DOCUMENT RESUME

ED 025 967 FL 000 784

By-Delattre, Pierre
Research Techniques for Phonetic Comparison of Languages.
Pub Date 63
Note-14p.
Journal Cit-IRAL; v1 n2 p85-97 1963
EDRS Price MF-$0.25 HC-$0.80
Descriptors-Articulation (Speech), Audio Equipment, *Comparative Analysis, Consonants, English, Films, French, German, *Language Research, Linguistics, *Modern Languages, Phonetic Analysis, *Phonetics, *Research Methodology, Second Language Learning, Spanish, Spectrograms, Syllables, Vowels

This is the first in a series of articles on general phonetic characteristics of American English. Language research techniques of spectrographic analysis, spectrographic synthesis, articulatory motion-picture study, and statistical calculation are briefly described. Thirty-five areas of phonetic difference for comparison of English to Spanish, German, and French are identified. These include general comparisons of vowel color, consonants, and declarative intonation shapes. Two spectrograms are pictured. For related documents see FL 000 781 and FL 000 782. (AF)

Sonderdruck aus IRAL Vol. I / 2

RESEARCH TECHNIQUES FOR PHONETIC COMPARISON OF LANGUAGES

Pierre Delattre, University of Colorado

Julius Groos Verlag Heidelberg

RESEARCH TECHNIQUES FOR PHONETIC COMPARISON OF LANGUAGES [1)]

Pierre Delattre, University of Colorado

Dans cet article, Delattre décrit le laboratoire de phonétique comparée établi à l'Université du Colorado. Ce laboratoire est équipé pour exécuter un programme intégré de recherches dans les domaines de l'analyse acoustique, la synthèse acoustique et la cinématographie, normale et à rayons X, de l'articulation. En outre, des analyses statistiques sont exécutées pour établir de façon quantitative les structures de son, d'accent et d'intonation qui prédominent dans une langue, par comparaison aux autres. Les langues actuellement en étude sont: l'anglais, le français, l'allemand et l'espagnol. Ces recherches ont comme but de fournir des renseignements phonétiques précis pour l'enseignement des langues.

La partie principale de l'article présente une série de brèves annotations sur les quelques trente-cinq aspects des différences phonétiques parmi les langues vers lesquels s'orientent les investigations du laboratoire. Dans des articles ultérieurs, les résultats déjà acquis seront présentés en trois parties: voyelles, consonnes et prosodie.

Delattre beschreibt das an der Universität Colorado errichtete Laboratorium für vergleichende Phonetik. Seine Ausrüstung ermöglicht Untersuchungen auf dem Gebiet der akustischen Analyse, Sprachsynthese und der Verfilmung von Artikulationsvorgängen auf gewöhnlichen und Röntgenfilmen. Weiterhin werden dort statistische Analysen gemacht, um quantitative Vergleichsmöglichkeiten zu Laut, Akzent und Intonation anderer Sprachen zu bieten. Die vorliegende Untersuchung erfaßt Englisch, Französisch, Deutsch und Spanisch mit dem Ziel, zuverlässige phonetische Daten für den Sprachunterricht zu liefern.

Der größte Teil des Aufsatzes befaßt sich mit einer Reihe kurzer Ausführungen über ungefähr 35 Aspekte phonetischer Unterschiede zwischen den Sprachen, die im Mittelpunkt der Untersuchungen des Laboratoriums stehen. In weiteren Artikeln werden die bereits erarbeiteten Ergebnisse in drei Teilen geboten werden: Vokale, Konsonanten und Prosodie.

INTRODUCTION

The teaching of modern languages in our schools and colleges would certainly improve if we could describe very objectively the new phonetic habits that are to be acquired by a student when, or before, he learns a second language. To reach this goal, advantage must be taken of the extraordinary technical advances made during the last decade or two in speech research centers, mainly the Haskins Laboratories of New York, but also the Massachusetts Institute of Technology Acoustics Laboratories of Boston, the Bell Telephone Laboratories of Murray Hill and the Royal Institute of Technology Speech Transmission Laboratory of Stockholm. While the phonemicists were absorbed in the formulation of rigorous methods for the structural analysis of languages as systems, some experimental phoneticians achieved gigantic progress, in those laboratories, concerning the acoustic cues – components of distinctive features – that correlate with the perception of vowels, consonants, and prosody. First, they analyzed the acoustic elements of speech by measurement on spectrograms. Spectrography makes the complex patterns of syllables accessible to the naked eye in a three-dimensional display of *time, frequency* and *intensity,* and shows a different pattern for every vowel, every consonant, as well as dialectal and

1) The research reported herein was performed pursuant to a contract with the U.S. Office of Education, Department of Health, Education, and Welfare.

