THE MAKING OF A SPANISH ORCHESTRA OF LANGUAGE BASED ON ROBSON'S THE ORCHESTRA OF LANGUAGE

Denise Munizaga Lagos, M.A.
Morehead State University, 1971

Director of Monograph: Dr. Ruth Barnes.

Ernest M. Robson provides in his book, The Orchestra of Language, a theory in which significant sound patterns, vowel and consonant combinations, are developed through a parallel with musical orchestral groups.

In his system Robson distributes the sound combinations of English in the same way in which the instruments of a musical orchestra are distributed. Thus, he considers in this orchestra of the language eleven orchestral groups arranged according to their auditory qualities. These groups are Percussions, Hisses, Noises and Overtones or Semi-vowels, high-toned vowels, middle-toned vowels, and low-toned vowels.

The most significant point of this orchestra of the language is the statement that these orchestral groups in combination tend to evoke meaning or emotions through their tone, striking power, and time duration.

The purpose of this monograph was to form a Spanish orchestra of language to parallel the English orchestra . In this Spanish orchestra of language eleven Spanish orchestral groups have also been distributed according to their auditory similarities with the English sounds. And it has been determined that these

Spanish groups, in combination, also evoke similar meaning or emotions to the ones evoked by the English sound-combinations through their tone, striking power, and time duration.

The comparison of both orchestras has been made by analyzing one poem in each language, one by W.B. Yeats, "The Lake Isle of Innisfree," and one by Fray Luis de Leon, "Vida Retirada."

The purpose of this analysis was to determine whether similar sound effect could be achieved by using the unique sound-combinations in existence in both languages.

"The Lake Isle of Innisfree" and "Vida Retirada" are two poems dealing with the subject of withdrawing from society and the desire to reach spiritual and physical peace in an isolated place. The mood of both poems is depressing. This depressing, melancholic mood is stressed by using low-toned vowels with consonant clusters of high striking power, a device which helps to stress even more the low tone of the poems.

The conclusion is that the same sound effect can be achieved in both languages by using the unique orchestral groups, and that this linguistic analysis proves to be a reliable method to determine the tone and striking power in language structures.

THE MAKING OF A SPANISH ORCHESTRA OF LANGUAGE BASED ON ROBSON'S THE ORCHESTRA OF LANGUAGE

A Monograph

Presented to

The Faculty of the Department of English

Morehead State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

---

by

Denise Munizaga Lagos

August 1971

Accepted by the faculty of the School of Humanities, Morehead State University, in partial fulfillment of the requirements for the Master of Arts degree.

R. W. Rowe
Director of Monograph

Master's Committee: R. W. Rowe, Chairman
Lyda Novinio
L. W. Rowe

August 6, 1971
(Date)

TABLE OF CONTENTS

Chapter 1

NATURE OF THE MONOGRAPH, PROCEDURE, PREVIOUS WORK, PURPOSE AND SPECIFIC ELEMENTS TO BE PROVEN, AND DEFINITIONS

NATURE OF THE MONOGRAPH

In this monograph the writer will propose an Spanish orchestra of language which can parallel Robson's English Orchestra of Language.[1] Some indication of his idea had been discussed before him, especially the idea that speech sounds--namely vowels and consonants--can be arranged in such combinations or patterns that are able to produce in the reader a state of hypnosis or magic.

Sound-combination is discussed in David I. Masson's "Vowel and Consonant Patterns in Poetry":

> The effect upon the reader or listener, mainly unconscious, is one of decoration when the pattern reaches a certain degree of crudity or, conversely, a certain degree of refinement. Between these limits, however, the pattern tends to exert a species of magical or hypnotic effect, like that of a ritual incantation.[2]

Although the reader may realize that one line sounds better to him than another, he does not generally realize how this effect is produced.

---

[1]Ernest Robson, The Orchestra of Language (New York: American Book-Stratford Press, 1959).

[2]David I. Masson, "Vowel and Consonant Patterns in Poetry," Essays of the Language of Literature, ed. Seymour Chatman and Samuel R. Levin (Boston: Houghton Mifflin Co., 1967), p. 4.

Some evidence indicates that these combinations of sounds are worked out consciously by the poet; however, some poets, such as Edgar Allan Poe, Hart Crane, and Walt Whitman, made clear their realization that certain sound-combinations help them bring about a certain effect better than others. Masson refers to some of these sound-combinations or patterns of particular poems that were used by the poets producing an effect on the reader or listener.

Moreover, all these combinations of sounds such as alliteration, disonance of vowels, consonance, forms, and images form a musical unity that evokes precise emotions. Analysis done in the works of some poets such as Milton, Spenser, Shelley, and Yeats shows the existence of systematic patterning. Furthermore, John Unterecker, in his book A Reader's Guide to William Butler Yeats, mentions Yeats's beliefs in relation to the nature of a poem:

> Yeats...sees the poem as a complex relationship of images, rhythms, and sounds which, in conjunction, become a symbol for emotional experiences otherwise inexpressible in words.[3]

Although the words do not carry the emotions, sound-combinations evoke experience. This concept is the same one developed more than sixty years later, a concept that reaches its climax in the idea of Ernest Robson of creating an orchestra of language.

A good piece of poetry or prose is created in the

---

[3]John Unterecker, A Reader's Guide to William Butler Yeats (New York: The Monday Press, 1959), p. 30.

same way that a piece of music is composed. The vowels and consonants are the instruments used by the poet or writer. Both consonants and vowels combined in certain clusters actually produce music, and this music may evoke in the will of the listener or reader distinct emotions. Ants Oras in his essay "Spenser and Milton: Some Parallels and Contrasts in the Handling of Sounds" comments about the music produced by sounds:

> Vowels with their variations in pitch, length, and overtones, almost literally provide "music," whereas consonants, much more tangible and produced with a more clearly perceptible effort of articulation, supply what may be metaphorically called body, mass, and weight.[4]

Vowels are the tones, the clear sounds; consonants are the noises, and in the same way that an orchestra needs strings, woodwinds and percussion instruments to compose a piece of music, the poet or writer needs vowels and consonants to create poetry or prose.

These ideas, collected and developed by Ernest M. Robson, are to provide the writers with patterns with which they can make words express emotions and ideas more effectively.

According to Robson's work, there are combinations of sounds that because of their tone, timbre and time--which is the duration they have in their enunciation--evoke certain feelings and emotions which are unique to that combination.

---

[4]Ants Oras, "Spenser and Milton: Some Parallels and Contrasts in the Handling of Sound," Essays on the Language of Literature, ed. Seymour Chatman and Samuel R. Levin (Boston: Houghton Mifflin Co., 1967), p. 20.

For example, the sound-combination in the word "struggle" evokes feelings of strength, power, and force mainly because of the use of initial consonant cluster "str" which has a high striking power. In this monograph the writer will determine whether such orchestral groups exist in the Spanish language, and if they do, whether they evoke the same range of feelings.

There are sounds that have a counterpart in another language: for example, one of the varieties of the sound /s/ in Spanish has its counterpart in the English sound /s/. They are both voiceless alveolar fricatives, so they would produce, in isolation, the same kind of response.

But, of course, the same emotions cannot be evoked by the same sound-combinations in different languages. The use of certain sound-combinations is conditioned by the language in which the poet or writer writes. Thus, in English the writer has more opportunities for using more vowel clusters because the English language has many vowel varieties. Spanish with just five vowel phonemes offers fewer opportunities for the writer to form different vowel clusters. English can use initial consonant clusters with at least three different consonant sounds, as in the word "strike." However, Spanish does not employ these initial consonant combinations. The same responses can be reached in the two different languages, of Spanish and English through unique orchestral groups that are parallel in principle, but unique in application.

PROCEDURE

This monograph is developed in four chapters. Chapter One deals with the nature of the monograph, together with the clarification of the linguistic concepts that constitute the basis for Robson's Orchestra of Language. Thus, the points discussed in this section are centered around the relationships that exist between sounds and music. Other points discussed in this chapter are procedures, purposes and specific elements to be proven, as well as definitions.

Chapter Two treats the most significant points in Robson's Orchestra of Language: time duration, striking power, and tone, together with the eleven orchestral groups which Robson distributed in his orchestra. A.C. Gimson's An Introduction to the Pronunciation of English[5] is consulted to support Robson's statements as to timbre, consonants, vowels, semivowels, and diphthongs. For this purpose Gimson's charts and description of sounds are presented at the end of this monograph in the appendices. Furthermore, Robson's charts are presented in the appendices for a better understanding of his theory.

Chapter Three is developed and devoted to the construction of a Spanish orchestra of language that parallels Robson's orchestra. The distribution of vowels, consonants, semivowels, and diphthongs is supported by T. Navarro Tomas's

---

[5]A.C. Gimson, An Introduction to the Pronunciation of English (London: Edward Arnold Ltd., 1969).

Pronunciación Española[6] whose description of sounds and charts are presented in the appendices.

In making an orchestra similar to that of Robson, I will show which Spanish sounds correspond to the English sounds and which do not. The fact that there are similar sounds in English and Spanish will serve as a starting point in the construction of this Spanish orchestra. For this purpose Robert P. Stockwell and J. Donald Bowen's The Sounds of English and Spanish,[7] R. Menéndez Pidal's Manual de Gramática Histórica Española,[8] and T. Navarro Tomás and Aurelio M. Espinosa's A Primer of Spanish Pronunciation[9] are cited.

Chapter Four has an analysis of William Butler Yeats's poem "The Lake Isle of Innisfree" taken from George K. Anderson and William E. Buckler's The Literature of England.[10] This analysis will be based on Robson's theory. The writer of this monograph will determine which responses are evoked

---

[6] T. Navarro Tomás, Pronunciación Española (Madrid: A. G. Fenix, 1959).

[7] Robert P. Stockwell and J. Donald Bowen, The Sounds of English and Spanish (Chicago: The University of Chicago Press, 1966).

[8] R. Menéndez Pidal, Manual de Gramática Histórica Española (Madrid: Espasa-Calpe, S.A., 1941).

[9] T. Navarro Tomas and Aurelio M. Espinosa, A Primer of Spanish Pronunciation (Chicago: Benj. H. Sanborn & Co., 1926).

[10] George K. Anderson and William E. Buckler, The Literature of England (Glenview, Illinois: Scott, Foresman and Co., 1968).

by the sound-combinations of this poem, using Robson's charts of striking power, tone, and time duration.

The reader may wonder why the writer of this monograph will show as evidence for the description of the Spanish sounds works done in the years 1926, 1941, 1959, and why not use newer works which can show the changes that have occurred in the later years.

The fact is that the Spanish language has suffered little or no change for many years because of the uniformity of the language. This uniformity of the Spanish language derives from the fact that the Spanish-speaking nations have one standard grammar which is none other than that one written by the Real Academia Española de la Lengua, or R.A.E.

The R.A.E. rules what is acceptable or not in the Spanish language and has allowed very little change since it was founded in 1713.

The role and effect of the R.A.E. has had in the Spanish language is commented on in Bryan Steel's "Contrasting Approaches to Spanish Lexicography":

> Since the foundation of the Real Academia Española de la Lengua in the eighteenth century, the basic approach to lexicography and grammar in Spain has been marked by its excessive caution and dogmatic purism. The existence of an official body whose duties include those of debating the claims to acceptability of words and syntactical patterns in common or specialized use has meant that both the D.R.A.E. and the G.R.A.E. lag perpetually behind in matters of current usage.[11]

Thus, the concepts and descriptions of sounds con-

---

[11]Bryan Steel, "Contrasting Approaches to Spanish Lexicography," Hispania, LVI (March, 1971), p. 46.

tained in the works of T. Navarro Tomas and R. Menendez Pidal are the ones accepted by the R.A.E.; therefore, no change has taken place lately.

Moreover, Chapter Four will deal with an analysis of Fray Luis de León's "Vida Retirada" taken from Real Academia Española's Biblioteca de Autores Españoles,[12] according to the Spanish orchestra of language, to show which responses can be evoked by the sound-combinations used by the author. The analysis of this poem will also indicate whether similar responses can be evoked in Spanish and English.

The analysis of this poem is based on the charts of the Spanish orchestra of language, and the tone, striking power, and time duration of the sound-combinations is determined from these charts.

PREVIOUS WORK DONE IN THE FIELD

After investigating in such works as the Linguistic Bibliography, the International Bibliography of Books and Articles on the Modern Languages and Literatures (M.L.A.), An Annotated Bibliography of Modern Language Teaching, Books and Articles 1946-1967; Books in Print (Authors), Books in Print (Titles and Publishers), The International Dissertation Abstracts, the American Doctoral Dissertations, and the P.M.L.A., I find that nothing other than this research has been done in relation to a construction of a Spanish orchestra of language.

---

[12] La Real Academia Española, Biblioteca de Autores Españoles (Madrid: Atlas, 1950), p. 3.

Nearly all the work done in Spanish is concerned with descriptions of sounds, syntactic structure, semantics, and literature, such as Amado Alonso's Estudios Linguísticos,[13] Hayward Keniston's Spanish Syntax List,[14] Emilio Alarcos Llorach's Fonología Española.[15]

In the bibliography previously-mentioned, the writer also looked for work done in English in relation to the subject of sound and music. Some English linguists have done some work related to music and sound, but none on the level of Robson's work. Among these works are David Abercrombie's Elements of General Phonetics,[16] David Abercrombie's Studies in Phonetics and Linguistics,[17] David I. Masson's Thematic Analysis of Sounds in Poetry,[18] David Littleton's Structures of Dialect in Short Stories and Poetry of Jessie Stuart.[19]

---

[13] Amado Alonso, Estudios Linguísticos (Madrid: Editorial Gredos, 1961).

[14] Hayward Keniston, Spanish Syntax List (New York: Holt, Rinehart and Winston, Inc., 1957).

[15] Emilio Alarcos Llorach, Fonología Española (Madrid: Editorial Gredos, 1954).

[16] David Abercrombie, Elements of Generals Phonetics (Chicago: Auline Publishing Company, 1967).

[17] David Abercrombie, Studies in Phonetics and Linguistics (London: Oxford University Press, 1965).

[18] David I. Masson, "Thematic Analysis of Sounds in Poetry," Essays on the Language of Literature, ed. Seymour Chatman and Samuel R. Levin (Boston: Houghton Mifflin Co., 1967).

[19] David A. Littleton, Structures of Dialect in Short Stories and Poetry of Jessie Stuart (Morehead, Kentucky: Morehead State University, July 18, 1970).

Furthermore, the writer inquired into bibliography concerning Robson's work, but as far as the writer can ascertain there is none.

PURPOSE AND SPECIFIC ELEMENTS TO BE PROVEN

The main purpose of this monograph is to construct a Spanish orchestra of language to be able to determine whether certain Spanish sounds can evoke similar responses to those evoked by the English sounds.

The writer specifies the Spanish sounds that are high-toned, medium-toned or low-toned, the sounds that are percussions, hisses; and, the other orchestral sound clusters--while taking into consideration the mechanics of their articulation. This monograph shows the similarities and differences between Spanish and English sounds, a comparison which is based mainly on their articulatory qualities.

The fact that some Spanish sounds are different in articulation from the English sounds, and that Spanish does not have some of the English sounds, changes somewhat the distribution of these sounds in the orchestra of language, but the more similar sounds will keep the same place as in Robson's orchestra.

The analyses of poems of William Butler Yeats and Fray Luis de León for these sounds shows which responses are attained by the different orchestral clusters in both languages, and it shows further whether these responses are different or similar.

ESSENTIAL DEFINITIONS

Because of the many different meanings that words carry in language, specific definitions will be given of the main concepts used in this monograph. A "phoneme" is defined as that significant speech sound that signals a change in meaning.

A "morpheme" is the smallest linguistic unit grammatically relevant; for example, the endings "-s" and "-'s" for the plural and the genitive case respectively.

A "word" is defined as the structure which carries a primary stress at least, and at least one vowel.

Some concepts used in Robson's Orchestra of Language need to be defined in order to avoid misinterpretations of such concepts. Thus, phonetic "timbre" is that distinctive quality in the sounds of words that enables us to sense and to identify the organs of speech involved in their articulation.

"Tones" represent the higher or lower vibration of vowels: they are divided into three tone categories: high, medium and low. These tones are produced in the mouth.

The average time extension, decibel loss between normal speech and the starting point of hearing and auditory sensitivity to frequency levels constitute what Robson calls "striking power."

These definitions will enable us to understand more

clearly the qualities given to sounds by Robson. An understanding of these definitions will further give the reader a view of the criteria according to which the sounds are distributed in the orchestra of language.

Chapter 2

SIGNIFICANCE OF ROBSON'S ANALYSIS OF LANGUAGE

Robson's The Orchestra of Language provides a synthesis of oral encoding, musicology, and written encoding of significant sound patterns in the English language. Written, somewhat ironically, for advertising people, the book has much value for semantic markers which rest appropriately enough on the phonology of the language.

Paul Mills, of the Barbizon Language School, John Kelley of the Bell Laboratories, Edgar Grisewood of the New York University Physics Department, Margaret Schlauch, linguist and author, and Rudolph Schramm, musicologist, synthesized their efforts to reveal how certain significant English phonemic combinations tend to evoke certain semantic associations, with the emphasis on the evocation of emotional states.

In forming this orchestra, the various vowel and consonant combinations are developed through a parallel with orchestral groups. Charts in the various appendices will reveal specifically the details of orchestration. Further, enough detailed information through charts and from the text of Robson's book itself is available to the reader who would consider the system itself in more detail. All is a matter of sounds, significant ones.

That sounds will vary in tonal, temporal, and power-

laden intensities are items of common knowledge. When one speaks of duration, he speaks of the length of time taken to articulate some element per unit time. The unit time used by Robson is that of decimal fractions of a second. In such a way, one can have some standard for measuring. That is, in combination--and in isolation--vowels and consonants--separately and together--must have some predictable difference in phonological length.

Now, one may assert, very well, that drawls and other dialect-ridden factors may vary so much as to make duration-measuring of little value. However, the majority of individuals using standard pronunciation is so substantial to make extremes of little or no consequence. As is well-pointed out, there are certain tendencies that are significant and that must be taken into account:

> There is no question that a speaker or singer can stretch out or condense. What is meant here is a natural tendency to speak so many words per minute, a normal rate. No matter how much an individual may vary his time duration deliberately, he can succeed only to the extent that he is always within the natural speaking range, with regard to duration, for example. If he exceeds the limits of the range, on purpose, he will speak only to himself.[20]

Thus, this duration is the result of the average of many phonological experiments. Now, without laboring the obvious, duration is caused by--and affected by--tone and power. Robson must be correct in asserting that monosyllabic

---

[20]Robson, op. cit., pp. 18-19.

words, for example, will have a mean time duration. That such is the case results from the fact that it takes time to articulate vowels and consonants. Where each vowel and consonant is given a striking power value in terms of decimal fractions of a second and where these are added together for each word, each word can be given an intensity or impact value.

Consider a phonetic item with a relative striking power value of 30, with a "relative mean time in seconds" value of .20 seconds. Then consider another item with a relative striking power value of 30, but with a relative mean time in second value of .10 seconds. The intensity rating for the second will be 30/10 or 300. The rating for the first will be 30/.20 or 150. All such items are developed in Robson's The Orchestra of Language.[21]

It is asserted that time is most important in language. What happens, semantically, for example, results from a difference in time. Where an adjective is used before a noun, the fact that one hears the adjective phonologically earlier than the noun often accounts for the psycholinguistical factor that one judges the whole in terms of the part--one judges the noun because he heard the adjective first.

It is true that this work will not attempt to parallel the Orchestra of Language in English and Spanish in sufficient detail to talk about psychological factors where adjectives are placed before or after nouns. However, the point is

---

[21]Robson, op. cit., see Appendix D.

being made that if there is a desire to focus on the adjective, rather than on the noun, the adjective with the greatest intensity rating will serve to have the adjective stand more completely for the noun. Robson moves from the point of time to that area of low and high sounds. Now, of course, the low is associated with a longer duration and the high with a shorter time duration. These matters are both implicit and explicit in the tables and appendices cited. Robson next moved to his orchestra itslef.

The orchestra is developed in terms of timbre, where "timbre" refers to the uniqueness of the sound:

> The timbre of words is the cast of the voice in them. This depends on the structure and physical traits of the tongue, teeth, lips, palate, vocal chords, and oral liquids that modulate the passage of breath when we talk.[22]

Robson goes on to parallel the auditory quality with the unique blends of color on the eye.

Finally, Robson reaches the point of detailing the orchestra.[23] There are eleven orchestral groups and the vowels, diphthongs, semivowels, and consonants must find their places there, and uniquely so. In the first group the high-toned vowels appear. In the second group, there are the middle-toned vowels. In the third group, there are the low-toned vowels.

Group 4 called "Noises with Overtones or Semi-Vowels." These are subdivided into "sonorous hums," "nasals," "breath

---

[22] Robson, op. cit., p. 30.

[23] See Table IV.

drops," "sonorous liquids," "breath lifts," and "roars." Group 5 consists of "Percussions," "Hisses," and "Throat Phonemes." Group 7 offers the "Straight Breath Sound;" Group 8, the "Straight Hisses;" Group 9, the "Percussions and Throat;" Group 10, "Percussion and Hiss," and Group 11, the "Straight Percussions." [24]

As the chart in Table IV will reveal, specific phonemes are included in each group, as well as words at point. When the comparison with the Spanish language is made, the orchestra will be arranged in parallel fashion. Robson points out that the sounds come from certain parts and features of man's sound reproducing apparati. As Sager indicates:

> Human voice features operate in many such ways to produce the light and the dark, the euphonious and the cacophonous, and do so with rising and falling tones, with soothing and stabbing thrusts, and with undulations with high and low crests.[25]

Support for Robson's statements as to timbre, vowels, diphthongs, semivowels and consonants is to be obtained generally, but specifically, from A.C. Gimson's An Introduction to the Pronunciation of English. Gimson's work is valuable in the area of English, and also valuable--and highly-respected--internationally. Because virtually no articles have been written specifically toward linking sound productions orchestrally-tiered in a semantic direction, it is as well

---

[24] See Table IV.

[25] E. J. Sager, "The Human Language Laboratory of Sounds," Stud. Sounds, I, 3 (1967), 83-90.

to rely on such work as Gimson's to support or reject Robson's initial assumptions.

Gimson's work offers adequate background support for all of Robson's assumptions as to timbre, duration, power, and intensity. While Gimson does not discuss these matters per se, in that he stops just short of psycholinguistical theory and applications, his Chapter 10, "Connected Speech," has adequate support for positions taken by Robson.[26]

Furthermore, Gimson's treatment of Acoustics and Auditory Aspects is relevant.[27] Sections from these areas are to be found in the appendices to this monograph.

Maruszewsky in "Psycho-biological Foundations of Speech, Word Productions, and Greater-Than-Word Production" makes the points that vowels color semantically the tones of significant speech utterances in manners physical, cerebral, and emotive. Consonants give the energy and the continuum that must result from the impulses giving rise to the entire speech-producing utterances. There are fundamental frequencies, tones and noises present. When the operations of speaking or writing are set in motion, the entire speech-producing system is composed of miniature items acting in concert on different levels, but all in terms of some unifying theme. When there is no unifying theme, one gathers, there is a personality fragmentation, or a conflict of movements originating closely in the same physical or psychological time. Each utterance,

---

[26] See Appendix B.

[27] See Appendix C.

microscopic or macroscopic,[28] has a connecting quality. This connecting quality is similar in kind and degree to that postulated by Gimson. Maruszewski's intensive treatment confirms Gimson's acoustic statements and points to Robson's emotive bent.[29] But, Robson's orchestral units now need specific comments as to their functional applications.

When a low-toned vowel combines with other orchestral groups, certain words are possible. These words, limited only by possible sound arrangements, tend to evoke meanings or emotions, or both, to the exclusion of others.[30] A fuller discussion of the specific can be found in Appendix D.

Robson's treatment of the orchestra itself[31] indicates that that which is low, huge, dark, and mysterious comes from the rear of the mouth. That which is clearer, sharper, and bright increases in proportion to the movement from rear to front. After noting that tones come from the mouth and not from the throat, at least in English, the thesis is developed that rather dramatic differences in evocation result from whether one combines with high-toned vowels, low-toned vowels, or middle-toned vowels. Some of the more significant observa-

---

[28]Microscopic refers to that less than the sentence. Macroscopic to that greater-than-the sentence.

