xvi - 359.

Koenig, W., Dunn, H. K., and Lacy, L. Y. "The Sound Spectrograph," The Journal of the Acoustical Society of America. Vol. XVIII, 1946. Pp. 19-49.

Lazarus, R. S., and McCleary, R. A. "Autonomic Discrimination without Awareness," Psychological Review. Vol. LVIII, 1951. Pp. 113-123.

Morgan, C. T., and Stellar, E. Physiological Psychology. New York: McGraw-Hill Book Company, Inc., 1950. Pp. ix - 609.

Navarro Tomás, Tomás. Fonología española. Syracuse: Syracuse University Press, 1946. Pp. 217.

Navarro Tomás, Tomás. Manual de pronunciación española. 4th ed. Madrid, 1932. Pp. 326.

Peterson, G. E. "The Information - Bearing Elements of Speech," The Journal of the Acoustical Society of America. Vol. XXIV, 1952. Pp. 629-637.

Peterson, G. E. "The Phonetic Value of Vowels," Language. Vol. XXVII, 1951. Pp. 541-553.

Peterson, G. E., and Barney, H. L. "Control Methods Used in a Study of the Vowels," The Journal of the Acoustical Society of America. Vol. XXIV, 1952. Pp. 175-184.

Pike, K. L. Language. Ann Arbor: Braun and Brumfield, 1954. Preliminary format, lithoprinted. Pp. 96.

Pike, K. L. Phonetics. Ann Arbor: University of Michigan Press, 1943. Pp. ix - 182.

Pike, K. L. "Operational Phonemics in Reference to Linguistic Relativity," The Journal of the Acoustical Society of America. Vol. XXIV, 1952. Pp. 618-625.

Postman, L., Bronson, W. C., and Gropper, G. L. "Is there a Mechanism of Perceptual Defense?" Journal of Abnormal Social Psychology. Vol. XLVIII, 1953. Pp. 215-225.

Potter, R. K., Kopp, G. A., and Green, H. C. Visible Speech. New York: D. Van Nostrand Company, Inc., 1947. Pp. xvi - 441.

Reed, D. W., Lado, R., and Shen, Y. "The Importance of the Native Language in Foreign Language Learning. Part I - The System of Significant Sounds." Language Learning. Vol. I, 1948. Pp. 17-23.

Solomon, R. L., and Howes, D. H. "Word Frequency, Personal Values and Visual Duration Thresholds," Psychological Review. Vol. LVIII, 1951. Pp. 256-270.

Solomon, R. L., and Postman, L. "Frequency of Usage as a Determinant of Recognition Thresholds for Words," Journal of Experimental Psychology. Vol. XLIII, 1952. Pp. 195-201.

Steinberg, John C. "Application of Sound Measuring Instruments to the Study of Phonetic Problems," The Journal of the Acoustical Society of America. Vol. VI, 1934. Pp. 16-24.

Swadesh, Morris. "Lexico-Statistic Dating of Prehistoric Ethnic Contacts," Amer. Philosophical Soc. Proceedings. Vol. XCVI, 1952. Pp. 402-463.

Tanner, W. P., Jr., Psychophysical Application of the Theory of Signal Detectability. Based on University of Michigan Engineering Research Institute, EDG Technical Report No. 18. February, 1954. Pp. 12.

Trager, George L. "The Phonemes of Castillian Spanish," Travaux du Cercle Linguistique de Prague. Vol. VIII, 1939. Pp. 217-222.