individual traits of pronunciation or marks of a foreign accent. On spectrograms, differences of duration, intensity, pitch, syllabication, diphthongization, affrication, aspiration, nasalization, vowel color, consonantal place of articulation, and so on, can be located and measured. Figure [2] will perhaps give an idea of what can be observed on a spectrogram to those who are not familiar with this kind of speech display. Second, experimental phoneticians synthesized the acoustic features of speech. All the information found on spectrograms is not linguistically relevant; some of it, for instance, refers to voice quality. It became necessary, therefore, to construct machines that could play back the spectrographic patterns of speech so that one could make omissions and modifications and judge the effects by ear. Such artificial speech synthesizers turned out to be ideal tools for research. By transforming hand-painted patterns into very intelligible speech they made it possible to isolate the acoustic cues of speech perception, to vary their dimensions of duration, frequency, and intensity, and determine their optimum values. This research technique has been so productive that a trained spectrographer can, today, paint, free hand, patterns of speech that will produce whatever dialectal traits he wishes, when passed through the synthesizer. Finally, experimental phoneticians were curious to relate the constantly changing patterns portrayed on spectrograms to the constantly changing cavities of the vocal tract, and they developed the use of motion picture X-rays, to replace the still X-rays. To reduce drastically the dangers of radiation, light intensifiers were invented. Then, to verify theories of speech production based on cineradiography, artificial mouths were built — electric analogs of the vocal tract — which produced excellent syllables by changing only three parameters: the mouth constriction, lip constriction and cavity volumes. As a result, the basic theory of speech production has become clear and simple. It involves a source of sound and cavity resonance of that sound. In vowels the source is at the glottis; for consonants it may be at any point of constriction between the glottis and the lips.

It occurred to us that the same technique as, for instance, that which was used at Haskins Laboratories to study the acoustic correlates of speech perception could be applied to the study of foreign accents, that is, the phonetic features that really make the difference between one language and another. Thus the most advanced electronic techniques would serve a practical purpose: better teaching of foreign languages. Fortunately, instrumental techniques of artificial speech synthesis were as familiar to us as the problems of language teaching. Under contract with the U.S. Office of Education, we undertook in 1959 the construction of appropriate equipment for the phonetic comparison of languages. By now this contruction is basically completed, except for such refinements as research alone can indicate, and considerable research has already been accomplished.

THE LABORATORY AND ITS TECHNIQUES

Our research equipment for speech can be described under three headings: acoustic analysis, acoustic synthesis, and articulatory study by motion-picture.

For acoustic *analysis,* we adapt sound spectrographs to our particular purpose. To the Kay Sonographs [3] we add a frequency marker, by means of a complex tone

[2]) See p. 96

[3]) [Editor's note]. The commercial form of the sound spectrograph is the Kay Sonograph, manufactured by the Kay Electric Co., Pine Brook, New Jersey, USA.

oscillator, to assure perfect calibration for all conditions of heating or zero line; a cooling system; and external vernier controls for the zero line. We modify the time scales of 5 in. per second to make it correspond to standard tape recording and to our synthesizer scale of 7.5 in. per second. To the 2000 cps per inch frequency scale we add one at 1200 cps to correspond to our synthesizers, and another at 200 cps, which enlarges the lower harmonics by ten times, to facilitate the study of intonation curves (the undulating rise and fall of the first or second harmonic generally offers a direct reading of frequency variations). We also connect to the sonograph two accessories: one is a variable-frequency power supply for speed-control of the sonograph turntable motor. It permits judging by ear and measuring objectively the frequency changes that result from varying the playback speed of a speech segment. One can observe the conditions under which an English [ɛ] changes to a German [ø] (a slight slowing down lowers both lower formants of [ɛ] to their frequency position for [ø]), or a Spanish [s] changes to a French [s] (a slight speeding up raises the frequency of the friction, which is higher in French because the s sounds are more fronted and less apical). Such speed-control would also show frequency relations among fricatives: as the frequency of a [ʃ] noise is raised by continuous acceleration, it is heard successively as [s], [θ], and [f]. The other accessory is a sound-gating device. This device makes it possible to separate from an utterance any portion as short as one centisecond or as long as 2400 centiseconds, and to listen to it without erasing the original utterance. Thus one can hear separately the various parts of a diphthong. One can also find the optimum point of separation between vowel and consonant and determine to what extent the consonant-to-vowel transition belongs to the perception of the consonant. This is done by displacing at will the point of separation and listening to both portions separately.

The preceding device operates on *time*-segments (vertical sections on a spectrogram). It is also possible to separate *frequency*-bands from an utterance (horizontal layers on a spectrogram). For this we use the analyzer part of the vocoder that we have built ourselves. This analyzer divides the utterance into twenty frequency bands, any one of which can be omitted in order to judge its contribution to the total perception. Thus it is found, for instance, that practically all vowel information is below 3000 cps, but the relevant consonant noises of the [s], [θ], [f] type are well above that.