[29]M. Maruszewski, "Psycho-biological Foundations of Speech," Rev. E.H. Lennberg, Psychol. Wychowawcza, trans., XII, 1 (1970), 7-29.

[30]See Appendix D.

[31]Ibid.

tions of Robson's work are cited. Using the "m" for a hum is most practical when the low-toned vowels are employed. Using the high-toned vowel and "m" results in less power. The most intense timbre of "n" is achieved through using high-toned vowels.

Different kinds of softness and liquidity are achieved through the various vowels and "l." Some combinations give the uplift effect, some the falling-off effect. The use of the right combination for the desired effect is what the writer or speaker should strive for, according to Robson.

If one desires that which flows rhythmically with power, the "w" in combination with the vowels is appropriate. Ultra-high tensions and explosiveness with speed and force come from the "y." Grim and shapeless roaring force comes from "r" in combination with "s" and with different vowel units.

If one desires an explosive percussion belt, there are available the /p/, /k/, and /t/. The lip drive comes from /p/, the tearing tooth quality from /t/, and the terrible hardness from /k/.

If there is a desire to have steady movement, the friction in the /s/+/t/+ different vowel groups is available. All this leads to the concept of striking power of sounds heard in speech and subsumed in decoding writing. This striking power is based on decibel intensity calculations by Harvey Fletcher.[32] The next element in the concept of the language and orchestration is that of learning by hearing relative differences.

---

[32]Robson, op. cit., p. 45.

Man is somewhat slow in realizing that we learn not so much by direct statement and direct observations of a single item, as by hearing, thinking, or feeling differences in two or more items when placed in the same class. Thus, if one desires to talk about relative striking power in any meaningful way, he might take synonyms and determine the difference not lexically but phonologically. Robson furnishes some paired words,[33] one set of which consists of the items "strike" - "hit." The meaning, lexically, carried is somewhat the same, but the intensity of "strike" to "hit" has the ratio 46/28.5, and the difference in intensity indicates, for the reader or listener, certain differences in meaning.

Without analyzing in depth, all the items in the Orchestra of Language, I indicate Robson's use of a vowel scale.[34] This scale indicates relative striking power of vowels in words. The range from low to high in articulating June-John-Jan-Jean is indicated.[35] The interesting aspect of the example cited is that one could not go--in English--from low to high in the sequence Jean-Jan-John-June. It is inevitable that Robson's work would move toward the sentence.

One of the most striking examples of the use of words in sentence utterance is the line from Macbeth: "Tomorrow and tomorrow, and tomorrow." An analysis of the orchestra and a view of the striking power of the words point or

---

[33] Robson, op. cit., p. 47.

[34] Ibid., p. 51.

[35] Ibid., pp. 51-53.

evoke a chord of bitter despair or hopelessness, achieved by the low and unvarying monotony of the sounds in sequence.[36] The remainder of the book is given over to demonstrating different kinds of sentences deliberately constructed from the assumptions made in the orchestra itself.

The next step is to consider whether or not Spanish has the kinds of sounds, as significant, that are found in Robson's "orchestra." If so, do they react uniquely with consonants to evoke certain psychological or emotive effects?

It is neither claimed nor postulated in any way that the same combinations of Spanish phonemes and English phonemes will produce the same results. But it is postulated that there are such parallel orchestral groups, that the vowels are unique in reacting with Spanish consonants, that rising and falling tones are produced, that explosiveness is produced, that monotony is produced, and that zip, fire, lethargy, and other emotive states are evoked, although not with the same phonemic combinations as found in English.

---

[36] Robson, op. cit., pp. 56-57.

Chapter 3

A SPANISH ORCHESTRA OF LANGUAGE

This chapter is aimed at developing an orchestra of language for Spanish. The lack of instruments--such as an spectograph which Robson used to measure sounds--and the lack of data about the Spanish sounds--such as mean duration of spoken phrases, of whispers and others--prevent the writer from providing an accurate information about the values of the sounds according to their striking power, tones, and time duration. Thus, the Spanish orchestra of language will be an approximation to Robson's orchestra of language which will be made taking as a basis the similarities that exist between the Spanish and English sounds, and following step by step the most important points in Robson's orchestra.

The first point considered by Robson to build his orchestra is the timbre of sounds. Robson has made his orchestra in terms of the timbre of sounds, so consider what timbre is in Spanish:

> Desde el punto de vista fisiológico, el timbre de las vocales resulta de la especial disposición que durante la producción del sonido adoptan los órganos articuladores, formando en cada caso, en la cavidad bucal, un resonador de forma y dimensiones determinadas.[37]

Timbre is considered the same for both Spanish and English.

---

[37] T. Navarro, op. cit., p. 35.

The next step taken by Robson was to distribute the sounds in an orchestra in eleven groups, in the same way as the instruments are distributed in a musical orchestra. This distribution was made according to the timbre of the sounds.

Together with the distribution of sounds in a Spanish orchestra of language, I will point out some important aspects relating the Spanish sounds.

The aspects of sounds that will be considered here will not be related to articulation, because it is not the aim of this work to describe the Spanish sounds, although their description will be provided in the Appendices E and F for the reader. The aspects the writer refers to at this point are related to classification of sounds.

The system of the Spanish vowels has been considered by at least two linguists as the essence of simplicity, namely Robert P. Stockwell and J. Donald Bowen. In Spanish there are five pure vowels: /a/, /e/, /i/, /o/, /u/ which are pronounced in the same way in any place they occur in the Spanish language. These five vowel sounds are represented phonetically in the same way they are represented orthographically. This characteristic is appropriate to which is a phonetic language.

The vowel sounds are the tones, the sounds that produce the music in the language, and for this reason Robson placed the vowels in the front of his orchestra and the consonants, that are noises, in the back:

Just as symphonic orchestra puts the stringed

> instruments and the wood winds in the front rows and the percussion devices such as drums, triangles, cymbals back in the rear rows, so the orchestra of language places the vowels that are tones in the front and the consonants that are noises in the back.[38]

Thus, the Spanish orchestra of language will also have the vowels in the front, together with diphthongs and triphthongs.

Then, the Spanish diphthongs are presented with their phonetic symbols:

Diphthongs	Phonetic Symbols	Examples
ai, ay	/ai̯/	aire, hay
au	/au̯/	causa
ei, ey	/ei̯/	deidad, ley
eu	/eu̯/	deuda
ou	/ou̯/	bou
oi, oy	/oi̯/	boina, hoy
uy	/ui̯/	muy
iu	/iu̯/	ciudad

The triphthongs in Spanish occur when the vowels "i" and "u" are in the same syllable, plus either vowels "a" or "e:"

> La presencia de los sonidos "i," "u" en una misma palabra, al principio y al fin de un grupo vocálico cuyo elemento central y predominante sea "a" o "e" da lugar a los triptongos.[39]

Thus, in Spanish there exist the following triphthongs:

iai	/jai̯/	despreciáis
iei	/jei̯/	limpiéis
uai	/wai̯/	averiguáis
uei	/wei̯/	buey

The vowel sounds with the diphthongs and triphthongs can be placed in the same way Robson distributed the vowel sounds and diphthongs in his orchestra according to high, low,

---

[38] Robson, op. cit., p. 31.

[39] T. Navarro Tomás, op. cit., p. 65.

and middle-toned vowels.

Accordingly, Group 1 will offer the high-toned vowels in Spanish: /i/, /ei/, /ai/, /iai/, /iei/, and /ui/. Group 2 consists of the middle-toned vowels which are /a/, /oi/, /e/, /uy/, /uei/, and /uai/. Group 3 consists of the low-toned vowels /o/, /u/, /au/, /eu/, /ou/.

This classification is made according to Robson's statement that high-toned vowels are the ones produced in the front of the mouth cavity; low-toned vowels are produced in the back of the mouth cavity; and, middle-toned vowels are produced in the medial part of the mouth cavity, such as the vowel /a/; and finally diphthongs are formed by low and high vowels. The four triphthongs that exist in Spanish are distributed in the high-toned vowels and the middle-toned vowels. The triphthong /iai/ is a combination of a middle vowel and two high-toned vowels, giving it the value of a high vowel because of the influence of the two /i/ sounds. Triphthong /iei/ is also in the high-toned vowels because it is formed by three high-toned vowels. Finally, the triphthongs /uai/ and /uei/ are placed in the middle-toned vowels because they are made by a low, middle, and high vowel, and a low, middle and high vowels, respectively.

In Group 4 Robson places the partial tones and the transient semivowels. The writer will determine whether a similar classification can be made in the Spanish orchestra.

The Spanish sounds /m/, /n/, /l/, and /ŋ/ have almost the same phonetic quality of the corresponding English sounds. It is important to point out, though, that the Spanish

sound /l/ does not represent the double "l" which corresponds to a different sound in Spanish. Further, the sound /n/ although it has a similar quality to the English sound /n/, does not represent the combination "ng" --because this combination is represented by two separate sounds in Spanish. The sound /n/ represents the letter "n."

Moreover, Robson places in this category the semivowel /j/. In Spanish there are two sounds similar to the English /j/: these two sounds are the semiconsonantal /j/ that represents the letter "i," and the sound /y/ whose orthographic representations are "y" and "hie."[40] The Spanish /j/ does not occur in initial position as the English sound, and /y/ although it is considered a consonant, has quality similar to the English /j/.

The sound /y/ occurs in initial position when not preceded by /n/ or /l/. The place where this sound occurs is of importance in Spanish because there is another sound which corresponds to the same spelling. Because of the place in which it occurs in the sound-combinations, /y/ has a completely different quality.

The semivowel /w/ is also in Group 4. In Spanish there are two similar sounds to the English /w/ which are /w/ and /g/. The second /w/, in Spanish, corresponds to the spellings "u" and "hu." The letter "h" in Spanish is a silent consonant. The sound /g/ has a quality similar to English /w/ when it occurs in combination with the sound /u/. Therefore, the combination /gu/ will be considered similar to the English /w/.

---

[40] See Appendices E and F.

Let us now examine the sound r/. In Spanish there are two sounds of /r/. One is a voiced-alveolar single vibrant sound whose phonetic symbol is /r/ whose articulation is quite different from the English sound /r/:

> In American English the tip of the tongue is turned back against the front palate, the tongue assumes a more concave shape, and the resulting sound, a more open, more vocalic, and longer /r/, is not even an approximation to Spanish r.[41]

Although this sound is not similar to the English /r/, both are roaring sounds; so, under this contention the writer will place this sound in the same place where the /r/, in English, is placed. This sound occurs in any position except the initial one and when the sound is not preceded by "l" or "n" and "s."

The other /r/ sound in Spanish has the phonetic symbol /r̄/ occurring in the initial position and preceded by "l," "n," or "s." This sound, produced by multiple vibrations, is the longest in Spanish.[42] This /r̄/ also carries roaring sound and a very strong one; thus, it is proper to place this sound in the category of English /r/.

Also placed in Group 4 is the sound /n/ whose written representation is "n." This sound, which does not exist in English, has a nasal quality, together with the quality of a /y/ phoneme.

In the Spanish orchestra of language there are the following sounds in Group 4: /m/, /n/, /l/, /r/, /r̄/, /j/, /y/,

---

[41] T. Navarro Tomas, op.cit., pp. 50-51.

[42] See Appendix F.

speech. The Spanish language considers the same mean as a normal loudness in conversational speech.

To build the scale of the striking power of the sounds Robson considering "th" as the lowest sound in English gave it a striking power of 1. In Spanish the sound /θ/ is also the weakest sound; this point was proven by measuring the power of the Spanish sounds in a tape recorder with a reading meter in decibels.

The rest of the English sounds have a striking power relative to that of "th," that is, Number 1.

Robson considers the highest striking power of English sounds equal to 30 for the sound /ai/. In the Spanish orchestra of language the writer will also consider 30 as the highest striking power for the sound /ai/. But the lowest striking power for the Spanish sounds is 20 for the sound /u/. Similarly, the English sound /u/ has a striking power of 20, although this sound is not the one with the lowest striking power in the English language.

The striking power of the Spanish consonants will be the same as those of the English consonants except for two sounds which have more power than their "similar" counterparts in English, the /x/ and the /r̄/.

The Spanish sound /x/ has more power because of the strong friction produced in the throat during its articulation. The consonant sound /r̄/ is a powerful sound in Spanish; when pronounced in a sound-combination, it yields a roaring effect like that of a tiger. Thus, the writer assigned a striking power of 2 to the /x/ sound and a striking power 10 to the

sound /r̄/. (There is no similar sound in English for the Spanish /r̄/.)

Tones of Sounds

The application of the tones to the Spanish orchestra will be done in a parallel fashion to the way in which the striking power was assigned; namely, the same principle applied by Robson to the English sounds will be applied to the Spanish sounds.

Vowels create tone; consonants do not:

> It's the vowels that give tone to words. Consonants and semivowels may blur tonal effects, make them decisive with the timbre in the casts of the voice, give them striking power, shift tones up and down a fraction or set tones and fix them more definitely by the device of rhyme. They do not create tone.[45]

Consonants and semivowels may influence the tone created by the vowels, but they certainly do not create tone by themselves.

Vowels carry the most significant importance in the orchestra of language because they help the writer stress the emotions he wants to convey; they also help him in achieving rhyme and style.

Let us examine how we can distribute the tones in the Spanish orchestra of language. There are some vowel sounds which do not exist in English. The writer will distribute the Spanish sounds in a similar scale from 2 to 26 as established by Robson.

---

[45]Robson, op. cit., p. 56.

Some sounds in Spanish, especially nearly all of the vowel sounds, have a higher intensity than the English sounds, a fact that may contribute to the formation of a higher tone in the vowel sounds. Let us consider the diphthong /ay/ in Spanish and English. Both languages have a similar diphthong, but in Spanish the diphthong is more tense and the glide is much longer than in English:

> The rise of the semivowel toward a higher fronter position is noticeably shorter and slower in English than in Spanish. English /ay/ can be made into a successful Spanish /ay/ by articulating more tensely throughout and by being certain that the glide is carried rapidly all the way forward and up.[46]

Thus, the Spanish sounds may be tenser and stronger than the English sounds.

Another point which is important to consider in the Spanish language is that Spanish allows the occurrence of sequence of two vowels in a word where both are strongly pronounced:

> Although sequences of two vowels rarely occur in English and when they do occur, a glottal stop must obligatorily break the sequence in Spanish such sequences are extremely frequent.[47]

This fact may contribute to the existence of a greater euphonious tone in the Spanish language.

All the Spanish sounds are distributed in the chart of tones according to the place they have in the orchestra of language; that is to say, high-toned vowels, middle-toned

---

[46] Stockwell and Bowen, op. cit., pp. 99-100.

[47] Ibid., p. 109.

and low-toned vowels. Nearly all these vowels are given the same values the English vowels have in the tones of Robson's orchestra with a few variations such as the values given to the triphthongs which do not exist in English, but all the values are given within the scale 2 and 26.

Time Duration

The final point considered in Robson's orchestra is the time duration of the sounds assessed at a speaking rate of 148 words per minute, a rate used to classify the time duration of the Spanish sounds.

The time duration of nearly all of the sounds will be taken from T. Navarro Tomas's Pronunciación Española in which Navarro gives the time duration of some Spanish sounds in sound-combinations in hundredths of a second.

In Spanish, there is no distinction between short and long vowels. Each vowel has only one time duration, one which can be slightly modified by the stress used in the vowel:

> Si se considera la duración normal de las vocales largas en otros idiomas, puede decirse que en la pronunciación ordinaria española no hay vocales propiamente largas.[48]

Thus, the Spanish vowels in normal conversation have only one characteristic. It is better to consider them "short vowels."

---

[48]Navarro, op. cit., p. 199.

T. Navarro Tomás gives some vowels with their time duration, such as:

o	-	0.14 sec.
a	-	0.12 "
e	-	0.8 "
i	-	0.12 "
u	-	0.10 "

For the determination of the time duration of the Spanish vowel sounds, the writer compared the similarities between the English vowel sounds to those of the Spanish language. Further, in order to determine the time duration of the Spanish vowel sounds, a stop watch was used to measure the vowel sounds both in isolation and in combination. Although not the most accurate method, it provides a useful approximation to the time duration of the Spanish vowel sounds.

Time duration of diphthongs and triphthongs

ai̯	-	0.21
ei̯	-	0.19
oi̯	-	0.21
iu̯	-	0.16
uy̯	-	0.18
au̯	-	0.19
eu̯	-	0.20
ou̯	-	0.22
iai̯	-	0.25
iei̯	-	0.23
uai̯	-	0.24
uei̯	-	0.22

Time duration of consonants

Consonants in Spanish are all pronounced (except for the letter "h"); therefore, Spanish does not have weaK or strong consonants. But some consonants have differences in their time duration, differences according to the position

they have in the word:

> Dichas diferencias dependen principalmente de la posición de las consonantes en el grupo fonético y de la naturaleza articulatoria de cada sonido.[49]

Thus, some sounds such as the /r/ and /r̄/ have different time durations. The /r̄/ sound is one of the longest consonants in Spanish, and /r/ is the shortest.

Consequently, the following time duration is given to the Spanish consonants:

b	-	0.7
ch	-	0.13
d	-	0.06
f	-	0.13
g	-	0.06
k	-	0.11
s	-	0.12
r	-	0.02
rr	-	0.14
y̆	-	0.07
l	-	0.08
m	-	0.13
n	-	0.10
ñ	-	0.06
p	-	0.02
/x/	-	0.11
/n̆/	-	0.10
/l̆/	-	0.18
/w/	-	0.07
/j/	-	0.06

Having determined the time duration of all sounds, it is now possible to make a scale grouping of the sounds according to their time average as Robson did in his orchestra.

The sounds are distributed in six groups according to the average of time duration in which each group can be classified. The writer may now form a chart with the time duration of the Spanish sounds similar to the one formed by Robson.

---

[49]Navarro, op. cit., p. 204.

Finally, it should be stressed that this Spanish orchestra of language is constructed according to the similarities that the Spanish sounds present in relation to the English sounds, although no sophisticated electronical instruments were used in this investigation.

Chapter 4

DETERMINATION OF SIMILAR EMOTIVE RESPONSES IN W.B. YEATS'S "THE LAKE ISLE OF INNISFREE" AND FRAY LUIS DE LEON'S "VIDA RETIRADA"

The purpose of this chapter is to determine whether similar responses, such as emotions and feelings, can be evoked by the sound-combinations formed in both English and Spanish. For this purpose, the writer has chosen two poems--one by W.B. Yeats, "The Lake Isle of Innisfree," and another one by Fray Luis de Leon, "Vida Retirada"--each based on the same subject, namely the subject of withdrawing from society to live in a quiet place. The speakers in these poems do not only express the desire of withdrawing from society just physically but also spiritually. At the same time, their longing for peace is physical as well as spiritual. Therefore, the same melancholic mood exists in both poems.

This analysis is made by using the English orchestra of language for Yeats's poem, and the Spanish orchestra of language for Fray Luis de Leon's poem.

Determination of the Tone in "The Lake Isle of Innisfree"

As it was pointed out in Chapter One, there are some poets who know how to use words to achieve effects that are entirely unique to that poem. This poem indicates this command of words allowing the speaker in the poem to create the effect he desires.

"The Lake Isle of Innisfree" has a sad, horizontal mood; therefore, the writer will determine whether the tone of the poem is also low in a horizontal way as the subject suggests.

The tone of the poem is determined by using Robson's chart. Robson indicates in his orchestra that the repetition of low vowel sounds such as /ʌ/, /u/, /ow/, /aw/, /er/, /o͝o/, /oh/, and /oo/ help to stress the low tone of the poem. The tone value of these vowels ranges between 6 and 2. Therefore, it is of great importance to determine whether the speaker in this poem uses vowel sounds to stress the tone of the poem.

First Stanza

12 23 6 12 6 2 6 6 2 2 24 6 25
I will arise and go now, and go to Innisfree
6 6 6 19 24 24 26 6 22
And a small cabin build there, of clay
6 6 6 21.
and wattles made.
12 25 2 24 12 22 26
Nine bean rows will I have there,
6 12 4 6 6 22 25
a hive for the honey bee,
6 22 6 2 23 6 25 2 22
And live alone in the bee-loud glade

Second Stanza

6 12 22 6 6 24 25
And I shall have some peace there,
4 26 6 6 12 2
for peace comes dropping slow,
6 12 6 6 21 6 6 6 12
Dropping from the veils of the morning
6 25 6 12 21 24
to where the cricket sings;
25 24 22 6 6 24 4 6
There midnight's all a glimmer, and
4 6 4 2
noon a purple glow
6 23 6 2 6 6 23 6 24
And evening full of the linnet's wings

Third Stanza

12 23 6 12 6 2 6 4 2 22
I will arise and go now, for always
12 6 22
night and day.
12 26 22 6 4 19 24 23 2 6
I hear lake water lapping with low sounds
12 6 6
by the shore;
12 12 19 2 6 2 22 2 2 6
while I stand on the roadway, or on the

21 6 22

pavements gray,

12 26 6 6 6 25 8 6

I hear it in the deep heart's core.

The speaker in the poem longs and dreams about leaving the society in which he lives. He is unhappy, and senses that he will find peace only in an isolated place such as the island of Innisfree.

The tone of this poem can be definitely considered "muted." The reader may notice the repeated use of low sounds in every line of the poem. The first stanza is full of sounds such as /ə/, /ow/, /ʌ/, /aw/, /o/, /ə:/ which tend to evoke sad, melancholic emotions. These low sounds are repeated again in the second and third stanzas and the repetition of these sounds contributes to keep the same mood in the whole poem.

The low tone of the poem is also achieved by using monosyllables containing low-toned vowels which help to give a falling tone to the poem. Notice in the first stanza the use of monosyllabic words with low-toned vowels such as "go," "to," "and," "of," "rows," "for," "the," "loud." In the second stanza Yeats uses the monosyllables "some," "slow," "all," "a," "the," "for," "of," "full," "moon," "glow," "comes," "and." In the third stanza he uses "go," "now," "for," "low," "sounds," "the," "shore," "on," "or," "it," "heart," and "core." Therefore, it can be ascertained that the use of low-toned vowels stresses depressing, melancholic feelings.

Determination of the Striking Power in "The Lake Isle of Innisfree"

Robson states that the use of striking power aids the writer in stressing the feeling he particularly wants to evoke in the reader. We already determined that the tone of the poem is low; therefore, it must be determined whether the use of striking power contributes to the tone of the poem.

A striking power number between 10 and 20 indicates an extremely weak syllable; numbers between 20 and 30 indicate an average syllable, and numbers between 30 and 40 indicate a syllable of high striking-power.

First Stanza

30	34	26	20	31	27	20	31
I	will	arise	and	go	now,	and	go
	23	24	28	36			
	to	Innisfree					
20	15	39	31	29	33	26	23
And	a	small	cabin	build	there		
	17	36	20	34	25	33	
	of	clay	and	wattles	made;		
36	31	38	34	30	31.5	26	23
Nine	bean	rows	will	I	have	there,	
	15	33.5	25	16	28	33	32
	a	hive	for	the	honey	bee,	
20	31	15	36	27	16	32	35	38
And	live	alone	in	the	bee-loud	glade	

Second Stanza

20 30 37 31.5 32 30 26 23
And I shall have some peace there,
25 30 38 39 30 35
for peace comes dropping slow,
39 30 28 16 37 17 16 40 31
Dropping from the veils of the morning
23 49.5 16 35 31 32
to where the cricket sings;
48 29 35 34 15 32 26
There midnight's all a glimmer,
20 26 15 32 36
and noon a purple glow,
20 28 31 26 17 16 29 33 35
And evening full of the linnet's wings

Third Stanza

30 34 26 20 31 27 25 34 35
I will arise and go now, for always
33 20 30
night and day
30 48.5 36 34 26 33 30 32
I hear lake water lapping with
33 33 32 16 33 23
low sounds by the shore;
31.5 30 38 32 16 38 33
While I stand on the roadway,

33 32 16 32 26 39

or on the pavements gray,

30 48.5 29 29 16 30 34.5 32 23

I hear it in the deep heart's core.