It should be noted here that separating formants from an utterance is very different from separating frequency bands. Formant frequencies rise and fall, and cut across the bands of frequency. In order to isolate whole formants and judge their contribution to speech perception, one needs not only an analyzer but an artificial speech synthesizer — a far more involved piece of equipment.

For acoustic *synthesis* we have constructed two sizable machines, both of which permit the isolation and manipulation of the acoustic factors of speech, the acoustic correlates of distinctive features as well as, for instance, the acoustic elements of a foreign accent. These two machines are conceived on different electronic principles and complement one another. Synthesizer I has 50 channels, the 50 harmonics of a series, produced by a tone wheel. Because of the great number of channels it is efficient in the use of formant transitions and produces very sharp consonants, but it cannot vary its fundamental frequency and speaks in a monotone. Synthesizer II uses the synthesizer part of a Vocoder and has only 18 frequency channels. Its consonants are not so sharp but it can vary the fundamental by simple hand-painting of frequency curves, making possible the experimental study of pitch variations (intonation). Both machines play back spectrographic patterns of speech by light reflection. The patterns can be made on spectrographs or be painted by hand. They must be white to permit reflection of the light from a powerful arc.

These two synthesizers are at the heart of our research equipment. Observing a difference between a German and a Spanish syllable on a spectrogram permits no more than hypothesizing. In order to ascertain the relevancy of this difference as a cross-linguistic contrasting factor, it is necessary to synthesize the utterance with and without the feature in question and compare the results by ear. To a considerable extent, then, the synthesizers permit the verification of every hypothetical difference between two languages, in isolation or in combination. It eliminates many of the speculative aspects that have marked this kind of research in the past.

The third section of our laboratory is for *motion-picture* study of tongue, jaws, velum action taken by X-ray or of lip action taken normally. In order to observe the *movements* of articulation that take place inside and outside of the mouth, as well as the *positions* of articulation, on motion-picture films, we have two types of viewers: 1) projectors with sound, which must be used at regular speed (24 frames per second) if the sound is to be heard without distortion; to facilitate observation, we have added stop, go and reverse push-buttons to such projectors; 2) projectors without sound but with a hand-crank permitting one either to stop at each frame or to watch at will in forward-backward motion short sequences of frames. On X-rays, considerable detail is made visible by motion which cannot be detected on a stopped frame. It is a combination of stop and motion which reveals the most. With two hand-cranked projectors, one can compare on the same screen the articulatory positions of a similar syllable in two languages, frame by frame or sequence by sequence. In addition to sound and silent projectors, we have film readers that project the frames on a flat surface, where tracing the X-ray shadows or the lip contours becomes possible. We are also able to print frames in any reasonable size.

Motion-picture X-rays are not taken in our laboratories but in appropriate hospitals that are equipped especially for that purpose and know what dosage of exposure a subject can receive without danger of radiation. X-ray laboratories use the language materials (lists of words and sentences) that we have prepared for comparison of English to foreign languages point by point. For each language these materials total about two pages and take three to four minutes to be spoken, a time that is too long for a single person. Here are examples of utterances for the comparison of English [e] and [o] to Spanish [e] and o] in stressed position: *It's a day still to face, Que le dé dos cafés; It's low pay to cope with, Don Lope da la copa.* Conditions of stress and phonic environment are kept as similar as possible in the two languages.

To summarize, our research technique appears to be three-fold: spectrographic analysis, spectrographic synthesis, and motion-picture analysis. This, however, is only from the instrumental aspect. A fourth category, which might be called statistical, plays an important role in showing how certain features prevail quantitatively. Statistical study applies, for instance, to vowels (What are the most frequently occurring vowels? How frequently do we meet those vowels that are very characteristic of a given language and not found in others ?), consonants, clusters, syllable sounds, syllable types (open, close), syllable length, place of stress, falling intonations, etc.

DIVISIONS OF RESEARCH

Preliminary investigations have led us to divide our comparison of English to Spanish, German and French into forty areas of phonetic difference. This does not mean exactly that, so far, we have found forty phonetic differences between English and any one of these three target languages – although this is not completely unrealistic – but that we have found it convenient to do our investigating separately under those many headings. Time will tell how many further subdivisions will be necessary. This dividing system is of interest in itself – we shall present the greatest part of it here with brief indications of the research techniques applicable to each area. Results that have already been obtained will be presented in later articles. Let it be clear that we are interested in all *phonetic* features of English and the target languages that can help in teaching the latter.

1. *General Comparison of vowel color.* To compare the vowel phonemes of a language to those of another in a manner that will be useful to the teacher, both the articulatory and the acoustic description must be presented. *Articulatory.* a) The profile of the tongue and lip positions for each vowel must be obtained not from X-ray stills, as in the past, but from X-ray motion pictures which show the same position with much more constancy. After comparing several subjects and taking into consideration the influence of various contiguous consonants, schematized tongue and lip shapes must be drawn. b) Phonetic charts of tongue position can be made in two dimensions: high-low, and front-back, using the highest point of the tongue hump. Two sets of such tongue charts are necessary, one with lip spreading, the other with lip rounding. These two can be joined in a double chart, correlating fairly well with the acoustic vowel charts (see below) if the rounded-vowel triangle (or quadrilateral) is placed somewhat lower and to the right of the unrounded vowel triangle. This is explained by the fact that lip rounding enlarges and closes the front cavity thereby lowering the frequency of the second formant, and that rounded vowels make their constriction slightly less fronted than the corresponding spread vowels, which reduces the volume of the back cavity and raises the frequency of the first formant. *Acoustic.* Vowels may be considered as musical chords, mainly characterized by the frequency (pitch) of their two lowest formants. Spectrograms show those two formants (different for every vowel) and make it possible to measure their frequency. a) The spectrographic aspect of the whole vowel system of each language can be schematized. b) An acoustic chart of vowels, quite similar to the triangular or quadrilateral articulatory chart based on the highest point of the tongue, can be made for each language by graphing on logarithmic paper the first formant frequencies in ordinate and the second formant frequencies in abscissa. Such acoustic charts have many advantages over the articulatory charts. They are more objective since formant frequencies are more subject to measurement than somewhat unstable tongue shapes; they represent more complete information since they rely not only on tongue position but on lip shape and any other articulator responsible for the

volume of vocal tract cavities. A comparison of acoustic charts among languages will give a true picture of relative qualities of vowel systems: their richness, their balance, their center of gravity, their regions of concentration such as low, front, back, center, etc. It will show relations, for instance, between duration and place of production, nasality and place of production.

[A technical parenthesis must be made here, however. The frequencies of characteristic formants cannot be obtained only by measurement of spectrograms; they must be verified by *synthesis.* To arrive at truly representative figures for a standard voice pitch (fundamental frequency), artificial vowels must be tested by ear. This is necessary because formant frequencies vary somewhat with individual voices. The formants of a feminine voice can be 10 to 15 percent higher than those of a male voice. Besides, verification by synthesis will prevent errors of all sorts that have crept into too many spectrographic studies of late, errors of formant location, errors of measurement, and errors due to faulty adjustment of the spectrograph itself.]

Correlation. The acoustic and the articulatory are naturally related. a) Formant frequencies actually are the notes of resonance of the vocal tract, which is generally divided into two cavities, one in the pharynx (back cavity) and one in the mouth (front cavity). The larger its volume, and the smaller and longer its opening, the lower the note of resonance of a cavity. Roughly speaking, therefore, observing changes in the frequencies of first and second formants is like observing modifications in volume and openings of the back and front cavities, respectively. b) Some practical relations may be more useful to the teacher than the one above: tongue fronting raises the second formant (by decreasing the front cavity volume); lip rounding lowers the second formant (by increasing the volume and closure of the front cavity); and tongue lowering raises the first formant (by decreasing the back cavity volume). But those relations are indirect. For instance, whenever there is tongue lowering, in series such as [i, e, ε, æ], [y, ø, œ, a], or [u, o, ɔ, ɑ], as the tongue lowers, the place of constriction moves toward the back until it is far down into the pharynx for the [a] family, leaving only a small back cavity between mid-pharynx and the glottis. It is the backing of the place of constriction which directly decreases the volume of the back cavity, thereby raising its resonant frequency. The lowering of the tongue is only an indirect factor. c) One can see now that the relation between the articulatory chart based on the highest point of the tongue and the acoustic chart based on the frequencies of formants is an indirect one: backing and lowering the highest point of the tongue changes the place of constriction, which in turn changes the volume and opening of the resonating cavities.

Dealing somewhat at length with a resonance aspect of vowel production will help us understand more rapidly many other features of speech.

2. *General comparison of consonants.* The perception of consonant noises, like that of vowels, depends on cavity resonance in the vocal tract. But whereas the source of sound (the sound to be resonated) is always far back at the glottis (vocal cords) for vowels, it can be anywhere between the glottis and the lips for consonants, and only the portion of the vocal tract that is located in front of the consonant constriction resonates the consonant noise quite efficiently. Consonant noises are mainly distinguished by their frequency: [ʃ] noise is lower than [s] noise, but higher than [x] noise; and that frequency depends on the volume and closure of the cavity in front

of the constriction. In a general way, therefore, the farther back the constriction, the larger the resonance cavity and the lower the consonant noise.