In the first stanza of the poem almost seventy-seven percent of the words has a high striking power which contributes to stress the idea of melancholy, and also the idea that the writer wants to escape from the world in which he lives and build himself a cabin in an isolated place where he can find his own self and the peace he has been seeking most of his life. The reader may notice the words containing a high striking power such as "go," "Innisfree," "small," "cabin," "build," "there," "clay," "wattles," "made," "nine," "bean," "rows," "will," "have," "hive," "honey," "bee," "live," "alone," "bee-loud," "glade."

The second stanza shows fifty-four percent of words with high striking power and the third stanza shows fifty-six per cent of the words with high striking power.

After analyzing the tone and striking power of this poem, it is stated that the poet has succeeded in stressing the feeling of frustration, and the depressive mood which he experiences. The low tone is achieved by the use of low-toned vowels, stressed by the use of clusters of consonants which contribute to the high striking power of the poem--a tool used to impress in the reader the poet's state of mind. It is necessary to determine now whether Fray Luis de Leon achieves the same effect with the sound-combinations he uses in his poem.

Fray Luis de Leon's "Vida Retirada"

Fray Luis de Leon, born in 1528 and died in 1591, belongs to one of the most important groups in the Spanish Renaissance which is "the mysticism." This movement represents the religious and spiritual literature of the XVI century.

The Spanish mysticism represents a balance between reality and idealism. Reality is sublimated by a desire for an ideal state:

> El sentimiento de la belleza terrestre se sublima por el ansia de lo absoluto, por la percepción y apetencia de Dios, creador e imagen perfecta, a la vez de esa belleza. Y la embriaguez de divinidad se humaniza por el sentimiento vivo de la realidad que el místico español, aún en los casos extremos de extasis, nunca pierde. Juntos van ardor y abandono, contemplación y activismo, ilusión y desengaño del mundo, sentimiento poético y rigor teológico, la doctrina y la experiencia personal.[50]

The mystics long for a better world in which they can achieve an ideal life, and their poetry is full of their feeling of frustration for the world, and the desire of escaping from reality.

Fray Luis de Leon's work presents this mysticism: all his poems and prose are centered around these subjects:

> a) ansia de soledad y retiro intimo en comunión la naturaleza;
> b) contemplación del orden natural - huerto, otoño, noche - y, a través de esa contemplación, escape de la realidad;
> c) anhelo de paz; y
> d) elevación hasta que llega a sentir la armonía

---

[50] Angel del Río, Historia de la Literatura Española (New York: Holt, Rinehart and Winston, 1963), pp. 251-252.

> divina, en consonancia con la calma serena de su propia alma.[51]

These points represent the feelings Fray Luis de Leon expressed in his poetry: the desire to escape from reality, the wish to live in a natural world, the desire for peace, and the need to reach harmony between the universe and divinity. His poem "Vida Retirada" is characterized by his desire to escape from reality and live in communion with nature.

Determination of the Tone in "Vida Retirada"

An analysis of the sounds for the type of tone of the poem has not been made; therefore, it is important to know whether the sound-combinations may produce the tone of the poem.

The poem contains seventeen stanzas: seven are analyzed to determine the tone of the poem. The first two stanzas express the idea of escaping from society and going to a quiet place to seek spiritual peace. The same idea is contained in Stanzas Three and Four. Stanzas Five and Six contain the idea of living by a river, and far from jealousy and hypocrisy. In Stanzas Seven and Eight he goes on with the same idea. In Stanza Nine he describes the place he would have near the river, and the description continues up to Stanza fifteen. In Stanzas Sixteen and Seventeen the author expresses the joy he would have in that ideal place where he would achieve peace.

---

[51] del Rio, op. cit., p. 266.

First Stanza

8 8 6 6 6 21 6

¡Qué descansada vida

6 8 8 2 8 8 2 6 6 2 6

la del que huye el mundanal ruïdo

21 21 8 6 8 6 21 6

y sigue la escondida

8 6 6 6 8 6 21 6

senda por donde han ido

6 6 6 6 6 8 8 8 2 6 6 21 6

los pocos sabios que en el mundo han sido!

Second Stanza

8 6 8 8 2 6 8 8 6

Que no le enturbia el pecho

8 6 6 8 6 6 8 8 8 6 6

de los soberbios grandes el estado,

21 8 6 6 6 8 6

ni del dorado techo

8 6 21 6 6 21 6 6

se admira, fabricado

8 6 6 6 6 8 6 8 2 8 6 6

del sabio Moro, en jaspes sustentado

Fifth Stanza

6 6 8 6 8 8 6 21 6
¡Oh monte, oh fuente, oh río,
6 8 8 6 8 2 6 8 22 6 6
oh secreto seguro deleitoso!
6 6 6 21 8 6 21 6
Roto casi el navío,
6 8 6 6 6 8 6 6
a vuestro almo reposo
2 6 8 6 8 8 6 8 8 6 6
huyo de aqueste mar tempestuoso.

Sixth Stanza

2 6 6 21 6 8 6
Un no rompido sueño,
2 21 6 2 6 6 8 8 21 8 8 6
un día puro, alegre, libre quiero;
6 8 6 8 8 8 6
no quiero ver el ceño
6 6 8 8 8 8 6
vanamente severo
8 6 8 6 6 8 8 6 6 6 8 21 8 6
de a quien la sangre ensalza o el dinero.

Ninth Stanza

8 6 8 8 6 6 8 6
Del monte en la ladera
6 21 6 6 8 6 2 8 6
por mi mano tengo un huerto,
8 6 6 21 6 8 6
que con la primavera
8 8 6 6 2 8 6
de bella flor cubierto
6 8 6 8 8 8 6 6 8 2 8 6
ya muestra en esperanza el fruto cierto

Sixteenth Stanza

21 8 6 21 8 6 8
Y mientras miserable-
8 8 8 8 6 6 6 6 6 6 6 6
mente se estan los otros abrasando
6 8 21 6 6 8
con sed insaciable
8 8 21 6 6 6 6
del peligroso mando,
8 21 6 6 6 6 6 6 8 8 6 6 6
tendido yo a la sombra esté cantando

Seventeenth Stanza

6 6 6 6 8 21 6
A la sombra tendido,
8 8 6 21 6 6 8 8 6 6 6 6 6
de hiedra y lauro eterno coronado,
8 6 8 6 8 6 6 6
puesto el atento oído
6 6 2 8 6 6 6 6
al son dulce, acordado,
8 8 6 6 6 8 8 8 8 6 6
del plectro sabiamente meneado.

The analysis of these seven stanzas shows very clearly the great number of low sound clusters used by the author. Nearly all of the combinations used contain the sounds /a/, /e/, /o/, and /u/. None of these sounds has a tone higher than 8. There are very few exceptions in which the author had to use vowels of high tone such as /i/ and /ei/.

Although the clusters used in Spanish differ from the clusters used in English, and even though English has more opportunities to produce more vowel combinations, the vowels in Spanish allow forming such combinations so as to produce the unique tone in the poem.

The complete version of the poem is presented in the appendices.

Determination of the Striking Power in "Vida Retirada"

The determination of the kind of striking power used in the words in this poem is ascertained by the analysis of the same stanzas which were analyzed for the determination of the tone.

Each syllable presents a sound-combination; therefore, the striking power of each syllable is considered, because, in Spanish, there are words with more than five syllables and if the value of each syllable is added, the determination of the striking power would not be accurate.

First Stanza

28 29 34 30 31 26 31
¡Qué descansada vida
33 32 28 48 30 26 30 36 34 31
la del que huye el mundanal ruïdo,
24 26 28 33 27 35 26 31
Y sigue la escondida
30 30 39 34 27 31 24 32
senda por donde han ido
36 31 34 30 37 28 28 30 26 31 31 26 32
los pocos sabios que en el mundo han sido!

Second Stanza

28 32 30 28 31 34 30 27 34
Que no le enturbia el pecho
27 36 31 35 37 42 29 30 27 31 32
de los soberbios grandes el estado,
27 32 31 36 31 28 34
ni del dorado techo
27 31 27 36 30 34 31 32
se admira, fabricado
32 30 35 32 37 28 32 29 24 31 31 32
del sabio Moro, en jaspes sustentado.

Fifth Stanza

29 35 28 29 35 28 29 63
¡oh monte, oh fuente, oh río
29 27 36 32 27 23 36 27 33 32 31
oh secreto seguro deleitoso!
39 32 31 26 30 31 55
Roto casi el navío,
28 35 40 33 32 35 31 31
a vuestro almo reposo
24 32 28 28 30 27 39 31 29 37 31
huyo de aqueste mar tempestuoso.

Sixth Stanza

23 32 42 26 32 32 33
Un no rompido sueño
23 54 22 37 28 30 36 29 35 32 37
un día puro, alegre, libre quiero;
32 32 37 35 30 27 33
no quiero ver el ceño
30 31 31 28 27 27 37
vanamente severo
27 28 35 33 33 36 28 35 30 29 30
de a quien la sangre ensalza o el
26 28 37
dinero.

Ninth Stanza

32 35 28 29 33 33 28 36
Del monte en la ladera
39 27 31 32 31 32 23 38 32
por mi mano tengo un huerto,
28 35 33 34 31 27 36
que con la primavera
27 27 33 44 23 39 32
de bella flor cubierto
31 35 38 29 27 27 39 30 30 30 32 39 32
ya muestra en esperanza el fruto cierto

Sixteenth Stanza

28 35 41 27 27 36 32
y mientras miserable -
31 28 27 27 35 36 29 42 28 38 32 31
mente se están los otros abrasando
35 30 26 30 34 32
con sed insaciable
32 27 29 40 31 34 31
del peligroso mando,
31 26 33 32 28 33 34 38 27 28 34 34 31
tendido yo a la sombra esté cantando.

Seventeenth Stanza

28 33 34 38 31 26 31
A la sombra tendido,
27 29 39 28 29 37 25 36 32 32 37 31 32
de hiedra y lauro eterno coronado ,
34 32 30 28 31 32 53 32
puesto el atento oido
33 35 27 27 28 40 31 32
al son dulce, acordado
32 35 40 30 34 31 28 28 56 38
del plectro sabiamente meneado.

The analysis of the striking power of this poem reveals that the author managed to use sound clusters which produce a very high striking power.

Moreover, I have considered the striking power of each

syllable without adding these syllables up. If the addition is made, nearly all the words show a striking power over Number 50. For example, the word "descansada" in the first line of the first stanza consists of four syllables whose striking power can be added "des/24/- can/34/- sa/30/ - da/31/." The addition of each syllable in this word shows a striking power of 124.

The high striking power of the poem is achieved by sound-combinations such as "can," "reir," "con," "don," "bios," "gran," "bia," "bri," and others. It is important to notice the striking power of the word "rio" in the first line of the fifth stanza which is 63. The high striking power is formed by the fact that both vowels are pronounced and the contribution of the high striking power of /r̄/. Nearly all words starting with /r̄/ have high striking power, such as: "roto," "reposo," and "rompido."

Clusters containing /r/ also have a high striking power, such as: "secreto," "mar," "sangre," "huerto," "sombra," and many others.

The analysis of this poem reveals that sound clusters can determine the tone of the poem, a tone which is stressed by the use of combinations with high striking power.

The analyses of both poems "The Lake Isle of Innisfree" by W.B. Yeats and "Vida Retirada" by Fray Luis de Leon have determined that the poems which deal with the same subject can evoke similar responses, and can be stressed by the unique orchestral clusters used in both languages.

This linguistic analysis for the determination of the tone of poems is one approach and apparently a useful method with practical application.

Chapter 5

SUMMARY

The construction of the Spanish orchestra of language based on Robson's orchestra and the analysis of the sounds in both languages--English and Spanish--reveal: that there exist similar sounds in both languages, as well as different ones.

The analysis of the vowel sounds in both languages shows that in Spanish there exist only five vowels, as opposed to the seven basic vowels in English. These five vowels have a consistent pronunciation in any combination where they occur; whereas the English vowels change their pronunciation--this change depends on the position in which they occur. For example, vowel "a" in Spanish is always /a/, while English /a/ may be /ʌ/, /a:/, /æ/, and /ei/.

Further, in Spanish the diphthongs are formed by two distinct vowels; whereas, in English, one vowel may represent a diphthong. For example, in English the vowel "o" may be the diphthong /ou/, but in Spanish this diphthong must be represented by vowels "o" and "u," that is to say, combination "ou" which is pronounced /ou/. On the other hand, the vowel "o" is always pronounced /o/ in Spanish.

Furthermore, the analysis of the consonant sounds in both languages indicates that although there are similar counterparts, virtually all of them differ in their point of articulation. Thus, the English /d/, for example, is an alveolar

plosive, while, in Spanish the /d/ sound can be dental plosive or interdental fricative. The English /tʃ/ and /dʒ/ are both fricative palato alveolar, and their Spanish counterparts /ĉ/ and /l̑/ are affricate palatal and lateral palatal, respectively. English /t/ is plosive alveolar, while Spanish /t/ is a plosive interdental. English /ð/ and Spanish /d̂/ are both fricatives but the English /ð/ is dental and the Spanish sound in interdental.

On the other hand, there exist in Spanish two varieties of consonant "r:" /r/ and /r̄/. The /r/ sound is the most similar one to the English sound /r/ although their articulations are somewhat different; the Spanish /r/ is a vibrant alveolar, while the English /r/ is a retroflex or semivowel post-alveolar. The /r̄/ sound is the strongest sound of Spanish and its quality is completely different to the English /r/. Actually, these sounds can be compared from only one point of view--the fact that they are both roaring sounds.

Additionally, there are some sounds in English that do not exist in Spanish, and sounds existing in Spanish which are non-existent in English. Thus, English sounds /ʃ̄ʃ/ and /z/ do not exist in Spanish. On the other hand, Spanish /n̑/, represented in the written language by "n," does not exist in English. Finally, it is important to point out that Spanish "b" and "v" are represented by only one sound--the /b/. There exists no distinction between "b" and "v"; therefore, sound /v/ does not exist in Spanish.

The construction of the Spanish orchestra of language is made according to the similarities of the sounds under

existing-immediate conditions in both languages. Those sounds that do not exist in English are distributed in the Spanish orchestra of language according to the kind of sounds they are--Hisses, Percussions, Liquids or Nasal sounds. Also, their manner of articulation is considered for this distribution. Thus, Spanish /n̰/ is placed among the nasals and semivowels because of its nasal quality and also because it is produced with a glide that makes it similar to a semivowel; its pronunciation is similar to the English combination /ny/.

Furthermore, the tones and the striking power given to the Spanish sounds are also assigned according to their similarities with the English sounds. At the same time, a tape recorder with a reading meter and a stop watch were used to compare the tone and striking power of sounds of both English and Spanish. (This procedure is followed in order to obtain the most accurate measurements possible).

Finally, in order to determine whether similar responses can be evoked by the sound-combinations existing in both languages, two poems are analyzed: W.B. Yeats's "The Lake Isle of Innisfree," and Fray Luis de Leon's "Vida Retirada."

In the analyses of these poems, it is found that both poets stress the depressing, melancholic mood of their poems through the use of low-toned vowels. They emphasize this low tone by using consonant clusters with high striking power.

This analysis shows that even though both languages have their unique orchestral clusters, it is possible for the poets to make use of these unique sound-combinations to stress particular emotions. Thus, similar responses can be evoked in Spanish

English by making use of the orchestral clusters which each language allows.

Both poets make use of the low-toned vowels, existing in each language, to stress the melancholic mood of the poems. Thus, Yeats frequently uses the vowel sound /ə/ in most of the lines of the poem analyzed. In Stanza One, Line One, he uses the sound /ə/ four times in words such as "and," "arise," and "Innisfree!" In Line Two of the same stanza, the same sound is used six times in monisyllabic words. Further, Yeats uses other low sounds such as /aw/ in Line Two, Stanza One in the words "small," and "wattles." The sound /aw/ is also used in line two of the second stanza in the words "dropping," and "morning." Throughout the poem, Yeats also uses sounds such as /ow/, /oo/, and /ʌ/.

On the other hand, Fray Luis de León's poem shows a very low tone by the extensive use the author makes of the low-toned vowels. For example, the sound /a/ appears four times in Line One, Stanza One of "Vida Retirada." In Line Three, Stanza One, de León uses the sound /u/ in the words "huye," mundanal," and "ruido." In the same stanza, Line Four, the sound /o/ is repeated three times in the words "por," "donde," and "ido."

It is interesting to notice that nearly all the consonants used by both poets carry striking powers of 2 or 3, such as /n/, /m/, /g/, /k/, /t/, /p/, /s/ and /b/. Also both poets use words containing the consonant "r" which is the sound with the highest striking power in both languages. Yeats uses the words "arise" and "Innisfree" in the first stanza, "dropping,"

"morning," "cricket" in the second stanza, and "arise" and "roadway" in the third stanza. Fray Luis de León uses the words "ruido" in the first stanza, "enturbia," "soberbios," "dorado," "admira," "fabricado," and "Moro" in the second stanza, "río," "secreto," "reposo," "mar" in the fifth stanza, and many other words with these sounds in the other stanzas. Therefore, it can be stated that the consonants clusters used in both languages are similar in relation to their striking-power values which are generally of 2 or 3.

The analysis made on these two poems has contributed to the determination of the existence of similar responses that can be evoked by the sound-combinations used in Spanish and English. Therefore, an analysis of this kind may prove to be a valid method to determine the artistic techniques used by the poets.

BIBLIOGRAPHY

BIBLIOGRAPHY

Abercrombie, David. Elements of General Phonetics. Chicago: Auline Publishing Company, 1967, xix, 203 pp.

Abercrombie, David. Studies in Phonetics and Linguistics. London: Oxford University Press, 1965, i, 15 pp.

Alarcos Llorach, Emilio. Fonología Española. Madrid: Editorial Gredos, 1954,

Alonso, Amado. Estudios Linguísticos. Madrid: Editorial Gredos, 1961, x, 286 pp.

Anderson, George K., and William E. Buckler. The Literature of England. Glenview, Illinois: Scott, Foresman and Co., 1968, v, 1861 pp.

del Río, Angel. Historia de la Literatura Española. New York: Holt, Rinehart and Winston, 1963, xxii, 488 pp.

Gimson, A.C. An Introduction to the Pronunciation of English. London: Edward Arnold Ltd., 1962, ix, 294 pp.

Keniston, Hayward. Spanish Syntax List. New York: Holt, Rinehart and Winston, 1957, xv, 278 pp.

Littleton, David A. Structures of Dialect in Short Stories and Poetry of Jessie Stuart. Morehead, Kentucky: Morehead State University, 1970, vii, 85 pp.

Maruszewski, M. "Psycho-biological Foundations of Speech," Rev. E.H. Lennberg, Psychol. Wychowawcza. trans., XII, 1 (1970), 7-29.

Masson, David I. "Thematic Analysis of Sounds in Poetry," Essays on the Language of Literature, ed. Seymour Chatman and Samuel R. Levin. Boston: Houghton Mifflin Co., 1967, viii, 450 pp.

Masson, David I. "Vowel and Consonant Patterns in Poetry," Essays on the Language of Literature, ed. Seymour Chatman and Samuel R. Levin. Boston: Houghton Mifflin Co., 1967, viii, 450 pp.

Menéndez Pidal, Ramón. Manual de Gramática Histórica Española. Madrid: Espasa-Calpe, S.A., 1941, vii, 369 pp.

Navarro Tomás, Tomás and Aurelio M. Espinosa. A Primer of Spanish Pronunciation. Chicago: Benj. H. Sanborn & Co., 1926, xv, 128 pp.

Navarro Tomás, Tomás. Manual de la Pronunciación Española. Madrid: A.G. Fénix, 1959, xi, 326 pp.

Oras, Ants. "Spenser and Milton: Some Parallels and Contrasts in the Handling of Sound," Essays on the Language of Literature, ed. Seymour Chatman and Samuel R. Levin. Boston: Houghton Mifflin Co., 1967, viii, 450 pp.

Real Academia Española. Biblioteca de Autores Españoles. Madrid: Atlas, 1950, xvi, 491 pp.

Robson, Ernest M. The Orchestra of Language. New York: Thomas Yoseloff, 1959, xii, 206 pp.

Sager, E.J. "The Human Language Laboratory of Sounds," Stud. Sounds, I, 3 (1967), 83-90.

Steel, Bryan. "Contrasting Approaches to Spanish Lexicography," Hispania, LVI (March, 1971), 46-52.

Stockwell, Robert P. and J. Donald Bowen. The Sounds of English and Spanish. Chicago: The University of Chicago Press, 1966, ix, 168 pp.

Unterecker, John. A Reader's Guide to William Butler Yeats. New York: The Monday Press, 1959, iii, 310 pp.

APPENDICES

APPENDIX A

GIMSON'S DESCRIPTION OF THE ENGLISH CONSONANTS

8.01 The Distinctive Consonants

It is possible to abstract from a continuous utterance of English by means of a process of commutation (see § 5.3) twenty-four distinctive units which are consonantal both in terms of their function (i.e. they tend to be non-central or marginal in the syllable—see § 5.5) and also, in the majority of cases, in terms of their phonetic nature (i.e. they have, at least in some of their realizations, articulations involving the obstructions or narrowings which produce, acoustically, a noise component—see § 4.2).

These twenty-four consonantal phonemes are classified in the table below in two general categories :—

A. Those articulations in which there is a total closure or a stricture causing friction, both groups being typically associated with a noise component; in this class there is a distinctive opposition between fortis and lenis types.

B. Those articulations in which there is only a partial closure or an unimpeded oral or nasal escape of air; such articulations, typically voiced and frequently frictionless (without a noise component) may share many phonetic characteristics with vowels.

	Bilabial.	*Labio-Dental.*	*Dental.*	*Alveolar.*	*Post-Alveolar.*	*Palato-Alveolar.*	*Palatal.*	*Velar.*	*Glottal.*
A. *Plosive*	p, b			t, d				k, g	
Affricate					(tr, dr)	tʃ, dʒ			
Fricative	(ʍ)	f, v	θ, ð	s, z		ʃ, ʒ			h
B. *Nasal*	m			n				ŋ	
Lateral				l					
Frictionless Continuant or glides (semi-vowels)	w				r		j		

Note

(1) In some types of RP it may be necessary to include the labio-velar fortis voiceless fricative [ʍ] as a phoneme.

(2) In practical teaching it may also be convenient to treat /tr/ and /dr/ as distinctive affricates as well as /tʃ/ and /dʒ/ (see § 8.10).

TABLE I

PHONETIC FEATURES OF THE PLOSIVE CONSONANTS

(3) The glottal stop [ʔ] has been excluded, since it is not phonemically distinctive in RP; its use as a reinforcement for vowels and its allophonic association with /p, t, k/ will be treated in § 8.09.

It will be seen from the arrangement above that :—

(*a*) the plosive and nasal phonemes fall into three contrastive groups as far as the place of articulation is concerned, i.e. bilabial, alveolar, and velar;

(*b*) the affricate, lateral, and /r/ phonemes have an alveolar basis;

(*c*) the fricatives have five areas of articulation, i.e. labio-dental, dental, alveolar, palato-alveolar, and glottal.

These basic areas of articulation, convenient for labelling the phoneme, may be extended in the various allophonic realizations of the phoneme, but for any particular situation in a context the number of oppositions involving the place of articulation will remain unchanged; thus, the allophones of /t/ may be dental or post-alveolar, or the allophones of /k/ palatal, without constituting additional distinctive areas of articulation, since such variants are conditioned by the context.

Class A

Consonants Involving Closure or Stricture

8.02 Plosives

The complete articulation of a pulmonic egressive plosive, or stop, consonant consists of three stages :—

(1) the *closing* stage, during which the articulating organs move together in order to form the obstruction; in this stage, there is often an *on-glide* or *transition* audible in a preceding sound segment and visible in an acoustic analysis as a characteristic curve of the formants of the preceding sound;

(2) the *hold* or *compression* stage, during which lung action compresses the air behind the closure; this stage may or may not be accompanied by voice, i.e. vibration of the vocal cords;

(3) the *release* or *explosion* stage, during which the organs forming the obstruction part rapidly, allowing the compressed air to escape abruptly; if stage (2) is voiced, the vocal cord vibration may continue in stage (3); if stage (2) is voiceless, stage (3) may also be voiceless (aspiration) before silence or before the onset of voice (as for a following vowel), or stage (3) may coincide with the onset of vocal cord vibration, as when a voiceless plosive is followed without intervening aspiration by a vowel; again, an *off-glide* or transition associates the plosive with a following sound.

Since a condition of plosive articulation is that the whole of the speech tract behind the primary closure should form a chamber sealed to the escape of air, and since the primary closures for the English plosives are normally made in the oral cavity, the soft palate must be held in its raised position for at least the first two stages of these articulations.[1]

8.03 The Significant Phonetic Features of English Plosives

The RP plosive phonemes comprise three pairs: /p, b/; /t, d/; /k, g/. The following words illustrate oppositions in word initial, medial, and final positions :—

	/p/	/b/	/t/	/d/	/k/	/g/
Initial .	pole	bowl	toll	dole	coal	goal
Medial .	riper	—	writer	rider	—	—
	—	—	bitter	bidder	bicker	bigger
	caper	caber	cater	—	—	—
	—	rubber	—	rudder	—	rugger
	lopping	lobbing	—	—	locking	logging
Final .	rip	rib	writ	rid	rick	rig

These oppositions may be realized by means of one or several of the following phonetic features :—

(1) *Place of articulation.*—/p, b/, bilabial; /t, d/, alveolar; /k, g/, velar.

(2) *Force of articulation.*—/p, t, k/ tend to be pronounced with more muscular energy and a stronger breath effort than /b, d, g/; the former are known as relatively strong or *fortis*, the latter as relatively weak or *lenis*.

(3) *Aspiration.*—The fortis series /p, t, k/, when initial in an accented syllable, are usually accompanied by aspiration, i.e. there is a voiceless interval consisting of strongly expelled breath between the release of the plosive and the onset of a following vowel, e.g. *pin, tin, kin* ['pʰɪn, 'tʰɪn, 'kʰɪn]. When /l, r, w, j/ follow /p, t, k/ in such positions, the aspiration is manifested in the devoicing of /l, r, w, j/, e.g. in *please, pray, try, clean, twice, quick, pew, tune, queue*; some devoicing may also occur in relatively unaccented situations, e.g. *apricot, atlas, applicant, heckler, buckram, vacuum*, etc. In other positions, i.e. preceding a vowel in an unaccented syllable and finally, such aspiration as may occur is relatively weak, e.g. /p/ in *polite, lip*. When /s/ precedes /p, t, k/ initially in a syllable, there is practically no aspiration, even when the syllable carries a strong accent, cf. *pin* ['pʰɪn] and *spin* ['spɪn] (see

[1] When a nasal consonant precedes a plosive, the soft palate is raised in the first stage of the plosive.

TABLE II

LENGTH OF SOUNDS PRECEDING

PLOSIVE CONSONANTS

§ 5.35).[1] In final positions, i.e. preceding silence, /p, t, k/ may have no audible release (see § 8.04 (1)). The lenis series /b, d, g/ is not normally aspirated (see, however, the next section).

(4) *Voicing.*—The lenis series /b, d, g/ may have full voice during their second stage when they occur in positions between voiced sounds, e.g. in *labour, leader, eager, rub out, read it, egg and . . ., to be, to do, to go.* In initial and especially in final positions, i.e. following or preceding silence, /b, d, g/, while remaining lenis, may be only partially voiced or completely voiceless, e.g. in *bill, done, game, cub, lid, bag,* /b, d, g/ may be realized, initially, with vocal cord vibration beginning only in the last portion of the second stage and, finally, as [b̥, d̥, g̊]. It is, therefore, unusual in normal speech for the release stage of /b, d, g/ in final positions to be accompanied by a voiced off-glide [ə]; with complete devoicing, any audible release will be accompanied by weak aspiration. The fortis series /p, t, k/ is not voiced.

(5) *Length of preceding sounds.*—When the RP plosives occur finally in a syllable, their value is determined largely (since the voicing factor is not strongly operative) by the length of the syllable which they close. It is a feature of RP that syllables closed by fortis consonants are considerably shorter than those which are open, or closed by a lenis consonant. We have seen in the chapter on vowels that this variation of length is particularly noticeable when the syllable contains a 'long' vowel or diphthong, cf. the fully long vowels or diphthongs in *robe, heard, league* (closed by lenis /b, d, g/) with the reduced values in *rope, hurt, leak* (closed by fortis /p, t, k/). Preceding consonants, notably /l, n, m/, are also shortened by a following /p, t/ or /k/, especially when the consonants are themselves preceded by a short vowel, e.g. compare the relatively long /l/ in *killed, Elbe,* /n/ in *wand,* and /m/ in *symbol* with the reduced varieties in *kilt, help, want, simple.* A phonemic transcription of *rope, robe,* as /rəʊp, rəʊb/ is, therefore, to be interpreted as indicating that the words are distinguished not only or even primarily by a difference of the final consonant, but rather by a complex of quantitative and qualitative contrasts extending over the greater part of the word. The same effect of reduction also operates when /p,

[1] Accented /p, t, k/ also show some loss of aspiration as a result of a preceding /s/ which is not in the same syllable, e.g. /k/ in *discussed* is only very weakly aspirated compared with the /k/ of *custard,* so that *discussed* may be distinguished from *disgust* only by the fortis nature of /k/. Other pre-stop consonants (especially fortis fricatives) appear to exert a certain absorption-effect upon the aspiration of initial /p, t, k/ in a following syllable or word—compare the degree of aspiration of the accented fortis plosives in *half-past, push past,* and *go past; brief talk, finish talking,* and *no talking; rough coat, fresh coat,* and *two coats.*

t, k/ occur medially in a word, cf. the length of /aɪ/ in *rider*, *writer*, although in this situation a more strongly contrastive voicing feature is likely to be present in /b, d, g/.

Summary.—The RP plosives may, therefore, be said to be distinguished :—

(*a*) by means of a three-term series in respect of place of articulation—bilabial v. alveolar v. velar ;

(*b*) at each point of articulation by the following phonetic features according to the situation (taking the bilabial oppositions as examples which are valid for the other two places of articulation) :—

Initial in accented syllable—/p/ voiceless fortis aspirated v. /b/ partially voiced or voiceless lenis unaspirated, e.g. *pole* v. *bowl*.

Medial, following accented syllable—/p/ voiceless fortis weakly aspirated v. /b/ voiced lenis unaspirated, e.g. *rapid* v. *rabid*.

Finally—/p/ voiceless fortis weakly aspirated, if released, reducing length of preceding sounds v. /b/ voiceless (or voiced only at the beginning of the compression stage) lenis weakly aspirated, if released, preceding sounds retaining relatively full length, e.g. *rope* v. *robe*.

It is clear that, initially in accented syllables, /p, t, k/ and /b, d, g/ are distinguished by the listener mainly through the presence or absence of aspiration, rather than through the presence or absence of voice.

Foreign learners are, therefore, advised to pay particular attention to the aspiration of /p, t, k/ when these phonemes occur initially in an accented syllable. If a word such as *pin* is pronounced [pɪn], instead of [pʰɪn], there is the danger that the English listener may understand *bin*, since he interprets lack of aspiration as a mark of the lenis /b/. The danger is particularly great for speakers of those languages, e.g. many in the Romance and Slav groups, where the opposition between lenis and fortis stops relies purely upon presence or absence of voice. (The aspiration distinction between /p, t, k/ and /b, d, g/ should also be retained, when /p, t, k/ are followed by /l, r, j, w/, by the devoicing of these latter, e.g. compare *plight*, *try*, *crate*, *tune*, *twelve*, with *blight*, *dry*, *great*, *dune*, *dwell*.) Such speakers should also avoid excessive voicing of the lenis series /b, d, g/, especially in final positions. On the other hand, speakers of some languages, e.g. of the Germanic type, tend to neutralize the /p/-/b/ kind of opposition in final positions, using a fortis plosive for both. It must be remembered that /b, d, g/ finally, though they may not be voiced, remain weak as compared with /p, t, k/, the preceding sounds retaining full length.

TABLE III

ACCOUSTIC FEATURES OF THE PLOSIVES

8.04 Acoustic Features of English Plosives

Perceptual cues, capable of being expressed in acoustic terms, may be provided by all three stages of plosive articulations, so that it is possible to distinguish : (1) plosives from other consonants, (2) /p, t, k/ from /b, d, g/, (3) the bilabial, alveolar, and velar types.

(1) Plosives differ from other consonants mainly in the stage corresponding to the articulatory ' hold '. This part of the consonant is generally characterized acoustically by a perceptible period of silence throughout the whole spectrum or, in the voiced /b, d, g/, an absence of energy except at a low frequency.

(2) /b, d, g/ may be distinguished from /p, t, k/ by means of a low frequency component present in the former, i.e. voice. Moreover, there is likely to be a marked rising bend of F1 of the adjacent vowel in the case of /b, d, g/, which is not as marked in the case of /p, t, k/. However, as we have seen, /b, d, g/ may often be voiceless, in which case they are distinguished from /p, t, k/—initially, by the comparatively weak burst of noise associated with the release stage and by the absence of the gap (aspiration) characterizing /p, t, k/; finally, by their influence on the duration of the preceding sounds; medially, by the longer closure period (absence of energy) required for /p, t, k/.[1]

(3) Cues to distinction as between the bilabial, alveolar, and velar plosives are provided, in some measure, by the noise burst related to their release stage. (Effective recognition is stimulated if the noise has a frequency span of some 480 cps and a duration of 15 msecs.) The actual situation of the burst of noise for each plosive category may, however, depend on the nature of the following sound. Thus, though the alveolars have a typical noise burst in the range of 3,000 to 4,000 cps, the bilabials may be characterized by a noise burst in the region of 360 cps in the vicinity of all vowels, but considerably higher frequency bursts (e.g. in the region of 2,000 cps) may produce an impression of a bilabial plosive when in the vicinity of the closer front and back vowels. Again, the situation of the release burst recognized as velar will depend very much upon the nature of the following vowel. It has been found that, before /u:/, velar recognition is achieved with a burst in the region of 700 cps; the frequency of the characteristic burst seems to rise throughout the series /u:/-/ɑ:/-/i:/, following the rising frequency of F2 of the vowels, and reaches about 3,000 cps for a

[1] See Leigh Lisker, ' Closure Duration and the Intervocalic Voice-Voiceless Distinction in English,' *Language*, vol. 33, no. 1, 1957, in which it is shown that intervocalic /b/ has an average duration of 75 msecs compared with an average of 120 msecs for an intervocalic /p/.

vowel of the type *C* [e], although in the case of the front vowels there is a second burst stimulus possible at considerably lower frequencies (matching to some extent the varying F1 of the vowels). The burst cues for velar plosives tend to be the inverse of those for bilabials in the range 300–2,000 cps.

A particular curvature (or transition) of F2 of the adjacent vowel also provides a primary cue for the distinction between the three types of English plosive articulation. Thus, in the case of the bilabials, F2 of a following vowel will typically curve upwards to its steady state and that of a preceding vowel downwards from its steady state (a *minus* transition); the alveolars exhibit small minus transitions in F2 of an adjacent /i:/ vowel and large plus transitions in F2 with an adjacent /u:/; the characteristic transitions associated with the velars seem to be largely the inverse of those for the alveolars, F2 having plus transitions but the greatest plus transition being related to an open vowel of the /ɑ:/ type. Such considerable variation in cues for velar plosives, both in the release burst and in the transitions, would appear to reflect the extensive articulatory variation associated with the English velar plosives (see § 8.08).

Any plosive + vowel or vowel + plosive sequence is, therefore, characterized by the curving of F2 of the vowel in a typical direction, as though from, or to, a fixed point in the spectrum known as the ' locus '. This point is defined experimentally when several vowels are employed. The transition does not, however, extend from the steady state to the locus, but merely points in the direction of the latter. It would appear from experiments with synthesized speech that the best recognition is achieved if the first half of the transition from plosive to vowel consists of silence; if the curve is extended too far, glides of the type /j, w/ may be perceived. Recent research suggests the possibility that the transition of F2 (and perhaps of F3) in the vowel is related to the *place* of articulation, whereas the nature of F1 is related to the *manner* of articulation.[1]

TABLE IV

ROBSON'S ORCHESTRA OF LANGUAGE

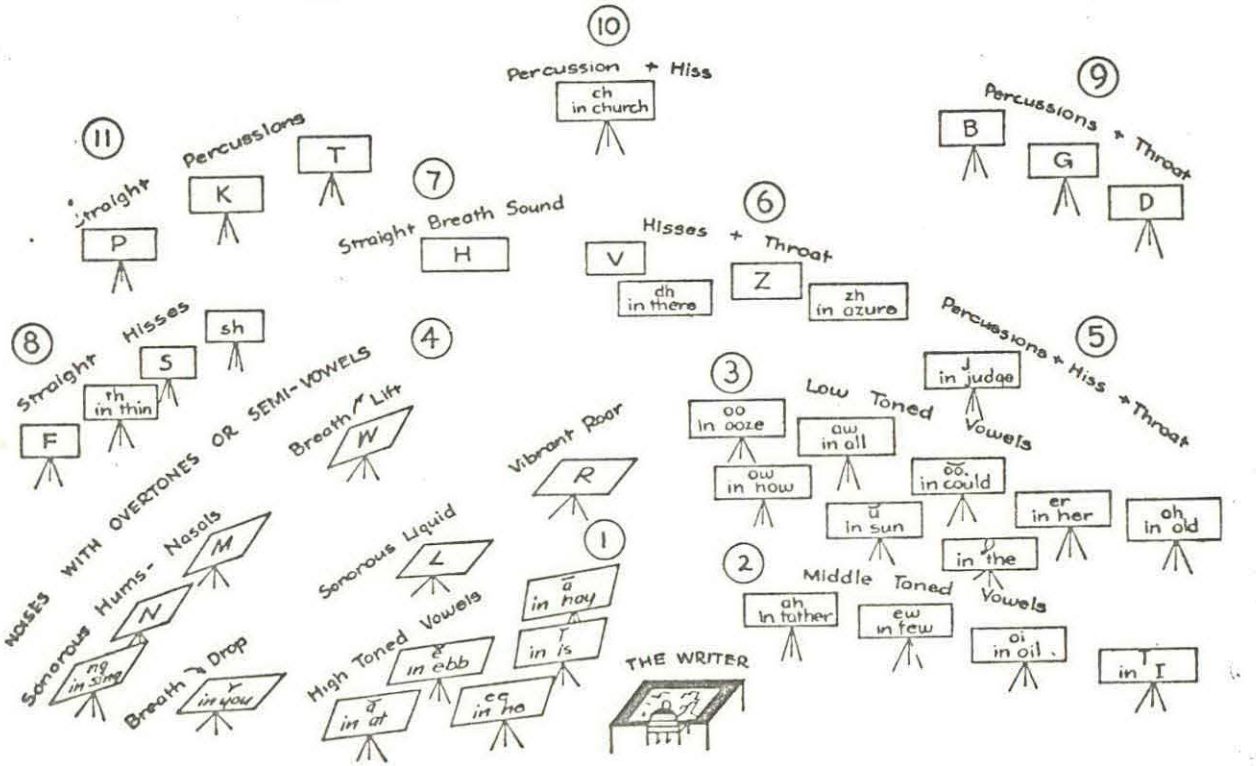


Figure 2

APPENDIX B

GIMSON'S ANALYSIS OF CONNECTED SPEECH

10.01 Accent

Connected speech, i.e. an utterance consisting of more than one word, exhibits features of accentuation that are in many ways comparable with those found in the polysyllabic word. Thus the character of a connected utterance may be said to be determined both by a changing pattern of successive qualities and quantities and also by the relationship of its parts, i.e. of the words composing the continuum. Some parts of the connected utterance will be made to stand out from their environment, in the same way that certain syllables of a polysyllabic word are more prominent than their neighbours; in both cases, accentuation has a contrastive function. Accentuation in connected speech differs, however, from the usual case of a polysyllable in that the situation of the accent in connected speech is determined largely by the meaning which the utterance is intended to convey. The meaningful, distinctive, function of accent in words is, on the other hand, of restricted application (§ 9.05). The following selected examples illustrate some of the similarities and differences of accentual patterns in the word and in connected speech :—

She can may be ˎ ˳ /ˈʃiː ˌkæn/ or . ˎ /ʃɪ ˈkæn/ (cf. the varying patterns of *insult* as verb or noun),
but not ˎ . /ˈʃiː kən/, as in *beacon*.

She can go may be ˎ . . (cf. *telephone*)
. ˎ ˳ (cf. *tobacco*)
. . ˎ or . . ˎ (cf. *cigarette*)
but not ˎ . . (cf. *bachelor*).

How can she? may be ˎ . . (cf. *quantity*)
. ˎ . (cf. *uncertain*)
but not . ˎ . (cf. *encounter*).

How can she do it? may be . . . ˎ . (cf. *interdependence*)
. . ˎ . . (cf. *aristocracy*)
but not . ˎ . . . (cf. *administrative*).

We put the case in the hall

may be . ● . ● . . ↘
↘ . . ● . . ●
. ↘ . ● . . ●
. ● . ↘ . . ●
. ↘ . ↘ . . ↘
. ● . ● ↘ . ●, etc.,
but not . ● . . . ↘ . (cf. *industrialization*)
with no accent on *case* and *hall*.

He understands the importance of practising—where *understands*, *importance*, *practising*, must have an accent of some kind on their appropriate syllables, whether or not *he*, for instance, is accented.

Such examples suggest that :—

(1) Although accentual patterns of connected speech are freer than those of the word and are largely determined by the meaning to be conveyed, some words are predisposed by their function in the language to receive accent. These *content* or *lexical* words are typically main verbs, adverbs, nouns, adjectives, demonstrative pronouns, etc. Other categories of words, such as auxiliary verbs, conjunctions, prepositions, pronouns, relative pronouns and articles (*form* or *grammatical* words), are more likely to be unaccented, although they, too, may be exceptionally accented if the meaning requires it.

The examples given above illustrate the freedom of accentual patterning in utterances taken without a context. But the meaning of any utterance is largely conditioned by the situation and context in which it occurs. Thus, it must be expected that the freedom of accentual patterning of the utterance and, in particular, of the situation of the primary (tonic) accent will be considerably curtailed by the constraints imposed by the contextual environment. In the case of an opening remark, or when a new topic is introduced into a conversation, there is greater scope for variations of meaning pointed by accentuation. Or again, it happens that some accentual freedom is possible in responses ; thus, in response to the statement *She came last week* (. ● ● ↘), an incredulous reaction *Last week* ! might have the pattern ↘ ○ [1] (i.e. 'Wasn't it the week before ? ') or ● ↘ (i.e. 'Don't you mean last month ? ') ; or, in response to *What was the weather like ?*, the reply *It rained every day* might have the pattern . ● ● . ↘ (emphasizing the continuous nature of the rain) or . ↘ ○ . ● (where the fact of

[1] Note that the sign ↘ here shows the place of the primary, tonic accent. It may be marked by the onset of a falling or, as here, a rising change of pitch direction.

raining is emphasized). But, in the following dialogue, such constraints are placed upon the accentuation, both by the context and by the nature of the content words, that little variation is possible (those words likely to be most strongly accented being printed in italic) :—

'Did you have a good *holiday*?'

'*Yes, very* good.'

'Was the *weather* all right?'

'It was *fine* for the *first* part, but for the *rest* of the time it was pretty *mixed*. We *enjoyed* ourselves though. We had the *car*, so we were able to do some *sight-seeing*, when it was too *wet* to go on the *beach*.'

(2) Many monosyllabic form words are subject to qualitative variation according to whether they receive the accent or not. Nevertheless, weakening of an unaccented form word may be inhibited by its position, e.g. *can* (/kən/ or /kn/ in unaccented initial and medial positions in the utterance, but /kæn/ in an unaccented final position). On the other hand, monosyllabic content words in relatively weakly accented positions (adjacent to a strong accent) retain their full vowel value and the prominence associated with it, e.g. *how* in *How can she?* (◦ ◝ .); or again, in *We put the case in the hall, case* and *hall* will always receive a secondary accent, wherever the primary accent is concentrated, because of the qualitative prominence of their full, strong vowels (and, frequently for the rhythmic beat which they carry), whereas *put* with its relatively weak vowel /ʊ/ may lose all accentual prominence, e.g. in the second pattern given for this sentence on p. 235.

(3) More than one word in an utterance may receive a primary (tonic, nuclear) accent. A deliberate, emphatic, or excited style of speaking often exhibits a proliferation of primary, tonic accents; a more rapid and matter-of-fact delivery is likely to show fewer primary and more secondary accents.

(4) Polysyllabic words, containing one accented syllable when said in isolation, carry an accent (primary or secondary) on that syllable in connected speech.

(5) In an extended dialogue in normal conversational style, the number of weak syllables (unaccented) tends to exceed that of those carrying an accent (primary or secondary).

10.02 Manifestation of Accent in Connected Speech

The same factors which we saw to be responsible for achieving a pattern of prominences in the word, i.e. variations of pitch, length, stress, and quality, contribute to the manifestation of the accented

parts of connected speech. Just as in the case of the word, it is necessary to distinguish those factors which are significant realities only for the speaker from those which, being perceived and interpreted by the listener, are significant in the function of speech as an act of communication.

10.03 Stress and Rhythm

The speaker is aware of the occurrence in the utterance of a number of strong stresses or beats corresponding to those parts of the utterance to which he wishes to attach particular accentual meaning and on which he expends relatively great articulatory energy; the remaining words or syllables are weakly and rapidly articulated. The number of syllables stressed by the speaker depends largely upon the nature of the words composing the utterance. Thus, an utterance containing a high percentage of content words is likely to receive more stresses than one with the same number of syllables but a higher proportion of form words: compare, for example, two sentences of eight syllables :—

The first six have all won a prize— . • • . • • . ↘
There were prizes for six of them— . . • . . ↘ . .

Stress of this kind constitutes a reality for the speaker, but is not, by itself, an efficient means of communicating accent in connected speech.

Nevertheless, the syllables uttered with the greatest stress constitute, for the speaker, hubs with which unstressed syllables will be associated to form rhythmic groups.[1] It is a feature of English that the utterance is delivered as a series of close-knit rhythmic groups, which override in importance on the phonetic level the significance of the word on the linguistic level. Compare the rhythmic groups, based on a strongly accented syllable, which are likely to occur in the following sentences :

They couldn't have | chosen a | better | time for their | holiday
. • . . | • . . | • . | • . . | ↘ . .

I want to | I want | to go | I want | to go | now
. ↘ . | . • | . ↘ | . • | . • | ↘

I want | to go | to-morrow
. • | . • | . ↘ .

[1] See W. Jassem, 'Stress in Modern English,' *Bulletin de la Société Polonaise de Linguistique*, Cracow, 1952; *Intonation of Conversational English*, Wroclaw, 1952; 'Indication of speech rhythm of educated southern English.' *Le Maître Phonétique*, no. 92, 1949.

Mary and	George	specially	want	to go	to-morrow
o . .	o	o .	o	. o	. ◝ .

It is noticeable that the rhythmic beats of an utterance occur at fairly equal intervals of time. As a result of this, the speed at which the unstressed syllables are uttered—and the length of each—will depend upon the number occurring between the strong beats. All the unaccented syllables occurring between two strong beats may not, however, be uttered with equal rapidity: those following the strong beat of a rhythmic group tend to occupy slightly more time than those which precede the strong beat. Thus, in the phrase *The authority | of the government | is in danger*, particular speed of utterance is associated with the unaccented *of the, is in*, compared with *-ity, -ernment*. The rhythmic grouping of unaccented syllables generally correlates with grammatical word-clusters; a rhythmic division will not normally fall within a word pattern, but several words may combine to form one rhythmic group. It often happens, however, that an unaccented word may equally well be assigned to either of two rhythmic groups, e.g. in *They couldn't have chosen*, the weak *have* may be the last syllable of the first group or the first syllable of the second group.

Syllables carrying a strong stress are usually signalled to a listener by means of pitch prominence (see §§ 10.10, 10.11), but it frequently happens that a syllable stressed by the speaker has no pitch prominence. In this case, the syllable (or word) is given some accentual prominence for the listener by a combination of any of the factors of loudness (relating to stress), sonority (relating to the quality and length of the syllabic sound), and the nature of the word ('content' as opposed to 'form'). The following sentences illustrate the occurrence of stressed syllables said on a monotone within an utterance or, following a primary accent initiating a rise, within a rising post-tonic sequence of syllables; such stressed syllables, without pitch prominence, may be said to carry for a listener secondary accent of a rhythmic, qualitative/quantitative, or semantic kind :—

seems in

.	o	.	◝	. .