The perception of consonants, however, does not depend only on consonant noise. It relies also on the rapidly changing formant frequencies (formant transitions) which reflect the rapid movements of articulation taking place from vowel to consonant or consonant to vowel. Consonant is movement; vowel is steady-state. In a syllable [do] for instance, only the relatively steady-state end of the formants belong to the vowel; the rapidly changing formant transitions that can be seen on a spectrogram between the [d] explosion and the [o] steady-state belong entirely to the perception of the [d]. *Articulatory.* To compare consonants among languages, a teacher will first need a) complete sets of profiles of tongue and lip positions as well as b) a schematized profile view of all the places of articulation for each language. X-ray motion pictures will provide this information, but not X-ray stills. *Acoustic.* a) A teacher also needs a chart of the frequencies of consonant noises in two parts: one for friction noises such as in [f, θ, s, ʃ, ç, x, X], another for burst noises such as in [p, t, c, k]. *Correlation* of these noise frequencies with the place-of-articulation profile will enable the teacher to show, for instance, why the [s] sounds of English are lower (in frequency) than those of French, and those of Spanish lower yet. b) Acoustic charts showing the role of formant transitions in the perception of consonants are probably too involved to have any practical value in teaching.

The frequencies of noises will be suggested by spectrographic analysis, but, as in the case of vowels, they will require confirmation by synthesis in artificial syllables.

3. *New sounds.* A language teacher will welcome a separate list, with acoustic and articulatory descriptions, of the sounds in the target language that are radically different from English, as well as a list of the sounds in English that are radically different from the target language. Acoustic and articulatory charts of these new sounds will bring out the regions and types of articulation that are the most characteristic and cause the greatest interference in learning.

4. *Distribution.* With the phonemes of a language, a teacher needs to know a) in which positions they occur, b) how they vary according to position. The former will show, for instance, that short vowels and voiced plosives are not found in final position in German; the latter that postvocalic [l] is not retroflex in French.

5. *Sound frequency.* To complement the information conveyed by comparative charts of distinctive sounds and their distribution, it is indispensable to know also the frequency of occurrence of each sound. a) The acoustic impression made by the target language on a new student depends on which sounds recur more than others, what the proportion of vowels to consonants is. Initial teaching material should perhaps make use of such knowledge. b) The importance given to "new sounds" in the early stages of learning should be considered in the light of their frequency of occurrence. Frequency counts must be based on conversational or dramatic material as well as on narrative material, and if the two types are found not to diverge markedly the results can be combined.

6. *Syllable characterization.* Another frequency count that may throw some light on the phonic characterization of a target language concerns a) the syllables that recur most often, their actual phonemic composition, their proportion of vowels to consonants; b) types that prevail in terms of "open syllables" (CV, CCV) or "closed syllables" (CVC, VC); c) the length of syllables, in terms of the number of phonemes, and the relation of length to frequency of occurrence and to speed of articulation. These last two areas (5 and 6) are strictly statistical.

7. *Typical tongue shape.* Each language seems to have a characteristic shape of the tongue, either in key consonants and vowels or in rest positions to which the tongue tends to return. Motion picture X-rays may make it possible to compare such a factor among languages.

8. *The neutral vowel.* When an American child hesitates, the vowel he utters is mid-central to mid-open-back with lips nearly unrounded. It is close to the unstressed schwa of sof*a*. A French child uses a mid-front well rounded vowel under the same circumstances. X-ray views of such articulatory positions should be very useful in teaching. This neutral vowel is a sort of home base to which the tongue returns frequently in the course of speech.

9. *Tension.* Objective measurement of the degree of tension or laxness in the muscles of articulation has not yet been possible. If such a thing exists as a language characteristic, it might correlate with such factors as the proportion of steady-state vs movement, the speed of movement, the time it takes to pass from fast change to relative steady-state, etc., which can be measured on spectrograms and tested by ear in synthesis.

10. *Duration systems.* Variations of vowel duration are of two very different kinds: those that are conditioned, automatic, unconscious and nondistinctive, such as the effect of stress, the rate of utterance, the influence of the following consonant -- [I] is shorter in *mitt* than in *mid;* and those that are learned and contribute to a phonemic distinction, such as in German *Stadt* vs *Staat,* in English *bit* vs *beat,* in French *tache* vs *tâche.* A teacher must be aware of both kinds in order not to confuse what should be taught and what should not. Measurements of length are made on spectrograms and verification of length perception by synthetic speech. The same two types of length variations are found in consonants: in intervocalic position single flap vs multiple flap is distinctive in Spanish, [karo/karro], whereas it is not in Italian [ieri-ierri].

11. *Declarative intonation shapes.* These include the two basic and most frequent modes of expression: continuation and finality. *Analysis.* To compare them among the four languages, we use spontaneous speech by intelligent speakers. Hundreds of feet of spectrograms are made in the 2000 cps per inch scale to show formants and high harmonics and in the 200 cps per inch scale to permit one to follow the first or second harmonic, which offers a direct reading of frequency variations. From a statistical study of such undulations, we extract the most typical shapes. *Synthesis.* Then we paint those shapes on Synthesizer II and judge by ear the effect of certain changes and omissions to discover which parts of the shapes are most relevant and what is the subjective impression (pitch) of those objective shapes (frequency) in

each language. We also impose the pitch shapes of one language on the words of another language to judge the degree of characterization. We would like our pitch *shapes* to be the phonetically relevant schematization of the objective frequency shapes. They are not to be confused with the subjective pitch *levels* (2—3—1↓) which can be distinctive within a language but tend to mask differences among languages. Phonetic shapes, which vary sharply from language to language, may be more useful than phonemic levels which use somewhat the same digital combinations in all languages.