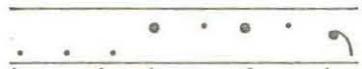
it seems impossible

wants, have in

.	o	.	o	.	◝	

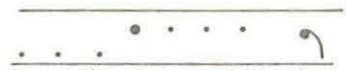
he wants to have it now

forty in



it was in nineteen forty-nine

(cf. *fifty* in



it was in nineteen fifty-nine

where the stressed /ɪ/ of *fifty* may not have the same qualitative prominence as the stressed /ɔ:/ of *forty*)

said, go in

they all said they would go

(where the vowels of both cases of *they* are likely to be weakened)

understand, say in

does he always understand what you say ?

(For variations of word accentual patterns in connected speech, see § 11.03.)

10.04 Quantity and Quality

(1) *Accented words.*—Vowels and continuant consonants in accented syllables which form the hub of a rhythmic group are shortened according to the number of unaccented syllables (especially following) in the group. Thus, the /aɪ/ of /taɪd/ (*tide*) shows progressive shortening in such rhythmic groups as *tidy, tidily, she tidied it,* etc.; the /ʌ/ and /m/ of /kʌm/ (*come*) are similarly shortened in *comfort, comfortable, come for me, circumference.* Or again, a comparable phonemic sequence will have slight variations of sound length according to the division into rhythmic groups : cf. *aboard a liner* /ə'bɔ:d ə'laɪnə/ and *a border-liner* /ə'bɔ:də ˌlaɪnə/—/ɔ:/ being slightly shorter in the second case ; *minor official* /'maɪnər ə'fɪʃl/ and *mine are official* /'maɪn ərə'fɪʃl/—/aɪ/ being longer in the second case. Such variation of rhythmic grouping, involving changes of quantity, constitutes a reality for the speaker, but it is doubtful whether slight modifications of this kind are markedly significant to a listener, since the choice of meaning for such similar phoneme sequences is normally determined by the context, such cues as are provided by rhythmic variation of quantity being redundant.

(2) *Unaccented words.*—A more marked effect is that which characterizes the quantity and quality of unaccented words. Content words (monosyllables and polysyllables) generally have in connected

	Unaccented	*Accented*
was	/wəz/	/wɒz/
we	/wɪ/ ([wi])	/wi:/
were	/wə/ + consonant	/wɜ:/
	/wər/ + vowel	/wɜ:r/
who	/hʊ, u:, ʊ/ ([hu])[1]	/hu:/
will	/l/	/wɪl/
would	/wəd, əd, d/	/wʊd/
you	/jʊ/ ([ju])	/ju:/

It should be noted that prepositions, e.g. *to, from, at, for*, apart from having a strong form when receiving a primary (nuclear, tonic) accent, also have a qualitative prominence when final and unaccented, e.g. *Where have they gone to?* (/tu:/, also /tʊ/, but not /tə/); *Where's he come from?* (/frɒm/ rather than /frəm/); *What are you laughing for, at?* (/fɔ:, æt/). When a preposition occurs before an unaccented pronoun, either the strong or the weak form may be used for the preposition, e.g. *I gave it to you* (/tə/ or /tu:/); *I've heard from him* (/frəm/ or /frɒm/); *I waited for you* (/fə/ or /fɔ:/); *I looked at her* (/ət/ or /æt/). Note, too, that certain form words, not normally possessing an alternative weak form for unaccented occurrences, may show such reductions in very rapid speech, e.g. *I* (/ə/) *don't know*; *What's your* (/jə/) *name?*; *I go by* (/bə/) *bus*; *Do you know my* (/mə/) *brother?*; *for love nor* (/nə/) *money*; *two or* (/ə/) *three*; *ever so* (/sə/) *many*. In the case of the disyllables *any, many*, a qualitative prominence may be retained on the first syllable under weak accent—/enɪ, menɪ/, but fully reduced, unaccented forms may be heard in rapid speech, e.g. *Have any more come?* /'hævnɪ 'mɔ: 'kʌm/; *How many do you want?* /'haʊ mnɪ dju: 'wɒnt/. Other monosyllabic form words normally retain their strong vowels in relatively unaccented positions, e.g. *on, when, then, one*, but again, in very rapid speech, reduced vowel forms may be heard, especially when the word is adjacent to a strongly accented syllable, e.g. *Then on* (/ən/) *Tuesday*; *When* (/wən/) *all's said and done*; *Then* (/ðən/) *after a time*; *one* (/wən/) *always hopes.*

It may be said that the more rapid the delivery the greater the tendency to reduction and obscuration of unaccented words. Even monosyllabic content words may be reduced in rapid speech, if they occur in a relatively unaccented situation adjacent to a primary accent, and especially if they contain a short vowel, e.g. /ɪ/, *You sit over here*

[1] A weak form with /h/ would normally be used when unaccented but following a pause.

/'ju: s(ə)t ˌəuvə 'hɪə/; /ʊ/, *He put it there* /'hi: p(ə)t ɪt ˌðɛə/; /ʌ/, *He'll come back* /'hi:l k(ə)m ˌbæk/; /e/, *Don't get lost* /'dəunt g(ə)t 'lɒst/; less frequently with the more prominent short vowels /æ, ɒ/, e.g. /æ/, *They all sat down on the floor* /ðeɪ 'ɔ:l sət 'dɑun ən ðə 'flɔ:/; /ɒ/, *We want to go* /'wi: wənt tə ˌgəʊ/; and, finally, the diphthong /əʊ/, with its dominant central [ə] element, is readily reducible to /ə/ under weak accent, e.g. *You can't go with him* /jʊ 'kɑ:nt gə 'wɪð ɪm/; *He's going to do it* /'hi:z gənə ˌdu: ɪt/.

10.10 Intonation [1]

The meaning of an English utterance, i.e. the information it conveys to a listener, derives not only from its changing sound pattern and the contrastive, accentual prominences already referred to, but also from associated variations of pitch. Such rises and falls in pitch level, or patterns of *intonation*, have two main functions :—

(1) *Accentual.*—Intonation changes are the most efficient means of rendering prominent for a listener those parts of an utterance on which the speaker wishes to concentrate attention; pitch change is especially significant as a cue for signalling the word or words carrying primary (nuclear) accent. It should be remembered, however, that the accentual pattern of a response is often largely conditioned by constraints imposed by the context (see § 10.02).

(2) *Non-accentual.*—In addition, intonation is used as a means for distinguishing different types of sentence, e.g. the same sequence of words may, with a falling intonation, be interpreted as a statement or, with a rising intonation, as a question. Moreover, a listener derives from a speaker's intonation information as to the latter's emotional attitude (to the listener or to the topic of a conversation) or personality, e.g. his intonation might reveal a patronizing attitude to the listener, an incredulous attitude to the topic or a querulous disposition.

In so far as a listener interprets correctly those parts of an utterance upon which the speaker wishes to concentrate attention, or is aware of the speaker's attitude to him, or makes judgments upon the personality of the speaker, the pattern of intonation used may be said to constitute a linguistic system which has a communicative function within a particular community. There seems no doubt that intonation in its accentual function and in that non-accentual function concerned with the distinction of sentence types, does constitute a linguistic reality of

[1] For a more complete treatment of intonation in these terms, see J. D. O'Connor and G. F. Arnold, *The Intonation of Colloquial English*, 1961, Longmans, and works referred to in the bibliography under Palmer, Kingdon, Schubiger, Jassem, Lee.

this kind. Since, however, we sometimes misinterpret the emotional attitude as conveyed by intonation, it may be said that non-accentual intonation patterns of this kind are less perfectly systematized, or that such linguistic systems are more numerous and applicable to smaller communities (regional or social) than phonological systems, so that a faulty judgment of emotional attitudes conveyed by intonation cues may derive from an interpretation of these cues in terms of our own, different, intonation usage in showing such attitudes (cf. the interpretative adjustments needed on the phonological level between speakers of two different types of English).

10.11 The Accentual Function of Intonation

The various degrees of accentuation in an utterance may be signalled by means of intonation in the following way :—

(*a*) *Primary* (*nuclear*) *accent*—by means of a change of pitch *direction* initiated by the syllable receiving the accent (marked ˏ, ˋ, ˏˋ, ˊ, ˇ, ˆ).

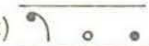
(*b*) *Secondary* (*pitch prominent*) *accent*—by means of a change of pitch *level* (higher or lower) on the accented syllable (marked ˈ).

(*c*) *Secondary accent without pitch prominence*—secondary accent on some words may be manifested by qualitative, quantitative, or rhythmic prominence, without pitch prominence (marked ˌ).

(*d*) *Unaccented syllables*—do not normally have pitch or other prominence and are unmarked (see, however, § 10.15).

10.12 Realization of Primary Accent

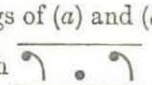
The primary accent (or accents) in a sentence is shown by initiating a change of pitch direction, with the *nucleus* (falling, rising, or a combination of the two) on the appropriate syllable of the word (or words) on which attention is particularly to be concentrated. The *situation* of the nucleus or nuclei is, therefore, of prime importance in conveying meaning, e.g. compare the meanings of the following sentence according to the shift of nucleus :—

(*a*) ˋJack ˌlikes ˌfish (i.e. not George, but Jack) 

(*b*) ˌJack ˋlikes ˌfish (i.e. there is no question of his *hating* fish) 

(*c*) ˌJack ˌlikes ˋfish (i.e. not meat or poultry, etc.) 

or again, the meanings of (*a*) and (*c*) may be combined in

ˋJack ˌlikes ˋfish 

In the sense that the nuclear syllable stands out from amongst its neighbours (both accented and unaccented syllables), the nucleus and its situation may be said to have a special contrastive function. But it may happen that special contrastive prominence is not attached to any of the accented syllables of a sentence—this is more common in reading than in normal speech, in which it is usual for one or more words in a sentence to be ' pointed '. Where no word is specially contrasted, the nuclear change of pitch direction will be initiated by the last accented syllable, e.g. :—

By the 'time we 'got to the ˌhouse, we were 'all 'wet ˏthrough.

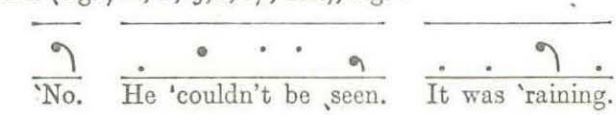
In a more animated style, more typical of conversation, the same sentence might have nuclear pitch changes on several accented syllables, e.g. :—

By the `time we `got to the ˋhouse, we were `all ˌwet `through.

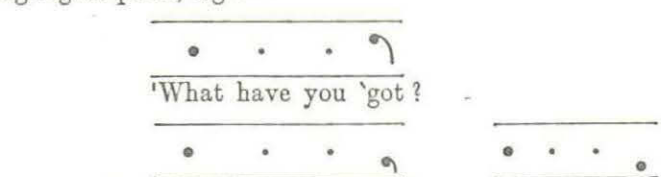
The way in which the change of pitch direction is realized depends on the type of change involved and on the nature and extent of the phonological sequence covered.

10.13 Types of Nucleus

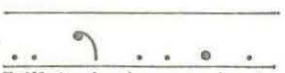
(1) *The falling nucleus* (ˌ, `).—The falling glide may start from the highest pitch of the speaking voice and fall to the lowest pitch (in the case of the *high-fall*), or from a mid pitch to the lowest pitch (in the case of the *low-fall*), or with variations of starting point according to the intonation context. The falling glide is most perceptible when it takes place on a syllable containing a long vowel or diphthong or a voiced continuant (e.g. /m, n, ŋ, l, z/, etc.), e.g. :—

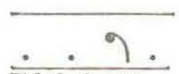


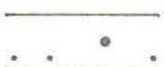
When a fall occurs on a syllable containing a short vowel with its limits formed by fortis, voiceless consonants (especially the stops /p, t, k/), the glide, particularly of a low-fall, is so rapid that it is not easily perceptible, or may be realized merely as a low level pitch in relation to a preceding higher pitch, e.g. :—

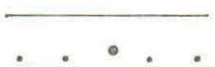


Again, when syllables follow the nucleus—the *tail*—the fall may be realized as the juxtaposition of relatively high pitch on the nuclear syllable and low pitches on the syllables of the tail, e.g. :—



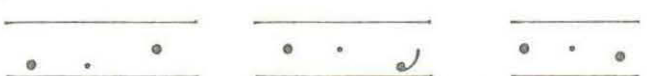
but 

or 

or 

(2) *The rising nucleus* (ˌ, ˊ).—In the same way, a rising glide, which may extend from low to mid, or from mid to high, or with other variations of starting and end points between low and high, is more easily perceptible when it occurs on a syllable containing a long vowel or diphthong or a voiced continuant consonant, e.g. :—

When a low-rising glide occurs on a short syllable, it must necessarily be accomplished much more rapidly, or may merely consist of a relatively high level pitch in relation to a preceding low pitch, or even of a slightly lowered level pitch in relation to a preceding mid or high pitch, e.g. :—

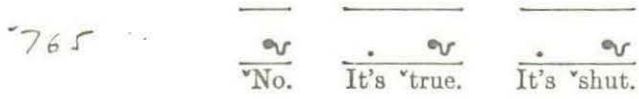


or

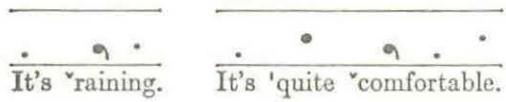
With a tail, the rise is achieved by means of a relatively low pitch on the nuclear syllable with an ascending scale on the following syllables, e.g. :

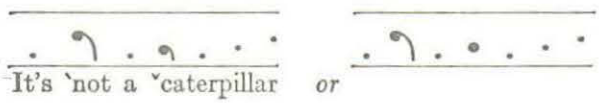
(3) *The falling-rising nucleus* (ˇ).—The fall and rise may be confined within one syllable, the glide beginning at about mid level and ending at the same level (or slightly above or below); in the case of a

short syllable, the dip in pitch is made extremely rapidly and may be realized as an instant of 'creaky' voice or even of cessation of voice, e.g. :—

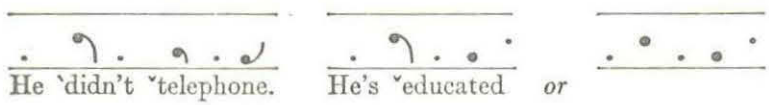


When an unaccented tail follows the nuclear syllable, the fall occurs on the nuclear syllable and the rise is spread over the tail, e.g. :—

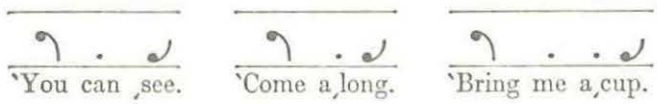


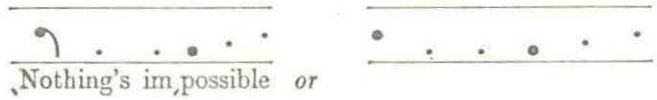


When a secondary accent follows the primary (nuclear) accent, the fall takes place on the nuclear syllable and the rise is initiated on the syllable carrying the secondary accent, e.g. :—

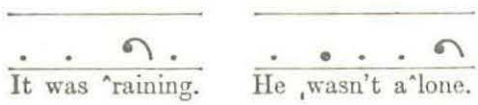


What amounts to the same pattern may occur where, within the same group, a word containing a falling syllable is followed by a word containing a rising syllable, both syllables carrying a primary accent, e.g. :—

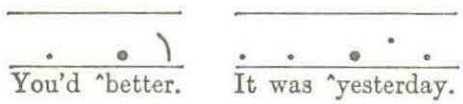




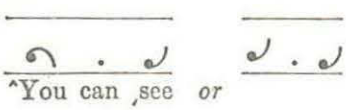
(4) *Rising reinforcement of a fall* (ˆ).—A fall may be reinforced by an introductory rise, especially on a long syllable containing voiced continuant consonants (which may be given extra length), e.g. :—



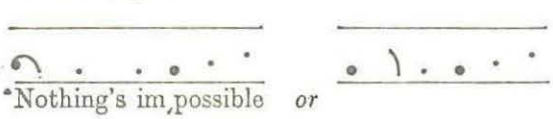
A reinforced short syllable followed by a tail may be realized as a low accented nuclear syllable followed by a fall on the tail, e.g. :—



When such reinforcement occurs on the falling part of a divided fall-rise (i.e. when the rise is situated on a later word), it has the effect of focusing attention on the falling nucleus, e.g. :—



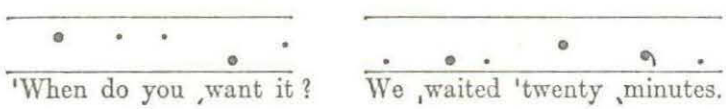
or



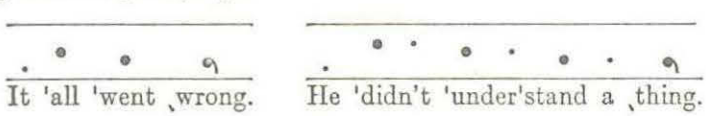
or

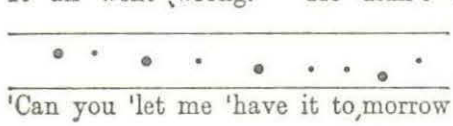
10.14 Realization of Secondary Accent

(1) *Pre-nuclear.*—Syllables preceding the nucleus may have *pitch prominence* (secondary accent, marked ') through being given a high level pitch when initial or a high level pitch in relation to preceding syllables, e.g. :—



The first pre-nuclear accented syllable is known as the *head* (a ' high head ' if given pitch prominence as above); any syllables occurring between the head and the nucleus constitute the *body*. Accented syllables within the body may be given pitch prominence by means of a step down in pitch, e.g. :—





Additional pitch prominence may also be given to an accented syllable in the body, in a sequence of such accented syllables, by means of a

step up (marked "), thus constituting in effect a supplementary head, e.g. :—

We 'came 'home to 'find "nobody ˌthere

'One of the 'two 'lorries "suddenly be'gan to ˌskid.

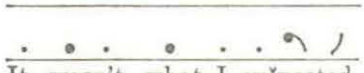
Such a pattern of pitch accentuation would have the effect of giving special prominence to *all* and *suddenly*. Alternative intonation patterns would divide these sentences into two groups :—

We 'came ˌhome | to ˌfind 'nobody ˌthere

'One of the 'two ˌlorries | 'suddenly be'gan to ˌskid.

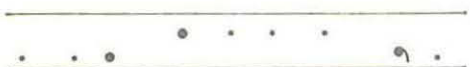
Pre-nuclear syllables may also be accented *without pitch prominence*, i.e. they are accented only for reasons of rhythmical stress, quality, or quantity, or because of the content nature of the word (marked ˌ, whether high or low in pitch). They may be the only accented syllables before the nucleus and be said on a relatively low level pitch (a ' low head '), e.g. :—

It ˌcan't be ˌhelped. ˌPut it ˌover `there.

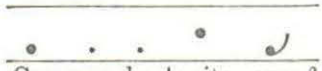


It ˌwasn't ˌwhat I exˇpected.

If such syllables precede a high head, they are relatively low and level, e.g. :—



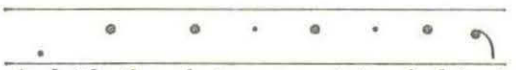
He aˌrrived 'just as we were ˌleaving.



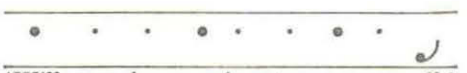
ˌCan you be 'quite ˌsure ?

(the high head on *just* and *quite* giving these words special prominence).

If such syllables occur within the body, they are kept on about the same pitch as the preceding pitch-prominent syllable, e.g. :—

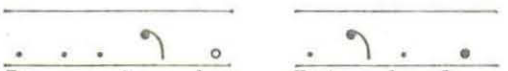


'And 'that's ˌwhat we ˌwant to ˌfind `out.



'Will you be ˌcoming to ˌsee us ˏoff?

(2) *Post-nuclear.*—After a falling nucleus, a secondary accent is manifested by rhythmic, qualitative, or quantitative prominence, the pitch remaining low (marked ˌ), e.g. :—



It was a `new ˌhat. It `can be ˌdone.

After a rising nucleus, however, syllables carrying secondary accent (also marked ˌ) continue on a rising pitch, e.g. :—

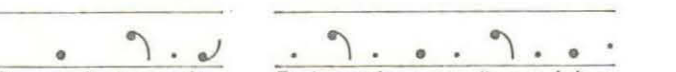
'Did you ˏsee him ˌyesterday? ˏI didn't ˌthink it was ˌfunny.

When final, they may have a certain additional pitch prominence, e.g. :

'Does he ˏlike ˌdogs? 'Did you 'see how ˏpale he ˌlooked?

cf. 'Have you 'put the ˏkettle ˌon? *and* 'Have you 'got the ˏcatalogue?

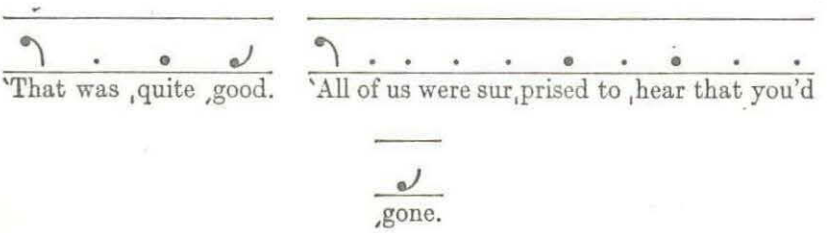
In the case of a falling-rising nucleus (whether on a single word where a secondary accent falls on the final syllable or in cases where the rise takes place on a following word), the pitch prominence (consisting of a rise) is very considerable, e.g. :—



I `haven't ˌmuch ˇappetite. It `wasn't ˌvery ˇappetizing.

I've `got the ˇcatalogue cf. I've `put the `kettle ˏon.

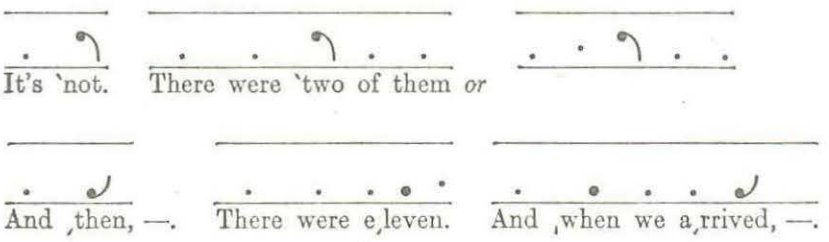
(3) *Secondary accent between fall and rise.*—Syllables carrying secondary accent, without pitch prominence, may occur between a fall and a rise; in this case, the unaccented syllables and those carrying secondary accent have a relatively low level pitch, e.g. :—



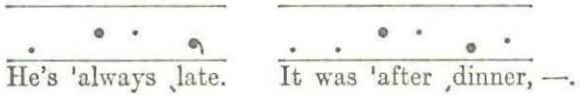
10.15 Realization of Unaccented Syllables

Unaccented syllables, in addition to the fact that they are said very rapidly and usually undergo some obscuration of their quality, do not normally have any pitch prominence. They may occur before the head or the nucleus, within the body, or after the nucleus.

(1) *Pre-nuclear.*—Unaccented syllables occurring before a nucleus (without a head), like syllables carrying secondary accent without pitch prominence, are normally relatively low, whether the nucleus is a fall or a rise, e.g. :—

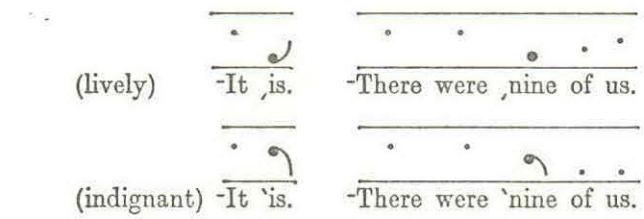


Unaccented syllables before a high head are usually said on a relatively low pitch, the head having contrastive prominence in relation to them, e.g. :—

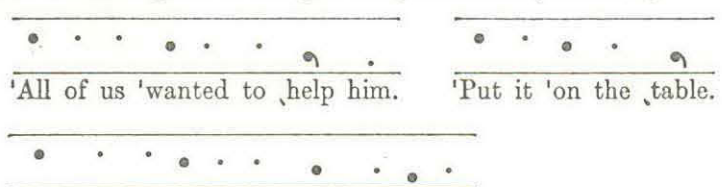


If pre-nuclear unaccented syllables, their weak quality remaining, are said on a relatively high pitch (marked ˉ), the utterance has a specially

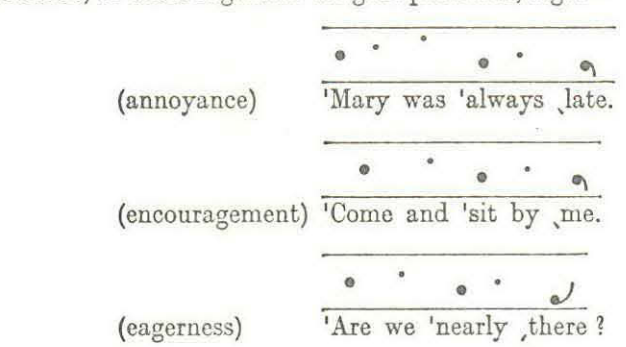
bright, lively, encouraging character before a rise, or an indignant, quarrelsome note before a fall, e.g. :—



(2) *Within body.*—Within the body, unaccented syllables remain on almost the same pitch as the preceding accented syllable, e.g. :—

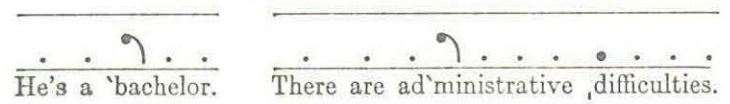


If the unaccented syllables are on a higher pitch than the preceding accented syllables, a special impression of liveliness, eagerness, impatience, or encouragement is again produced, e.g. :—



It is to be noted that in these cases the accented syllables within the body receive extra relative pitch-prominence, the pattern being equivalent to a series of rising nuclei.

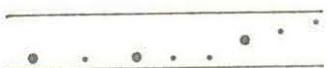
(3) *Post-nuclear.*—Unaccented syllables following a falling nucleus remain on a low level, e.g. :—



After a rising nucleus, unaccented syllables continue (or effect) the rise (the last syllable of all having a short rising glide sometimes, which gives it an extra prominence without contrastive significance), e.g. :—


'Put it 'on the ˌtable.

It's ˌnot imˌpossible.


ˌHas he ˌgot a soˈlicitor ?

The rise of a falling-rising nucleus may be spread over the following unaccented syllables, e.g. :—


He 'doesn't 'like ˇcriticism.

Between a fall and a rise, unaccented syllables remain relatively low, e.g. :—


ˈYesterday was ˌfine.

10.20 The Attitudinal Function of Intonation

As was pointed out in § 10.10, intonation, in addition to its function of providing a means of accentuation, may also serve to distinguish sentence types (e.g. statement and question) and to indicate the emotional attitude of the speaker. Such a function applies equally to utterances consisting of more than one word and to those of a single word (of one or more syllables). In these cases, it is not so much the situation of the nucleus which is of importance, but rather the type of nucleus employed, i.e. whether a fall, rise-fall, rise, or fall-rise is used. Thus, a statement form of words may be made into a question if a rise is used instead of a fall, e.g. ' He's 'not ˌthere ' (statement) *v.* ' He's ˌnot 'there ? ' (question) ; or, in such a case as that quoted by H. E. Palmer,[1] the sentence ' He doesn't lend his books to anybody ' may have two meanings according to whether *anybody* is said with a falling nucleus (i.e. he lends them to nobody) or with a falling-rising nucleus (i.e. he does lend them to some people). This semantic function of intonation

[1] See H. E. Palmer, *English Intonation with systematic exercises*, Heffer, 1922 ; also W. R. Lee, *An English Intonation Reader*, Macmillan, 1960.

(typified by the last example) occurs less frequently than that which shows the speaker's emotional attitude: compare, for instance, ' It's ˌall ˏright ' (a plain statement of fact), ' It's 'all ˏright ' (an encouragement), and ' It's ˌall ˏright ' (a grudging or petulant agreement). In this connection, in addition to the type of nucleus used, the pitch of that part of the utterance preceding the nucleus (the head) is also of great importance (and sometimes the pitch of unaccented syllables—see § 10.15 (1)).

In the examples given in the following sections, an attempt is made to assign generalized verbal descriptions to the attitudes conveyed by intonation patterns in respect of various types of utterance; the main types of utterance are:—

(*a*) assertions
(*b*) questions containing an interrogative word (X-questions)
(*c*) questions expecting ' Yes ' or ' No ' as an answer (Yes/No questions)
(*d*) question tags
(*e*) commands, requests, etc.
(*f*) exclamations, greetings, etc.

The examples are given as isolated utterances, but it should be remembered that the attitudinal meaning of an utterance must always be interpreted within a context, both of the situation and also of the speaker's personality. It may well happen that an intonation which is neutral in one set of circumstances might be, for instance, offensive or patronizing when used by another person or in other circumstances.

10.21 Falling Nuclei

A falling nucleus, whether ˎ or ', is generally matter-of-fact, separative, and assertive, the higher the fall the more vigorous the degree of finality implied. No explicit appeal is made to the listener, yet the pattern is not necessarily impolite; a conversation amongst people who are intimately acquainted might, for instance, exhibit a preponderance of falling intonations, without the exchange being querulous or lacking the social courtesies of speech.

(1) *Low-falling nucleus*

Pattern—nucleus (tail)

(*a*) ˌNo. ˎYesterday (detached, unexcited)
(*b*) ˌWhen ? (curt) ˌHow are you ˌgoing to ˌdo it ? (weak insistence on *how*)

(*c*) ˌAre you ˌcoming ? ˌDoes he ˌwant to ? (curt, impatient, testy)
(*d*) (He ˌdoes,) ˌdoesn't he ? (ˌLovely,) ˌisn't it ? (calmly presupposing agreement)
(*e*) ˌCount them. ˌGet it ˌthen (calm, detached, peremptory)
(*f*) ˌTragic ! (quietly sympathetic ; or distant and unmoved, especially if the fall does not reach the lowest level) ; ˌMorning (perfunctory)

Pattern—high head, nucleus (tail), the high pre-nuclear pattern showing a degree of lively interest.

(*a*) The 'parcel a'rrived on ˌThursday (matter-of-fact, but interested)
(*b*) 'What do you 'want to ˌdo ? (blunt to strangers, but a common unemotional form amongst intimates)
(*c*) 'Are you ˌgoing ? (peremptory, impatient)
(*e*) 'Put it 'over ˌthere (polite, neutral)
(*f*) 'What a ˌmess (phlegmatic, mild)
'(-)Good ˌmorning (the high pitch on *good*, whether it is an accented head or an unaccented high pre-nuclear syllable, makes the greeting pleasant, though routine)

Pattern—low head, nucleus (tail) ; the low pre-nuclear pattern throws the nucleus into greater prominence and often shows lack of interest.

(*a*) It's ˌall we could exˌpect (surly, uninterested)
(*b*) ˌWhat are we ˌgoing to ˌdo ? (resigned, bored)
(*c*) ˌHave you ˌgot the ˌtickets ? (uninvolved, perfunctory)
(*e*) ˌLeave it ˌon the ˌtable (pre-occupied, expecting to be obeyed as a matter of course)
(*f*) ˌHow aˌnnoying (bored, unconcerned, sarcastic)
ˌGood ˌmorning (routine, perfunctory greeting)

(2) *High-falling nucleus.*—Strongly contrastive or contradictory ; often showing strong indignation or excitement ; very common in ordinary colloquial speech.

Pattern—(pre-head), nucleus (tail).

(*a*) 'No. It 'was. Of 'course it ˌisn't (vigorous agreement or contradiction)
(*b*) 'Why ? 'How ˌcan she ? (surprise, indignation, incredulity)
(*c*) Do 'you ˌlike it ? You're 'sure ? (insistence, demands an answer)
(*d*) (She 'doesn't), 'does she ? (He 'does), 'doesn't he ? (demands agreement)
(*e*) 'Drop it. 'Cancel it ˌthen (energetic, angry command)
(*f*) 'What a surˌprise ! (strong surprise, indignation)
'Good ˌmorning (hearty greeting)

Pattern—low head, nucleus (tail); a high-falling nucleus usually depends for its effect on a relatively low pattern of any preceding syllables, all the accentual force being concentrated in the nucleus; the following sentences said with high heads and bodies would have less contrast associated with the nuclear word.

(*a*) ˌNot at ˋall. I ˌquite aˋgree.
(*b*) ˌWhy do you ˋwant it ?
(*c*) ˌCan we ˌhave it ˋnow ?
(*e*) ˌGo and ˋfind it.
(*f*) ˌWhat a ˌstupid misˋtake !
ˌGood ˋmorning (a bright, cheerful greeting)

(3) *Rising-falling nucleus.*—The initial rise may reinforce the meaning of any high fall, often with additional warmth, indignation, sarcasm, etc. A preceding head will usually be relatively low.

Pattern—(low head) nucleus (tail).

(*a*) ˆYes. It was aˆpalling. Aˌbout ˆtime (all may show enthusiastic agreement or enthusiasm; but ' ˆYes ', said slowly, expresses doubt, /ˆpɔːl/, prolonged, shows horror, and ' ˆtime ' may reveal sarcasm or indignation)
(*b*) ˌWhat does his ˆfather ˌdo ? (suspicious, indignant interest)
(*c*) ˌCan you be ˆsure ? (mocking, knowing, suspicious)
(*d*) ˆIs she ? ˆDoesn't he ? (incredulous, mocking, indignant)
(*e*) ˌGet aˆnother one (pert, haughty, sulky, petulant)
(*f*) ˌOh, inˆdeed. ˌHow ˆnice for you (sarcastic)
ˌGood ˆmorning (portentous, ironical greeting)

10.22 Rising Nuclei

(1) *Low-rising nucleus.*—Essentially unfinished and continuative, often with overtones of politeness, encouragement, pleading, diffidence, suspicion, etc.

Pattern—nucleus (tail).

(*a*) ˌNo. ˌPossibly (tentative, grudging, encouraging). Note also the introductory usage: ' ˌThen, —,' and adverbially after a main clause: ' (We were aˋlone), ˌluckily.'
(*b*) ˌHow did you ˌdo it ? (insistence on ' How '—the lower the starting point of the rise, the greater the insistence)
(*c*) ˌAre they ˌcoming ? (insistence on ' Are ')
ˌIs he ? (doubtful, indifferent)
(*d*) (He's ˋgot one), ˌhasn't he ? (It isn't ˋthere), ˌis it ? (doubtful, asking for information)

(*e*) ˌWait. ˌHold it (gentle command or request)

(*f*) ˌWell (introducing a topic, or an uninterested question form)

Pattern—high head, nucleus (tail); a relatively high pre-nuclear pattern gives an effect of a fresh thought, lively interest, appeal, encouragement, etc.

(*a*) It's 'all ˌright. It 'doesn't ˌmatter. She 'won't be ˌlong (reassuring statements). Also dependent clauses occurring before the main clause, e.g. ' 'When we aˌrrived, (they'd al'ready ˌgone) '; following a main clause containing a falling nucleus, the pre-nuclear pattern of this dependent clause is likely to be low, e.g. ' (They'd al'ready `gone), ˌwhen we aˌrrived '.

(*b*) 'What's the ˌtime ? (polite inquiry)

(*c*) 'Can you ˌcome ? (polite, interested)

(*e*) 'Sit ˌdown. 'Come over ˌhere (pleasant, encouraging invitation)

(*f*) 'All the ˌbest. 'Good ˌluck (cheerful good wishes). 'Good ˌmorning (cheerful, friendly greeting)

Pattern—low head, nucleus (tail); a relatively low pre-nuclear pattern often signifies complaint, suspicion, veiled threat, lack of interest or enthusiasm, etc.

(*a*) It's ˌall ˌright. It's ˌnot imˌportant (resigned, disgruntled, long-suffering, bored). Also in dependent clauses, before or after the main clause, e.g. " If ˌJack's ˌhome, (we'll ˌcome at `eight) '; ' (We'll ˌcome at `eight), if ˌJack's ˌhome '; in the first case, ' ˌJack ' has not the contrastive accentual force which would be associated with ' 'Jack '.

(*b*) ˌWhat have you been ˌdoing ? (unsympathetic, menacing, threatening)

ˌWhat's your ˌname ? (peremptory, routine cross-examination)

(*c*) ˌCan you ˌcome ˌnext ˌweek ? (uninterested, disgruntled, bored)

(*e*) ˌTry aˌgain (routine request, peremptory)

ˌDon't ˌleave the ˌdoor ˌopen (long-suffering, complaining)

(*f*) ˌGood ˌmorning (polite but perfunctory greeting)

(2) *High-rising nucleus.*—This nucleus, rising to a high pitch, is associated essentially with questions, e.g. :—

(i) An elliptical question (showing eagerness, brightness, enthusiasm, or asking for a repetition) :—

'Coffee ? (= 'Will you have some more coffee ?' or 'Did you say coffee ? ')

'Like it ? (= ' Do you like it ? ')
'Monday ? (= ' What about Monday ? ' or ' Did you say Monday ? ')
'John ? (= ' I'm calling you—are you there ? ' or ' Is that you, John ? ' or ' Did you say " John " ? ')
'When ? (= ' I can hardly believe what you said—would you repeat it ? ')

(ii) A question showing great eagerness, excitement, concern, indignation, etc. :—

It 'is ? 'You 'did (surprise, incredulity)
ˌCan you 'come ? (eager expectancy)
You ˌactually 'saw him ? ˌWhat 'me ? (indignation, surprise, horror)
ˌCan we a'fford it ? (concern, expectancy, apprehension)
It 'wasn't 'yours ! (dismay, surprise, indignation)

10.23 Falling-Rising Nuclei

The falling-rising nucleus combines the dominant effect of the fall (contradictory, contrastive) with any of the emotional or meaningful attitudes (not expressed verbally) associated with the rise. Both fall and rise may occur within one word, e.g. :—

Pattern—(accented or unaccented pre-nuclear syllables), nuclear word (tail).

(*a*) ˇNo (doubtful or encouraging)
He ˇcould (forcefully reproachful)
ˇSometimes (encouraging)
I 'haven't 'much ˇappetite (' but I'll join you to be polite ' ; if the pre-nuclear pattern is low, the effect is less agreeable—disgruntled)
I'm ˇwaiting (' so do hurry up ')
ˇLately, (it's been ˌtoo ˋwet) (lively, introductory, continuative)
If he ˇcould, (we'd be deˋlighted) (forceful dependent clause)

(*b*) ˇWhen ? ˇHow ? (forceful, encouraging, prompting)

(*e*), (*f*) ˇJohn ! ˇLook ! (an appealing, inviting summons)
ˇGently ! (encouraging, soothing, warning)

or the fall and the rise may occur on different, accented, words, e.g. :—

Pattern—(accented or unaccented pre-nuclear syllables), word containing fall (intervening syllables, with or without accent), word containing rise (tail).

(*a*) 'He ˌcould (' but I doubt whether *you* could ')
It ˌwasn't 'your ˌfault (reasoning, sympathetic)
'That's the ˌway to ˌdo it (encouraging, pleasant)
(He'll 'have to), 'one ˌday (insistence on *one*, followed by the rise of the dependent clause)
On 'Saturday ˌmorning, (we'll be ˌgoing 'out) (contrastive fall on *Saturday* followed by continuative, introductory rise)

(*b*) 'When can you ˌcome ? (a polite questioning rise, preceded by a fall on *when* to insist on the precise time)

(*c*) Will 'you be ˌcoming on ˌSunday ? (a polite questioning rise, preceded by a fall on *you* to focus attention on *your* intended action)
'Can ˌyou ? (insistence on your *ability* to do something associated with the polite questioning rise)

(*e*) 'Do ˌsit ˌdown. 'See if you ˌcan (compulsive request)
'Mind ˌhow you ˌgo (strong but sympathetic warning)

(*f*) 'Well ˌdone ! (warm, sympathetically appreciative)

10.24 Multi-nuclear Patterns

It often happens that, in addition to the composite nuclear patterns consisting of a fall + a rise, a number of nuclei occur in a single syntactic group. Such a series of nuclear syllables, often occurring rhythmically at equal intervals of time, may produce an effect of a categoric, downright, hectoring, insistent, self-assertive way of speaking. Nevertheless, such multi-nuclear patterns are extremely common in ordinary conversation and often serve no other purpose than to produce a lively, animated effect and to focus attention on the important words of an utterance, e.g. :—

(1) *A series of falling nuclei preceding a fall.*—An emphatic equivalent of a head + body + falling nucleus.

(*a*) We 'never 'thought he 'had a 'chance
It was 'all we could 'do to 'stop him

(*b*) 'What do you 'think we can 'do ?

(*c*) 'Can she 'do the 'work ?

(*e*) 'Don't be 'such a 'fool
'Go and 'try a'gain.

(*f*) 'What an 'awful 'mess. ˌGood 'after'noon

(2) *A series of falling nuclei preceding a rise.*—An emphatic equivalent of a head + body + rising nucleus (or fall-rise).

(*a*) I `always `thought he ,would
I was `most im`pressed that `they should ,want to
It `looks as `black as `night ,out˅side
(*b*) `How many `seats were you `able to ,get ?
(*c*) Are you `sure that `George and `Mary ,know ?
(*e*) `Mind you `put your `hat and `coat ,on
`Come and `see us `when you ˅can
(*f*) `What a `very `silly ,little ,boy !

(3) *A series of rising nuclei preceding a fall.*—An emphatic, impatient, or patronizing equivalent of a head + body + falling nucleus.

(*a*) And ,now you ,both must ,go to `bed (decisive, but humouring)
It's ,only ,in your ,own `interests (strongly persuasive, somewhat impatient)
(*b*) ,What on ,earth do you ,think you're `doing ?
(*c*) ,Can you ex,pect them to ,do it a`lone ?
(*e*) ,Don't be ,such a ,silly `fool !
(The last three examples showing impatience, exasperation.)

(4) *A series of rising nuclei preceding a rise.*—A repetitive indication of the attitudes associated with a rising nucleus (appeal, complaint, sarcasm, etc.), being the equivalent of a head + body + low rise.

(*a*) You ,know we ,only ,wanted to ,help (patronizing, grudging)
(*b*) ,When were you ,thinking of ,paying it ,back ? (patronizing, sarcastic)
(*c*) ,Will you be ,coming to ,see us on ,Monday ? (apprehensive, unenthusiastic)
(*e*) ,Don't ,make ,such a ,noise ! (quietly pleading, the appeal being stronger as a result of the succession of rises)

The same sentences said with a series of higher rises would give an effect of greater animation, excitement, urgency, etc.

10.25 Other Devices for Signalling Attitudes

Many other devices exist for expressing in sound the mood of a speaker in addition to the actual words which he uses. A rapid rate of delivery, for instance, may express irritation or urgency, whereas a slower rate may show hesitancy, doubt, or boredom in statements, or sympathy or encouragement in questions and commands; a repeated alveolar click denotes discontent; an egressive voiceless air-stream with friction at the rounded lips and a falling pitch (produced by the adjustment of the mouth cavity by movement of the tongue) expresses

surprise, admiration, relief, whereas an ingressive air-stream of the same type may, in addition, be used to signify pain or pleasurable anticipation; and the utterance may be punctuated by sighs, denoting boredom, impatience, or sorrow. The intonation of a sentence, however, may indicate attitudes of the speaker other than those associated with the type of nucleus used or the choice of a relatively low or high pre-nuclear pattern of syllables:—

(1) *Range of intonation.*—The degree of enthusiasm, liveliness, interest, in those tunes where the pre-nuclear pattern is relatively high will be greater the higher the pitch of the head, e.g. with an extra-high pitch on the syllable italicized in the following examples and a consequent greater pitch separation of any accented syllables between the head and nucleus:—

(*a*) It's '*aw*fully ˌgood
'*Ev*eryone 'thought it was ˌmarvellous
It's '*not* as 'bad as you ˌthink

(*b*) '*What's* the ˌtime ?

(*c*) '*Are* you ˌsure ?

(*e*) '*Look* at ˌthat

(*f*) '*Well* I ˌnever
'*Good* 'afterˌnoon

(2) *Continuous level pitch.*—In certain situations (excluding the monotonous chant of prayers sometimes heard in church), a continuous level pitch may be used to express a curt, detached, preoccupied attitude, which may give the impression of overbearing superiority if the pitch is relatively high, or of surly impatience if the pitch is relatively low:—

ˌThat'll be ˌall. It's ˌnot a ˌbit of ˌgood
ˌWhat do you ˌwant ? ˌAre you ˌcoming ?
ˌCome ˌin. ˌHow do you ˌdo

(The accented syllables are here marked with ˌ, since, whether they are high or low, they do not have pitch prominence.)

10.30 Advice to Foreign Learners

Of all the features of accent in connected speech, the foreign learner should pay particular attention to:—

(1) The choice of words upon which the speaker's stress and accent (primary and secondary) fall and the situation of the nucleus. This placing of the accent in connected speech will usually be determined by the meaning which is to be conveyed, subject to the constraints imposed by the context and total situation.

(2) The rhythmic nature of connected English. The accented syllables follow each other at roughly equal intervals of time; any intervening unaccented syllables are said rapidly, the greater the number of such syllables the more rapidly they are uttered; unaccented syllables are associated with the hub of an accented syllable, so that an utterance is divisible into a series of rhythmic groups, themselves closely related usually to syntactic groups.

(3) The related weakening of unaccented words, e.g. the appropriate use of weak forms. A strong form in what is intended to be an unaccented position will give an impression of accent to the English ear. Thus, *He was late*, pronounced /hi: wɒz ˈleɪt/ instead of /hɪ (or [hi]) wəz ˈleɪt/, gives undue prominence to *was* and may appear to contradict some such statement as *He wasn't late.*

He should also be aware of the attitudinal implications of pitch variation. Frequently, the tunes used in English to distinguish sentence types will, in their broad pattern, be similar to those of his own language; but he should note, for instance, the high frequency of occurrence among English RP speakers of falling-rising tunes (often within a single word), which are less commonly encountered in other languages. He should be acquainted, too, with the English usage of falls and rises to signify the mood of the speaker, so that an over-use of rises will not give an unintentional impression of, for example, diffidence or complaint, and too many falls create an unwitting effect of impolite assertiveness.

APPENDIX C

GIMSON'S ANALYSIS OF THE ACCOUSTIC AND AUDITORY ASPECTS OF THE SOUNDS OF SPEECH

3.1 To complete an act of communication, it is not normally sufficient that our speech mechanism should simply function in such a way as to produce sounds; these in turn must be received by a hearing mechanism and interpreted, after having been transmitted through a medium, such as the air, which is capable of conveying sounds. We must now, therefore, examine briefly the nature of the sounds which we hear, the characteristics of the transmission phase of these sounds, and the way in which these sounds are perceived by a listener.

3.2 When we listen to a continuous utterance, we perceive an ever-changing pattern of sound. As we have seen, when it is a question of our own language, we are not conscious of all the complexities of pattern which reach our ears: we tend consciously to perceive and interpret only those sound features which are relevant to the intelligibility of our language. Nevertheless, despite this linguistic selection which we ultimately make, we are aware that this changing pattern consists of variations of different kinds: of *sound quality*—we hear a variety of vowels and consonants; of *pitch*—we appreciate the melody, or intonation, of the utterance; of *loudness*—we will agree that some sounds or syllables sound 'louder' than others; and of *length*—some sounds will be appreciably longer to our ears than others. These are judgments made by a listener in respect of a sound continuum emitted by a speaker and, if the sound stimulus from the speaker and response from the listener are made in terms of the same linguistic system, then the utterance will be meaningful for speaker and listener alike. It is reasonable to assume, therefore, that there is some constant relationship between the speaker's articulation and the listener's reception of sound variations. In other words, it should be possible to link through the transmission phase the listener's impressions of changes of quality, pitch, loudness, and length to some articulatory activity on the part of the speaker. It will in fact be seen that an exact parallelism or correlation between the production, transmission, and reception phases of speech is not always easy to establish, the investigation of such relationships being one of the tasks of present-day phonetic studies.