12. *Nondeclarative intonation shapes.* These will include frequent attitudes such as command, question, interrogation, implication, parenthesis, echo, exclamation. The techniques of analysis are the same as for declaration.

13. *The place of logical stress.* Here, the method applied is statistical. We analyze a few thousand running *words* of narrative and conversational material in each language, and we extract for the teacher two sorts of information: a) how variable or predictable the place of word stress is; b) the characteristic place of stress in one-, two-, three-, and four-syllable words. In a later section on rhythmic patterns, we will apply nearly the same method to finding the place of stress in sense groups.

14. *The nature of logical stress.* To discover the relative contributions of excesses of intensity, duration, and frequency (objective factors) to giving prominence (subjective impression) to a syllable, the speech *synthesizers* are incomparable tools. The three objective factors can be varied individually and in combination, and the effects judged by ear by natives of each language. Spectrographic *analysis* is limited, here, to making hypotheses.

15. *The place of emphatic stress.* In many languages it is different from the place of logical stress. In French, for instance, emotive stress bears on the first or second syllable, whereas logical stress is on the last. Statistical methods are applied.

16. *The nature of emphatic stress.* Pitch and intensity are used more here than in logical stress. The teacher needs to know how differently these ingredients are mixed in the target language and in his native language. The synthesizers are the main research tools.

17. *Rhythmical patterns.* A statistical study of the place of stress in sense groups is also necessary, because word rhythm (Section 14) and sense-group rhythm turn out to be markedly different in languages with variable place of stress.

18. *Phonetic syllabication.* It must be studied at two levels: a) its phonetic realization when it is distinctive — the German constrast *zum einen* vs *zu meinen* is made quite differently from the English *an aim* vs *a name*; b) its phonetic characteristics when it is nondistinctive — English *aiding* vs French *aider.* A synthesizer makes it possible to isolate the acoustic correlates of syllabication and to vary them at will. They contribute considerably to making a phonic sequence sound German, English, Spanish, or French.

19. *Variations in syllable weight.* Here we study the distribution of intensity, duration and frequency in *all* syllables of a text, unstressed as well as stressed, in relation to

such factors as their position in the phrase, or their grammatical function. This involves spectrographic measurements, statistics, and verification by synthesizers of the character it contributes to each language. In English, for instance, final syllables in unstressed position are found to have less intensity but as much or more duration than non-final stressed syllables.

20. *Loss of vowel color.* To what extent, in each language, do vowels retain the same color under different degrees of stress? How different are the *o*'s in a*bo*lish and abo*li*tion? Here vowel formants are being measured on spectrograms.

21. *Diphthongization.* Languages vary considerably from one another in the way so-called monophthongs actually change in color. Two factors seem relevant here: the amount of change, and the direction of change (toward center, toward ends). These changes can profitably be transferred from spectrograms to vowel charts. They require a thorough acquaintance with consonant spectrography to avoid confusing vowel changes with formant transitions that are related to the perception of consonants.

22. *Diphthongs.* Spanish and German have three diphthongs in common with English: [ai], [au], and [ɔi]. But the phonetic realization of these diphthongs is different in each language in at least three ways: the color of initial and target vowels, the relative amount of steady-state of each portion, and the decrease of intensity from beginning to end. Spectrograms, correlated with X-ray motion pictures of the tongue movements plus front- and side-view motion pictures of the lip movements, are being studied.

23. *Vowel nasality.* Its phonetic realization as a distinctive feature (as in French) must be studied separately from its occurrence as a nondistinctive characterization (as in American English or in Spanish). Both can be seen on spectrograms as well as on X-rays.

24. *Vowel color and syllable type.* The type of syllable — closed or open — does not affect a vowel to the same extent in all languages. For instance, the tendency for *ai* to open in closed syllable (*j'aime)* and to close in open syllable *,'ai)* is very marked in French. Similar tendencies also exist in other languages, but to a lesser extent. Spectrographic measurements of formants will specify the differences.

25. *Vowel attack and release.* The sharp attack, with or without vocal cord explosion, is used more in Germanic than in Latin languages. Various degrees of attack have to be studied experimentally by spectrographic analysis, vocal cord study, and vowel synthesis.

26. *Time dimensions of consonants.* One character of the consonants that differs from language to language is relative speed of closing, duration of closure, and speed of opening. These dimensions are being studied by motion pictures of lip and tongue (X-ray) movements in comparative sentences such as: *Two galley bees are dozing in a candy pail,* vs *Tout son cadeau est un habit payé à la gare.* To make use of the maximal extent of formant transitions, velars must be before central vowels, dentals before back vowels, and labials before front vowels, as in these sentences.