3.3 The formation of any sound requires that a vibrating medium should be set in motion by some kind of energy. We have seen that in the case of the human speech mechanism the function of vibrator is often fulfilled by the vocal cords and that these are activated by air pressure from the lungs. In addition, any such sound produced in the larynx is modified by the resonating chambers of the pharynx, mouth and, in certain cases, the nasal cavity. The listener's impression of *sound quality* will be determined by the way in which the speaker's vibrator and resonators function together.

3.31 Speech sounds, like other sounds, are conveyed to our ears by means of waves of compression and rarefaction of the air molecules (the commonest medium of communication). These variations in pressure, initiated by the action of the vibrator, are propagated in all directions from the source, the air molecules themselves vibrating at the same rate (or frequency) as the original vibrator. In speech, these vibrations may be of a complex but regular pattern, producing ' tone ' such as may be heard in a vowel sound ; or they may be of an irregular kind, producing ' noise ', such as we have in the consonant *s* ; or there may be both regular and irregular vibrations present, i.e. a combination of tone and noise, as in *z*. In the production of normal vowels, the vibrator is normally provided by the vocal cords ; in the case of many consonant articulations, however, a source of air disturbance is provided by constriction at a point above the larynx, with or without accompanying vocal cord vibrations.

Despite the fact that the basis of all normal vowels is the glottal tone, we are all capable of distinguishing a large number of vowel qualities. Yet there is no reason to suppose that the glottal vibrations in the case of [ɑ:] are very different from those for [i:], when both vowels are said with the same pitch. The modifications in quality which we perceive are due to the action of the supra-glottal resonators which we have previously described. To understand this action, it is necessary to consider a little more closely the nature of the glottal vibrations.

3.32 It has already been mentioned that the glottal tone is the result of a complex, but mainly regular, vibratory motion. In fact, the vocal cords vibrate in such a way as to produce, in addition to a basic vibration over their whole length (the *fundamental* frequency), a number of overtones or *harmonics* having frequencies which are simple multiples of the fundamental or first harmonic. Thus, if there is a fundamental frequency of vibration of 100 cycles per second (cps), the upper harmonics will be of the order of 200, 300, 400, etc., cps. Indeed, there may be no energy at the fundamental frequency, but merely the

harmonics of higher frequency such as 200, 300, 400 cps. Nevertheless, we still perceive a pitch which is appropriate to a fundamental frequency of 100 cps, i.e. the fundamental frequency is the highest common factor of all the frequencies present, whether or not it is present itself.

The number and strength of the component frequencies of this complex glottal tone will differ from one individual to another and this accounts at least in part for the differences of voice quality by which we are able to recognize speakers. But we can all modify the glottal tone so as to produce at will vowels as different as [iː] and [ɑː], so that despite our divergences of voice quality we can convey the distinction between two words such as *key* and *car*. This variation of *quality*, or timbre, of the glottal tone is achieved by the shapes which we give the resonators above the larynx—the pharynx, mouth, and nasal cavity. These chambers are capable of assuming an infinite number of shapes, each of which will have a characteristic vibrating resonance of its own. Those harmonics of the glottal tone which coincide with the chamber's own resonance are very considerably amplified. Thus, certain bands of strongly reinforced harmonics are characteristic of a particular arrangement of the resonating chambers which produces, for instance, a certain vowel sound. Moreover, these bands of frequencies will be reinforced whatever the fundamental frequency. In other words, whatever the pitch on which we say, for instance, the vowel [ɑː], the shaping of the resonators and their resonances will be very much the same, so that it is still possible, except on extremely high or low pitches, to recognize the quality intended. It is found that the vowel [ɑː] has one such characteristic band of strong components in the region of 800 cps and another at about 1,100 cps. The vowel [iː] has bands of energy at about 280 and 2,500 cps.

3.33 This complex range of frequencies of varying intensity which go to make up the quality of a sound is known as the *acoustic spectrum*; those bands of energy which are characteristic of a particular sound are known as the sound's *formants*. Thus, formants of [ɑː] are said to occur in the region 800 and 1,100 cps. Such a complex wave pattern may be analysed by means of a number of instrumental techniques, most of which involve lengthy calculations; e.g. the number of fundamental vibrations may be counted on an oscillographic tracing. But recently the instrument known as the *sound spectrograph*, originally designed to render speech patterns visible to the deaf, has made possible a relatively rapid and visual presentation of the spectrum and the various intensities of sound contained in it. In an instrument of this sort in common use, a number of filters, covering a range of frequencies

from 0–8,000 cps, respond to the varying sound intensities at different frequencies and ultimately produce a length of paper giving a three-dimensional display of the acoustic spectrum : frequency is shown on the vertical axis, time on the horizontal, and the energy at any frequency level by means of the degree of blackening made by the tracing pen on the paper. Thus, the concentrations of energy at particular frequency bands (the formants) stand out very clearly (see Fig. 3 for the spectrogram of the vowels [ɑ:] and [i:]). Moreover, since the machine permits the analysis of an utterance lasting up to about $2\frac{1}{2}$ seconds, the ever-changing pattern of the spectrum of continuous speech (see Fig. 3) and the difficulty of dividing up an utterance into separate (discrete) segments are made obvious. Such spectrographic analysis provides a great deal of acoustic information in a convenient form and, in addition, the instrument itself is relatively simple for a phonetician without special training to operate. Nevertheless, much of the information given is, in fact, irrelevant to our understanding of speech and the phonetician is obliged to establish by other methods the elements of the spectrum which are essential to speech communication.

Although there remain a number of problems to be solved by spectrographic analysis, much is already known about the acoustic structure of *vowels*. For instance, two, or at the most three, formants appear to be sufficient for their correct identification. As far as the English vowels are concerned, the first three formants are all included in the frequency range 0–4,000 cps, so that the spectrum above 4,000 cps would appear to be largely irrelevant to the recognition of our vowels. It is true that on a telephone system, which may have a frequency range of about 300–3,000 cps, we find little difficulty in identifying the sound patterns used by a speaker and are even able to recognize voice qualities. Indeed, when we are dealing with a complete utterance in a given context, where, as we have seen above (pp. 3–4), there is a multiplicity of cues to help our understanding, a high degree of intelligibility may be retained even when there are no frequencies above 1,500 cps.

As one would suspect, there appear to be certain relationships between the formants of vowels and the cavities of the vocal tract (i.e. the shapes taken on by the resonators, notably the relation of the oral and pharyngeal cavities). Thus, the 1st formant appears to be low when the tongue is high in the mouth : e.g. [i:] and [u:], having high tongue positions, have 1st formants of the order of 280–300 cps, whereas [ɑ:] and [ɒ] have their 1st formants in the region 600–800 cps, their tongue positions being relatively low. On the other hand, the 2nd formant seems to be inversely related to the length of the front cavity : thus

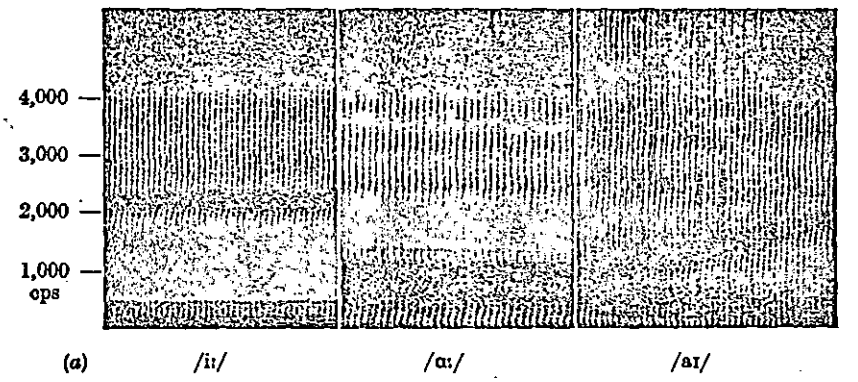
[i:], where the tongue is raised high in the front of the mouth, has a 2nd formant at about 2,500 cps, whereas [u:], where the tongue is raised at the back of the mouth and lips are rounded, has a relatively low 2nd formant around 900 cps.

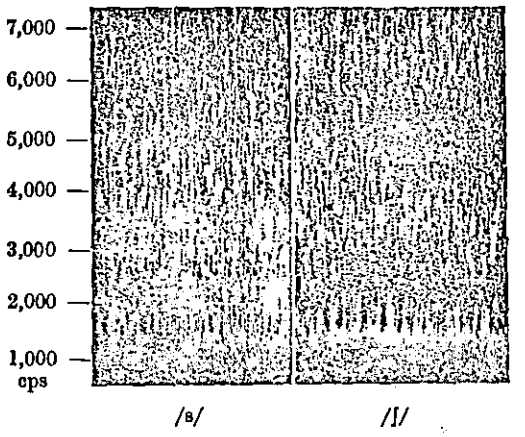
It is also confirmed from spectrographic analysis that a diphthong, such as that in *my*, is indeed a glide between two vowel elements (besides involving a perceptible articulatory movement), since the formants bend from those positions typical of one vowel to those characteristic of another (see Fig. 3).

Acoustic information concerning *consonant* articulations is at the moment less complete than that for vowels, but for many (e.g. the initial sounds in *pin, tin, kin, thin, fin, sin, shin*, in which the glottal vibrations play no part) there is an essential noise component, deriving from an obstruction or constriction within the mouth, approximately within the range 2,000–8,000 cps (see Fig. 3). This noise component is also present in analogous articulations in which the vocal note is present, as in the final sounds of *ruse* and *rouge* where, as we have seen, we are dealing with sounds which consist of a combination of glottal tone and noise. Relevant acoustic data concerning both vowel and consonant articulations will be given in the sections dealing with individual English sounds (Chapters 7 and 8).

Spectrographic analysis also reveals the way in which there tends, on the acoustic level, to be a merging of features of units which, linguistically, we treat separately. Thus, our discrimination of [f] and [θ] sounds would appear to depend not only on the frequency and duration of the noise component but also upon a characteristic bending of the formants of the adjacent vowel. Indeed, in the case of such consonants as [p, t, k], which involve a complete obstruction of the air-stream and whose release is characterized acoustically by a relatively brief burst of noise, the vowel *transition* between the noise and the steady state of the vowel appears to be of prime importance for our recognition of the consonant. This overlapping of vowels and consonants would suggest that an analysis of speech based solely upon acoustic data would find it convenient to operate with units larger than the sound segment.

3.4 Our perception of the *pitch* of a speech sound depends directly upon the frequency of vibration of the vocal cords. Thus, we are normally conscious of the pitch caused by the 'voiced' sounds, especially vowels; pitch judgments made on voiceless or whispered sounds, without the glottal tone, are limited in comparison with those made on voiced sounds, and are induced mainly by variations of





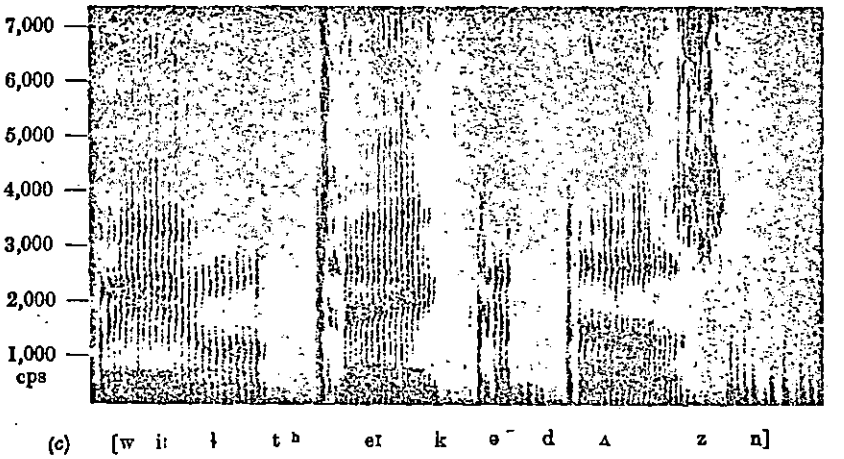


FIG. 3.—Spectrograms of /iː/, /ɑː/, /aɪ/, /s/, /ʃ/, and *We'll take a dozen.*

intensity or by the dominance of certain harmonics brought about by the dispositions of the resonating cavities.

The higher the glottal fundamental frequency, the higher our impression of pitch. A male voice may have an average pitch level of about 150 cps and a female voice a level in the region of 240 cps. The pitch level of voices, however, will vary a great deal between individuals and also within the speech of one speaker, the total range of a speaking voice being liable to have a range as extensive as 80–350 cps. Yet our perception of pitch change extends further than the limits of glottal fundamental frequency, since our recognition of quality depends upon frequencies of a much higher order. In fact, the human ear perceives frequencies from as low as 16 cps to about 20,000 and in some cases even higher. As one becomes older, this upper limit may fall considerably, so that at the age of fifty it may extend no higher than about 10,000 cps. As we have seen, such a reduced range is no impediment to perfect understanding of speech, since a high percentage of acoustic cues for speech recognition fall within the range 0–4,000 cps.

Our perception of pitch is not, however, solely dependent upon fundamental frequency. Variations of intensity on the same frequency may induce impressions of a change of pitch ; and again, tones of very high or very low frequency, if they are to be audible at all, require greater intensity than those in a middle range of frequencies.

As far as the variation of pitch in speech is concerned, it is by no means easy to measure the continuous changes. It is, of course, possible to count the fundamental vibrations within a selected segment in a trace such as is provided by an oscillograph. But, besides being a laborious process, such a method is likely to give us a pattern which is much more complicated than our auditory impression of pitch change. Our impression of pitch changes in speech results largely from fundamental frequencies carried by vowels and other voiced sounds, the voiceless sounds being discounted. With pitch, as with quality, an instrumental analysis does not make this human linguistic discrimination and may, therefore, provide us with a good deal of irrelevant information. Several instrumental procedures do, however, exist which go far towards showing us the grosser pattern which is relevant to linguistic work, e.g. a general pitch curve can be seen by tracing a harmonic on an expanded spectrographic display. The second harmonic is a convenient one to choose, since this is low enough usually not to be obscured by formant intensity nor very much deformed by formant transitions.

3.5 Our sensation of the relative *loudness* of sounds may depend on

several factors. A sound or syllable may appear to stand out from its neighbours—be 'louder'—because a marked pitch change is associated with it or because it is longer than its neighbours. It is better to use a term such as *prominence* to cover these general listener-impressions of variations in the perceptibility of sounds. More strictly, what is 'loudness' at the receiving end should be related to *intensity* at the production stage, which in turn is related to the size or *amplitude* of the vibration and the speaker's feeling for 'stress'. Generally speaking, it may be said that an increase in amplitude of vibration, with its resultant impression of greater loudness, is brought about by an increase of air-pressure from the lungs. When a sound or syllable is stressed, it is being uttered with more muscular effort, increased air pressure, and greater amplitude of vibration. As we shall see (§ 9.02), this greater intensity is not in itself usually the most important factor in rendering a sound prominent in English.

Moreover, all other things being equal, some sounds appear by their nature to be louder or more sonorous than others: e.g. the vowel in *barn* has more carrying power than that in *bean*, and vowels generally are more powerful than consonants.

The judgments we make concerning loudness are not as fine as those made for either quality or pitch. We may judge which of two sounds is the louder, but we find it difficult to express the extent of the difference. Indeed, in terms of our linguistic system, we need perceive and interpret only gross differences of loudness, despite the fact that when we judge quality we are, in recognizing the formant structure of a sound, reacting to characteristic regions of strong intensity in the spectrum.

3.6 In addition to affording different auditory impressions of quality, pitch, and loudness, sounds may appear to a listener to be of different *length*. Clearly, whenever it is possible to establish the boundaries of sounds or syllables, it will be possible to measure their duration by means of such traces as are provided by oscillograms or spectrograms. Such delimitation of units, in both the articulatory and acoustic sense, may be difficult, as we shall see when we deal with the segmentation of the utterance. But, even when it can be done, variations of duration in acoustic terms may not correspond to our linguistic judgments of length. We shall, for instance, refer later to the 'long' vowels of English such as those of *bean* and *barn*, as compared with the 'short' vowel in *bin*. But, in making such statements, we shall not be referring to absolute duration values, since the vowel of *barn* may in fact, in any given utterance, be very considerably longer than that of *bean*. In the English system, however, we know that only two degrees

it is important to be aware of the limitations of such listeners, so as to be able to make a proper evaluation of their judgments. A listener's reactions are normally conditioned by his experience of handling his own language. Thus, if there are only five significant vowel units in his language, he is liable to allow a great deal more latitude in his assessment of what is the 'same' vowel sound than if he has twenty. An Englishman, for instance, having a complex vowel system and being accustomed to distinguishing such vowels as those in *sit, set, sat,* will be fairly precise in his judgment of an *e* type of vowel. A Spaniard, however, whose vowel system is made up of fewer significant units, is likely for this reason to be disposed to be more tolerant of variation of quality, while still identifying a sound of the *e* type. Or again, if a listener is presented with a system of synthetic vowels which is numerically the same as his own, he is able to make allowance for considerable variations of quality between his and the synthetic system and still identify the vowels correctly—by their 'place' in a system rather than by their precise quality; this is what he does when he listens to and understands his language as used by a speaker of a different dialect. Tests involving synthetic speech and concerned with the recognition of absolute vowel qualities would have, therefore, to be presented to a listener in such a way that his own linguistic prejudices have the least possible influence on his judgments.

Our hearing mechanism also plays an important part in monitoring our own speech; it places a control upon our speech production which is complementary to our motor, articulatory habits. If this 'feed-back' control is disturbed, e.g. by the imposition of an artificial delay upon our reception of our own speech, disturbance in the production of our utterance is likely to result. Those who are born deaf or who become deaf before the acquisition of speech habits are rarely able to learn normal speech; similarly, a severe hearing loss later in life is likely to lead eventually to a deterioration of speech.

APPENDIX D

ROBSON'S DESCRIPTION OF THE ENGLISH ORCHESTRA OF THE LANGUAGE

The first step in this direction is to reduce the complexities of articulation qualities to a number of auditory impressions or casts of the voice upon hearing as though they were a set of physiological colors. From this viewpoint, the writer's point of view, the timbres of language are the *audible* casts of the voice in words such as: the boniness or hardness of the consonant *k* (carries a bony roof-of-the-mouth color); the liquid soft bubbliness of *b* (stamped with the interior moisture of bulbous-shaped lips); the harsh raucousness and the oral roar of the vocal chords and the throatal resonance of *r*; the almost pure, breathlike airiness of *h*; and the minute sharpness in the dental detonation of *t*. Further simplification of these casts of the voice in words is to classify them as phonetic instruments in a complete orchestra of the language whose intelligent use depends on the direction of the writer.

Please turn to Figure 2. Here the sounds of speech have been arranged in the image of a 41-piece orchestra waiting to be organized by the auditory imagination of the writer, the director, and the composer of language.

The illustration shows vowels, diphthongs, semivowels, and consonants arranged in eleven orchestral groups. These musical images will help the writer or speaker to visualize the letters of the alphabet as symbols of timbre and instruments of phonetic intelligence. Just as a symphonic orchestra puts the stringed instruments and the wood winds in the front rows and the percussion devices such as drums, triangles, cymbals back in the rear rows, so the orchestra of the language places the vowels that are tones in the front and the consonants that are noises in the back. Although this arrangement grew out of an analogy with music, it is based on the valid acoustic distinction between a tone and a noise. This auditory distinction leads the writer to look at the alphabet as he must look at it if he intends to use the timbres of speech to develop audio-scriptic intelligence.

The first three groups are tone divisions. They separate the

vowels into three tone levels: high, middle, and low. Since the vowels are the most potent creators of tone in language, this scheme prepares the writer to think about vowels and to hear their vibrations as notes in a vowel scale. Later, the vowel scale and its uses in writing will be discussed in detail. At present, the writer should keep one fact in mind. These "tones" are produced in the mouth, not in the throat or the vocal chords. The "high" vowels *ee*, *ĕ*, *ĭ*, etc., are made in the front one of the two chambers that the tongue makes by separating the hollow of the mouth into two resonating rooms. The tongue acts like a wall. The "low" tone vowels such as *oo*, *oh*, *aw* are produced by drawing back and dropping the tongue-wall so that the oral chamber is lengthened, deepened, and enlarged. The strongest overtones of the low-toned vowels come from the rear of the mouth, out of whose depths they carry to our hearing a feeling of lowness, largeness, and darkness. The "middle" vowels are compromises. This group includes the diphthongs *oi* and *ew* that fuse a low and high vowel, and the vowel *ah* whose oral structure lies midway between the high and low vowels.

Let us look again at the illustrated orchestra.

Behind the full-toned vowels in the first three groups appear the partial tones, the transient semivowels in Group 4: *w*, *y*, *r*, *l*, *m*, *n*, *ng*. The last three of these sounds, *m*, *n*, *ng*, resonate with vibrations in the nasal cavities that give them a humming quality. The hum of *m*, lower in tone level than *n*, is heard most clearly in words with low-toned vowels such as *hum*, *room*, *strum*, *mum*, *mumbles*, *murmur*, *mom*, *momentum*. When *m* borders on a low- or middle-tone vowel, its hum color is more powerful than in the high-vowel environment of words such as *maim*, *member*, *mammy*, *mimic*, *minimize*. The higher vowels break up sharply the low throb in *m*'s resonance and mask its hums. This observation suggests a tendency in many semivowels to sound clearer when they blend with vowels on the same tone level. The vowel sustains the altitude of the transient

tone. Otherwise, hearing is compelled to make too many adjustments in too short a period of time. We see here how delicate these qualities are and how necessary it is to use all the elements of auditory intelligence in a related pattern to make a definite impression on the ear.

The hum-timbre of *n*, higher in tone than *m*, is more impressive in high-voweled words: *keen, kin, pin, tin-pan, Dane, insane, winnow, wind, stain.* If we listen closely to the humming resonance of *ng* in *bang, cling, dung, wrung, ding-dong, clang, gong, whang, singsong,* the denseness of the nasal quality is noticeable. The peculiar, blurred, dense hums of *ng* are the effect of its closely packed overtones. The dense timbre of *ng* is uninfluenced by the tone levels of neighboring vowels.

There is something of the plasticity, softness, and liquidity of the tongue in the timbre of *l* that tinges words such as *lingual, loll, wallow, swill, lolli* in lollipop. When the low-tone level of *l* precedes a vowel and introduces a word, it conveys a sense of uplift in such expressions as *left, lift, lilt, laugh, lady, light, lick, leap, lip.* But the low-tone level of *l* may have an opposite effect when it follows a vowel or terminates a word. Then it can give the impression of a physical letdown, a leveling off, or a long drop: i.e., *lull, fall, tail, stale, spill, fallow, keel, dull, pall, shawl, kill, cool, sprawl.*

These examples should show the writer that the use of phonetic timbre is a creative opportunity, and not a gadgetlike gimmick of mechanical character. Its value is a potential for the writer's imagination, and not a guarantee.

Three of the more powerful semivowels in Group 4 of the illustrated orchestra are *y, r,* and *w.* The low-toned timbre of open, loose-lipped *w* sweeps its upcurve of breath out of the low chordal regions of resonance, and glides into vowels at high speed, and with force. The power and the tubal openness in the swift upswerve of this sound has creative possibilities for the writer. The timbre of *w* may be used to suggest the open-

flowing, the rhythmic, the forceful, or combinations of these qualities. These possibilities are conveyed in *win, wild, water, wings, wobble, waves, wind, wanton, whiff, warble, whistle, swish, swing, sway, willow, wail, will,* and (again with deference to Dizzy Dean), *swang.* The power and the conspicuous quality of *w* can put a definitive pulse into phrases, e.g., *the wishing well, wanton wenches, the Wailing Wall, wishy-washy, winning ways, wild women, werewolf.* The timbre of transient-toned *y*, with its speed and gliding power suddenly emitted from the high tone region in the mouth, strikes the ear with some of the chordal tension of the vowel *ee.* This explosion of high tension may be used to express the assertive, the abrupt, the final, or the swiftly rising impetus in *yield, youth, yank, yell, yip, yap, yes, yeast.* The throatal, low roar of *r*, full of pharyngeal friction, has a resonance that is reinforced by the sympathetic vibration of the tip or blade of the tongue, near the center of the hard palate; *r* carries its low-toned roughness, power, and its dark throat color into low-toned words such as *strong, storm, harsh, war, rough, burr, hard, rock, hoarse, char, gore, gorge, wrong.* The sense of irritation and power displayed in these words is partly due to the chordal friction and the tremendous intensity in the ear-color of *r.* Although *r* is low-toned, its timbre is not destroyed by gliding into high-toned vowels in words such as *leer, rage, rile, liar, cheer, rasp, rant, ram, wretch, bear, rape.* Here its power is evident. Engineering data on the acoustic dimensions of *r* show it to be the most powerful of all nonvowel speech sounds. That is the probable explanation of why the timbre of *r* is not destroyed by running into vowels on quite different tone levels.

The straight percussion sounds that strike the ear like drums appear in the noise section of the orchestra of the language, high up in the last row. They are the breath-exploding consonants *p, k,* and *t.* They do explode. The pentup air pressure suddenly released by the noise of *p* has a lip timbre; *t*, a teeth quality; and *k*, a hard texture that echoes from the bony roof of the

mouth. These are the swiftest sounds of speech, whose time durations are so tiny that the ear has little opportunity to detect tone. The timbre of the percussions in orchestral Group 11 is not influenced by the tones of adjacent vowels. The lip explosiveness of *p* sounds through *nipple, pucker, pop, pout, lip, pulp, flip, pipe*. The sharp, quick, abrupt, dental cast of *t* carries its keen timbre into *taut, meat, tear, cut, tattle, teeth, white, tart, tight, tack*; and its quality of the diminutive and the swift into *tot, trout, teat, tut-tut, titter*. The boniness of the roof of the mouth enters into the articulation of *k* and passes its quality of hardness on with speed and sharpness in *crack, truck, cluck, conk, kick, click, knock, creak, squeak, ack-ack, nick*.

If you try to pronounce *p, k, t* with resonance in the throat, your articulations begin to sound like *b, g, d*. These resonated plosives constitute orchestral Group 9. Their throatal vibrations lengthen the durations of *b, g, d*, and this extra time for hearing permits the ear to perceive more tone than in the snappy, unvoiced percussives, *p, k, t*. The timbre of *b* is impressed with a sense of the softness of bulbous-shaped lips. It has the quality of a relatively slow explosion moving softly, amorphously, with low tone through a wet bag. Words such as *lob, blimp, bubble, bobble, blab, blubber, sob, babble, boob, slob, burp, bib, bulb, babe* communicate a sense of full-blown moisture, or slowness, or lack of structure, or a combination of these characteristics. The source of *g* is about five times as far back from the surface of the front teeth as its relative *k*. The timbre of *g* is an explosive gutturalness shot through with power, duration, and indefiniteness in tone. When *g* occurs twice in a word it gives the effect of a guttural noise bouncing up and down like a stammering in the throat in *goggles, gag, groggy, gargle, gig, agog, gee-gaw*. These words are double acoustic images of the guttural colored emissions of breath that the back of the tongue explodes in the glottal regions. The sudden tongue-dropped, open-mouthed throatiness of *g* casts its timbre like a shadow of the

voice upon the expression of words like *gawk, gasp, gook, gape.*

The hissing consonants in orchestral Group 3 have the frictional quality of breath blowing tiny, aerial bubbles over the lips and tongue and through the teeth. They are *th, sh, f,* and *s.* Some of these sounds are long-durational and all of them have a soft texture, except *s,* which carries the sharp imprint of the teeth. Their timbre of fricative breath, beating in the ear like an atomized gas, pulses through words such as *fast, swift, slash, swish, hiss, fizzle, fuss, sift, soufflé, shuffle, scuffle.* The longer time duration of steady breath friction reveals its dimension in *stiff, staff, suffer, shaft, smooth, stuff.* Notice how the soft-lipped, air-rippled quality of *f* enters expressions such as *fluff, ruffle, foam, muff,* and how the carnivorous teeth emotionally color the breath with the sharpness and snap of *s* in *hiss, snip, sizzle, cuss, slash, sik'em.*

These examples can be multiplied to show that the timbre groups in the illustrated orchestra of the language will aid the writer to select more powerful, more sensuous, and phonetically more intelligent words. But the writer might remember this fact. The timbre of consonants and semivowels was usually considered in conjunction with the other dimensions of audio intelligence: time, tone, power. Although timbre is a tool for making language decisive, its impressions on hearing are limited. Sharp contours of timbre depend on all the dimensions of sound to communicate their full meaning. So does audio-scriptic intelligence.

We can draw these conclusions: The greater the number of sound dimensions the writer can control, the greater his potential for creating sound images rich with phonetic information. The greater the concentration of auditory signals throughout a flow of words written to produce a clear auditory effect, the greater the probability that the intended effect will be heard. Concentration is important. The reason for its importance is the weak energy of conversational speech.

APPENDIX E

TRIÁNGULO VOCÁLICO

La articulación de las vocales con arreglo a la posición de la lengua puede representarse por medio del triángulo ideado por el alemán Hellwag, 1781. En dicho triángulo, dispuesto de manera invertida, los vértices superiores van ocupados por la i (vértice palatal) y por la u (vértice velar), correspondiendo el vértice inferior a la vocal a. Entre la a y la i se colocan la e y las demás vocales intermedias palatales, y entre la a y la u, las velares. En el campo de dicho triángulo cabe asimismo situar el lugar de articulación de las vocales mixtas considerándolas por encima del punto de la a, entre la i y la u. Claro es que el punto de articulación de cada vocal se deduce considerando en conjunto la inclinación del cuerpo de la lengua en un sentido o en otro, pudiendo servir de guía en todo caso la mayor elevación que sea posible advertir en la curva del dorso de la lengua.

Navarro, op. cit., pp. 37-38.

APPENDIX F

DESCRIPTION OF THE SPANISH CONSONANTS ACCORDING TO T. NAVARRO TOMAS

71. TENSIÓN MUSCULAR. — El grado de tensión con que se articulan las consonantes españolas varía segun diferentes circunstancias, y principalmente según la posición del sonido con respecto al acento de intensidad; es más tensa, por ejmplo, la θ en *cierto*-θjérto, que en *certidumbre*-θɐrtiđúmbre, y asimismo la s en *jamás*-xamás, es más tensa que en *lunes*-lúnəs, etc. Las oclusivas p, t, k acentuadas se debilitan en muchos casos, convirtiéndose en fricativas más o menos sonoras: *eclipsar*-eklipsáɾ, *aritmética*-ariđmétika, *tecnicismo*-tɐgniθizmo, mientras que en casos semejantes estas mismas consonantes se pronuncian frecuentemente como oclusivas cuando las sílabas en que ellas se encuentran llevan acento fuerte: *eclipse*-eklipse, *ritmo*-rítmo, *técnica*-tɐ́knika. Entre las consonantes b, d, g, ŷ, r y sus fricativas correspondientes b, d, g, y, ɾ, las modificaciones del acento de intensidad producen ciertas vacilaciones de articulación, de las cuales iremos dando cuenta en los párrafos que tratan de la pronunciación de dichos sonidos.

72. LA TENSIÓN SEGÚN LA POSICIÓN DEL SONIDO EN EL GRUPO.—En igualdad de circunstancias respecto al acento, la tensión articulatoria de las consonantes varía tam-

bién sensiblemente según su posición en el grupo fonético. Los principales grados o matices que a este respecto conviene en la práctica saber distinguir son dos: *a)* posición inicial, en que la tensión articulatoria es relativamente fuerte, sobre todo si se trata de la posición inicial absoluta: *cima*-θime, *seda*-séđe, *racimo*-raθimo, *casero*-kaséro; *b)* posición final, en que la tensión es menor, sobre todo por lo que se refiere a la final absoluta: *bizco*-biθko, *pesca*-péske, *perdiz*-pɛrdiθ, *francés*-franθés. En el caso de una consonante fricativa final, ante pausa, *-s*, *-z*, *-d*, etc., no sólo disminuye la tensión muscular, sino también el impulso de la corriente espirada, resultando un sonido relajado con una fricación muy débil. Los extranjeros, no advertidos sobre este punto, dan de ordinario al sonido de las finales españolas *z* y *s* una fuerza y una duración excesivas. Las oclusivas *p*, *k* finales de sílaba, ante otra consonante oclusiva se reducen a articulaciones meramente implosivas y se pronuncian asimismo con tensión menor que en posición inicial: *indocto*-inḍokto, *concepto*-kɔnθépto, *aspecto*-aspékto, *adoptar*-aḍoptár, etc.; a veces, en estos casos, llegan también a pronunciarse simplemente como fricativas más o menos sonoras.

73. Oclusivas puras y oclusivas aspiradas. — La pronunciación de las consonantes p, t, k, iniciales de sílaba, resulta pura o aspirada, según el momento en que, terminada propiamente la articulación de la consonante, empiezan a vibrar las cuerdas vocales; en uno y otro caso la oclusión es igualmente sorda; la diferencia consiste en el modo de producirse la explosión. En las oclusivas, puras, apenas cesa el contacto de los órganos bucales, empiezan las vibraciones de la glotis, resultando la explosión completamente o en su mayor parte sonora: *padre*-pádrə, *todo*-tódo, *casa*-káse; en las oclusivas aspira-

das la sonoridad empieza un poco más tarde, percibiéndose durante la explosión un tenue soplo sordo, como una breve *h* aspirada, que se intercala entre la tensión de la consonante oclusiva y el sonido siguiente: *padre*-p^{h}ádre, *todo*-t^{h}ódo, *casa*-k^{h}ása. La pronunciación correcta española emplea únicamente las formas oclusivas puras [1]; las formas aspiradas, frecuentes entre alemanes e ingleses, deben evitarse cuidadosamente [2].

74. Oclusivas sonoras. — En las oclusivas sonoras b, d, g, iniciales absolutas, las vibraciones laríngeas empiezan normalmente en español unas seis o siete centésimas de segundo antes de la explosión. Algunos extranjeros, alemanes e ingleses principalmente, pronuncian estas consonantes en dicha posición con vibraciones laríngeas demasiado tardías o demasiado débiles, de modo que oyéndoles decir, por ejemplo, *baño, bollo, doma, deja, gasto, goma,* resulta para nuestro oído casi como si dijesen *paño, pollo, toma, teja, casto, coma.* Para adquirir la pronunciación española, que en este punto coincide, en general, con la francesa, con la italiana y con la de las demás lenguas neolatinas, debe moderarse un poco la tensión muscular y debe procurarse, ante todo, que las

[1] Sobre la duración, intensidad y sonoridad de la explosión de estos sonidos y sobre algunas pequeñas diferencias que entre ellos aparecen, véase S. Gili, *Algunas observaciones sobre la explosión de las oclusivas sordas,* en *Revista de Filología Española,* 1918, V, 45-49.

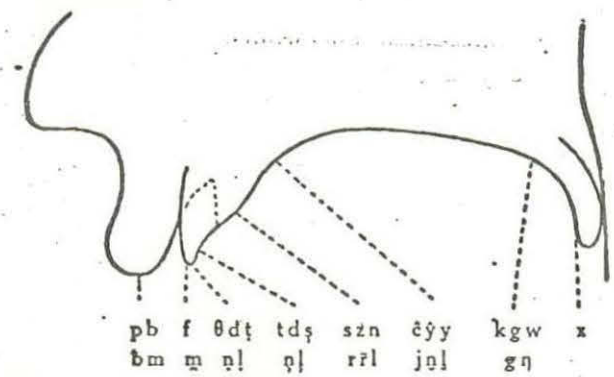
[2] Pronunciando palabras como *papa, tapa, pata, capa,* etc., un papel de fumar o la llama de una cerilla mantenidos a poca distancia de los labios, experimentarán, con las oclusivas aspiradas, una sacudida brusca y violenta; mientras que en pronunciación correcta española la salida del aire durante la explosión de dichas consonantes sólo produce en el papel o en la llama un movimiento pequeño y suave.

vibraciones laríngeas sean claramente perceptibles antes de la explosión de la consonante, dejando al mismo tiempo una cierta elasticidad a las paredes bucales para que en su cavidad cerrada pueda almacenarse el aire que se escape por la glotis durante la oclusión. Entretanto, el velo del paladar debe mantenerse elevado, impidiendo la salida del aire por las fosas nasales, pues el poner m, n, ŋ delante de b, d, g, respectivamente, como en algunos libros se aconseja, a fin de que estas últimas resulten sonoras, sólo es una torpe e inaceptable imitación de la pronunciación normal.

75. Las fricativas b, d, g. — Tres articulaciones particularmente características de la lengua española son las que se representan fonéticamente con los signos b, d, g; se trata de su pronunciación en los §§ 81, 100 y 127; son articulaciones generalmente desconocidas en francés, en inglés y en otros muchos idiomas; en español, por el contrario, son tan frecuentes que apenas hay frase en que no aparezcan varias veces, siendo muchos los casos en que algunos de estos sonidos se repiten o se combinan aun dentro de una misma palabra: *obligado*-obligádo, *agradable*-agradábla, *avinagrado*-abinagrádo, *comedido*-komədído, etc. Su uso es, sin duda, en nuestra pronunciación mucho más frecuente que el de las oclusivas b, d, g, con las cuales alternan de una manera regular, sin llegar a confundirse con ellas sino en casos excepcionales; pero el hecho de ir representadas unas y otras en la escritura corriente por los mismos signos *b, d, g*, y sobre todo el abandono en que, en general, se halla en nuestras escuelas el estudio de la pronunciación, hacen que tales fonemas, no obstante su importancia, sean comúnmente ignorados o mal conocidos aun por aquellos que se dedican a la enseñanza del idio-

ma. En cuanto a los extranjeros que pretendan hablar español, puede asegurarse que sin el dominio de estos sonidos su lenguaje se hallará siempre muy lejos de la pronunciación española correcta.

76. Otros sonidos españoles que no tienen representación en la ortografía corriente. — Entre los sonidos que se describen a continuación, hay varios — m̥, n̥, l̥, t̥, z, ɾ̥, y, ŋ — que, como las consonantes fricativas b̆, d̆, ğ, se usan inconscientemente; algunos, como t̥, l̥, no aparecen con mucha frecuencia; otros, en cambio, como z, ɾ, n̥, ŋ, etc., son, sin duda, mucho más abundantes; pero todos ellos, dentro de las circunstancias que a cada uno corresponden, se producen de una



Esquema de la clasificación de las consonantes españolas según el punto de articulación.

manera constante y regular; su enumeración en un catálogo de los sonidos españoles resulta, por consiguiente, indispensable.

77. *H* muda. — En la pronunciación correcta española, la *h* no representa la aspiración laríngea que en otros idiomas le corresponde y que en determinados

casos tuvo también, en otro tiempo, en nuestra lengua; la *h* ortográfica es actualmente en nuestra escritura una letra muda sin ningún valor fónico: *hoja*-o̧xe, *ahora*-ạo̧ra, *alcohol*-ạlkǫ̣l, *huerta*-wę̣rte, *hueco*-wėko, *ahuecar*-awekáı, etc.; la antigua aspiración aparece aún, sin embargo, en palabras como *humo*, *horno*, etc., en pronunciación dialectal [1].

APPENDIX G

TABLE I

STRIKING POWER

Consonants \ Vowels	30 ai	30 iai	30 oi	29 o	28 a	28 ou	28 ei	28 iu	27 uai	26 iei	25 e	25 eu	25 uei	25 au	24 i	24 uy	20 u
/r̄/ rr,r																	
/r/ r																	
/l/ l			MOST		POWERFUL												
/g/ gu																	
/w/ hu, u																	
/ĉ/ ch																	
/ñ/ ñ																	
/n/ n																	
/y/ hi																	
/j/ i																	
/m/ m																	
/n/ n																	
/l/ ll					MIDDLE		AREA										
/d/ d																	
/ŷ/ y,hi																	
/g/ g																	
/k/ k,c,q																	
/t/ t																	
/f/ f																	
/b/ b								AREA	OF	MINIMUM	POWER						
/d/ d																	
/p/ p																	
/s/ s,z																	
/x/ j,g																	
/θ/ z																	

TABLE II

TONE VOCABULARY

	High Vowels					Middle Vowels						Low Vowels					
	26	24	22	21	19	12	10	10	8	6	6	6	6	4	4	2	2
Consonants	iai	iei	ei	i	uy	ai	oi	iu	e	a	uai	au	o	eu	uei	ou	u
High																	
s																	
ch	HIGH TONE SOUNDS																
ŷ																	
y																	
j																	
Middle High																	
d	MIDDLE HIGH TONE SOUNDS																
n																	
t																	
Middle Low																	
rr	MIDDLE LOW TONE SOUNDS																
r																	
w																	
g																	
θ																	
d																	
Low																	
f																	
p	LOWEST TONE WORDS																
b																	
m																	
l																	

TABLE III

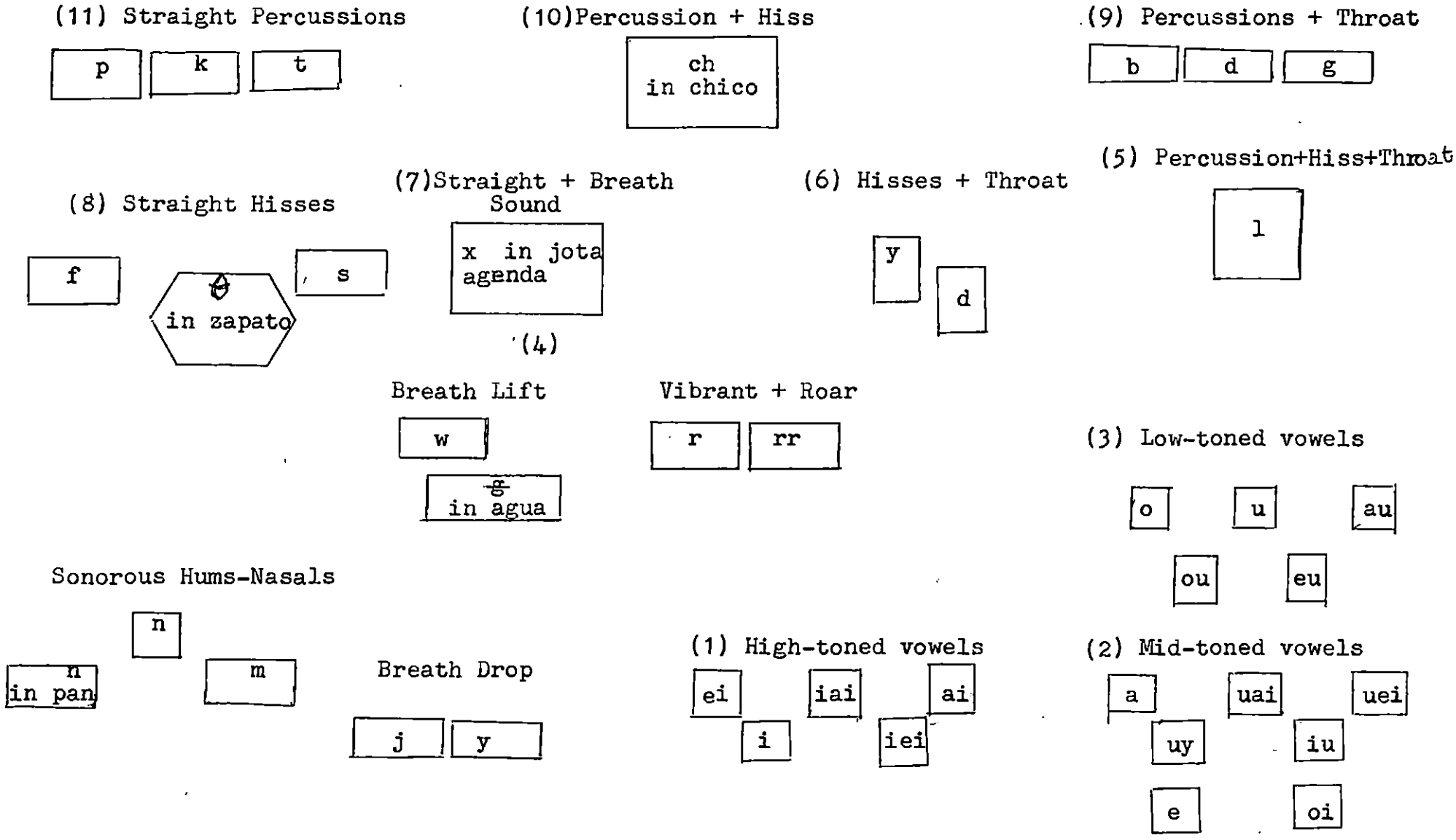
TIME DURATION OF THE SPANISH SOUNDS

AVERAGE TIME OF GROUPS OF VOWELS AND CONSONANTS OF SPANISH SPEECH AT A SPEAKING RATE OF 148 WORDS PER MINUTE

Group 1 Average 0.02 Sec.		Group 2 Average 0.07 Sec.		Group 3 Average 0.12 Sec.		Group 4 Average 0.14 Sec.		Group 5 Average 0.17 Sec.		Group 6 Average 0.22 Sec.	
p	0.02	d	0.06	ŋ	0.10	r̄	0.14	iu	0.16	eu	0.20
r	0.02	g̃	0.06	n	0.10	o	0.14	uy	0.18	ai	0.21
		y̆	0.07	/x/	0.11			ei	0.19	ou	0.22
		l̃	0.08	m	0.13			au	0.19	iei	0.23
		ñ	0.06	s	0.12					uei	0.22
		w	0.07	k	0.11					uai	0.24
		/j/	0.06	f	0.13					iai	0.25
				ch	0.13					oi	0.21
				u	0.10						
				i	0.12						
				a	0.12						

APPENDIX H

THE SPANISH ORCHESTRA OF LANGUAGE



APPENDIX I

FRAY LUIS DE LEON'S POEM

VIDA RETIRADA

1 ¡Qué descansada vida
la del que huye el mundanal ruïdo,
y sigue la escondida
senda por donde han ido
los pocos sabios que en el mundo han sido!

2 Que no le enturbia el pecho
de los soberbios grandes el estado,
ni del dorado techo
se admira, fabricado
del sabio Moro, en jaspes sustentado.

3 No cura si la fama
canta con voz su nombre pregonera,
ni cura si encarama
la lengua lisonjera
lo que condena la verdad sincera.

4 ¿Qué presta a mi contento
si soy del vano dedo señalado,
si en busca de este viento
ando desalentado,
con ansias vivas, con mortal cuidado?

5 ¡Oh monte, oh fuente, oh río,
oh secreto seguro deleitoso!
Roto casi el navío,
a vuestro almo reposo
huyo de aqueste mar tempestuoso.

6 Un no rompido sueño,
un día puro, alegre, libre quiero;
no quiero ver el ceño
vanamente severo
de a quien la sangre ensalza o el dinero.

7 Despiértenme las aves
con su cantar sabroso no aprendido,
no los cuidados graves
de que es siempre seguido
el que al ajeno arbitrio está atenido.

8 Vivir quiero conmigo,
gozar quiero del bien que debo al cielo,
a solas, sin testigo,
libre de amor, de celo,
de odio, de esperanzas, de recelo.

9 Del monte en la ladera
por mi mano plantado tengo un huerto,
que con la primavera
de bella flor cubierto
ya muestra en esperanza el fruto cierto.

10 Y como codiciosa
por ver y acrecentar su hermosura,
desde la cumbre airosa
una fontana pura
hasta llegar corriendo se apresura.

11 Y luego sosegada
el paso entre los árboles torciendo,
el suelo de pasada
de verdura vistiendo
y con diversas flores va esparciendo.

12 El aire el huerto orea
y ofrece mil olores al sentido,
los árboles menea
con un manso ruïdo
que del oro y del cetro pone olvido.

13 Ténganse su tesoro
los que de un falso leño se confían;
no es mío ver el lloro
de los que desconfían
cuando el cierzo y el ábrego porfían.

14 La combatida antena
cruje, y en ciega noche el claro día
se torna; al cielo suena
confusa vocería,
y la mar enriquecen a porfía.

15 A mí una pobrecilla
mesa de amable paz bien abastada
me basta, y la vajilla
de fino oro labrada
sea de quien la mar no teme airada.

16 Y mientras miserablemente se están los otros abrasando
con sed insaciable
del peligroso mando,
tendido yo a la sombra esté cantando.

17 A la sombra tendido,
de hiedra y lauro eterno coronado,
puesto el atento oído
al son dulce, acordado,
del plectro sabiamente meneado.