27. *Consonant fronting.* The degree of tongue-tip fronting for all the apical consonants contributes markedly to the characteristic overall resonance of a language, for it affects not only the consonants themselves but the contiguous vowels as well. It is being studied by means of X-rays and spectrograms in comparative sentences such as: *Take the same native shale from the lake in the dale,* vs *Les aînés ont aidé chez les Leger l'été passé.*

28. *Consonant aspiration.* How late the glottis remains open after the plosive burst of noise can be measured on spectrograms and reproduced on synthesizers. Comparative sentences such as *The publication of a text*... vs *La publication d'un texte*... are used.

29. *Consonant fricativation.* This is mostly found in Spanish, when voiced plosives stand between vowels. The acoustic and articulatory nature of those fricativated plosives needs to be compared with actual plosives and actual fricatives in the other languages, also between vowels. Comparative sentences: *Dan was good about bidding long ago,* vs *Gobemo el abogado sin duda en Vegas.*

30. *Consonant affrication.* Affrication can be distinctive as in *tico* vs *chico* in Spanish, or nondistinctive as in English *tick, kit* vs French *tic, quitte* with less friction after the burst than in English. To study this phenomenon we correlate X-rays with spectrograms.

31. *Consonant palatalization.* English tends to combine a yod with a contiguous lingual consonant: *miss you* is often pronounced [mIʃu]. Other languages may behave quite differently under the same circumstances. X-rays and spectrograms make the comparison possible among languages.

32. *Consonant release.* The degree of release of final consonants differs from language to language. Spectrograms show this and indicate how consonant closure correlates with it. Tests of the contribution of release to perception of place of articulation can be made by synthesis as well as by tape-splicing manipulation.

33. *Consonant anticipation.* Post-vocalic consonants show a tendency to close earlier in some languages than in others, and after certain vowels than after others: German *bitte* vs *biete.* Arresting transitions of the formants show this by their length and their rate of change. It can also be studied in motion pictures. Marked consonant anticipation gives a jerky impression to the foreign listener.

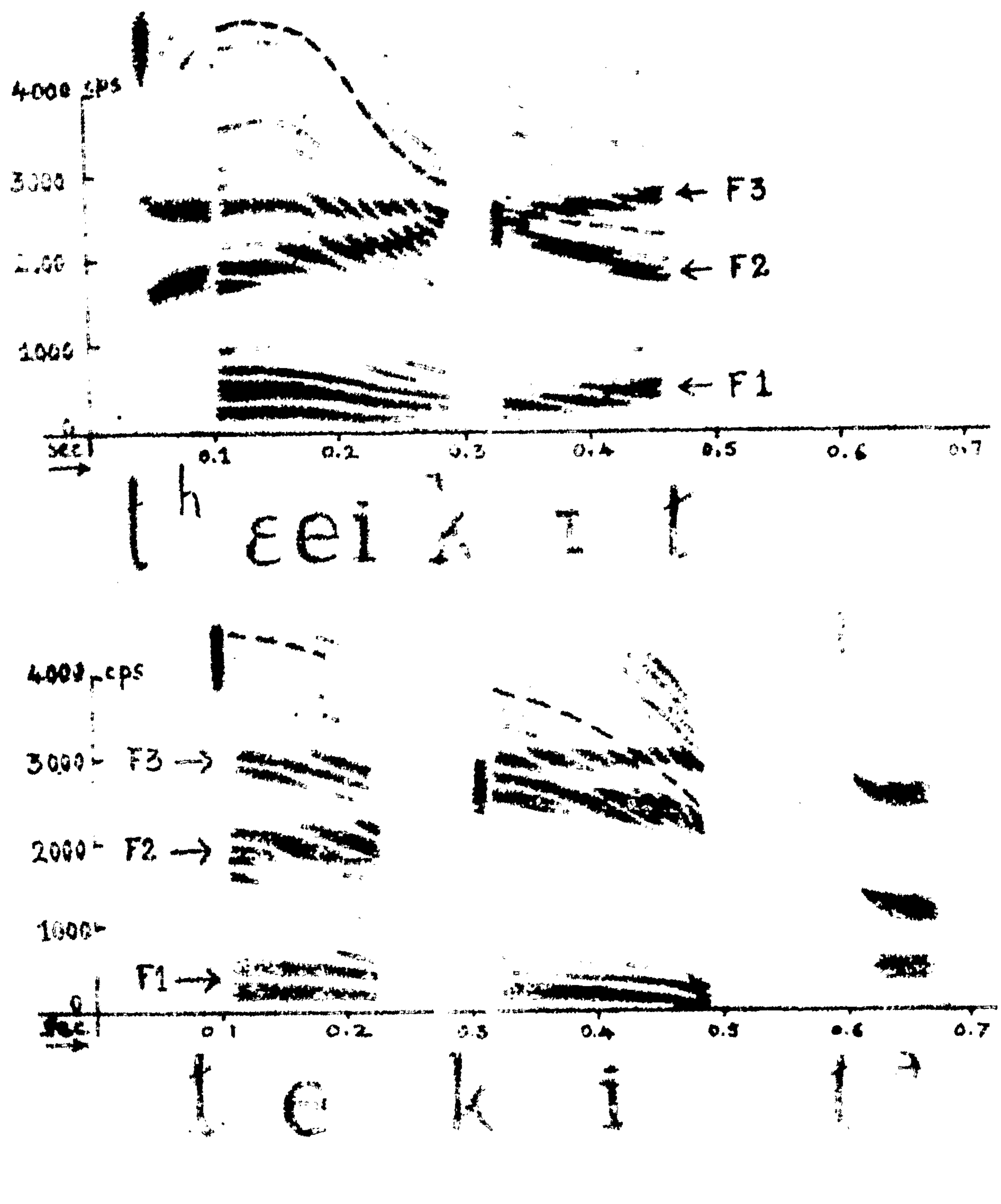
34. *Vowel anticipation.* The degree to which one takes the position of a vowel while, or even before, articulating the preceding consonant, varies considerably among languages. Sentences such as the following permit motion picture comparison between English and German: *The boat sailed down the Po,* vs *Das Boot schwimmt auf dem Po.*

35. *Consonant voicing.* It appears that initial [d] voices earlier in Spanish than in English and conversely intervocalic [t] voices more in English than in Spanish. Such voicing and unvoicing conditions in all positions are being compared among languages as they pose important problems of interference in learning. The acoustic and articulatory correlates of voicing are found to be extremely complex, however, and cannot be limited to the presence of a "voice bar" on spectrograms.

SUMMARY AND CONCLUSION

Our purpose is the improvement of language teaching. We have described our four-way research technique – spectrographic analysis, spectrographic synthesis, articulatory motion-picture study, and statistical calculation – for the comparison of the phonetic characteristics of English, Spanish, German, and French; and we have very briefly introduced the application of such techniques to some thirty-five phonetic differences among those languages – all differences, the knowledge of which can be invaluable for better teaching. In later articles the actual results that have already been obtained will be presented in three parts: vowels, consonants, and prosody.

Figure 1



Schematized spectrograms of the command "Take it" are shown as uttered by an American (above) and by a Frenchman (below). The thin lines represent the rich range of harmonics produced by the vocal cords. Those harmonics are thicker (more intense) whenever their frequency coincides with the continuously changing notes of resonance of the vocal tract cavities. From left to right, the darker lines formed by the thicker portions of the harmonics are the "formants." *Phonemically,* note that the first formant (F1) is higher, and the second lower, for [e] or [I] than for [i]; for [t]'s, the bursts are relatively high and the second and third formants originate respectively (or point to) lower and higher than the center frequency of the [e]'s; for [k]'s, the bursts are at the second-formant onset, therefore lower than for [t]'s, and the second and third formants tend to meet. In general, only the first and second formants are necessary to recognize vowels, but the third-formant transitions play an important part in the identification of consonants. *Phonetically,* we note the following differences between the two speakers. 1. Place of Stress. The Am. stresses the first syllable; the Fr. the last. 2. Nature of stress. The Am. shows more syllabic duration, more syllabic intensity, more aspiration, higher pitch (the dashed line) on the first syllable; the French shows more syllabic duration on the last syllable, but not more intensity, or higher pitch. 3. Intonation. It can be followed from left to right on any harmonic, but the higher the harmonic the easier the reading of pitch. The high harmonic that is marked with dashed lines indicates that in order to express command the Am. uses a concave, reversed-S shape; the Fr. uses a convex shape. Boch shapes are falling and could be called: 2—1↓ phonemically, yet they are very different phonetically. 4. Aspiration. The Am. shows it before the stressed vowel; the French shows it nowhere. Before the unstressed vowel, the Am. [k] is lenis — although there are no vocal cord vibrations (the first harmonic is not present) the closure is short enough to create a semi-impression of voicing. 5. Release of the final consonant is shown in Fr.; the Am. speaker does not show any in this rather typical case. Note that voiceless [ə] of consonant release has a first formant, whereas [h] of aspiration does not (this is a phonemic, not a phonetic remark). 6. Am. diphthongization of [e] is recognizable by the falling frequency of the first formant and the rising frequency of the second formant [ɛei] in their mid sections (the faster rises at the beginning and end of the second formant belong not to the diphthongization of the vowel but to the perception of the [t] and [k]. 7. Consonant anticipation is visible in [tɛeik] by the joining of the arresting transitions of the second and third formants by the Am.; the arresting transitions of the Fr. [k] join less than its releasing transitions after the [k] burst.