PERCEPTION OF VOWEL NASALIZATION IN FRENCH

by

André Gustave Lafargue

B.A.,B.Sc., Memorial University of Newfoundland, 1971

M.A., Memorial University of Newfoundland, 1975

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

Department of Paediatrics

Division of Audiology and Speech Sciences

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

May, 1976

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my Department or by his representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of PAEDIATRICS Div. of Audiology & Speech Sciences

The University of British Columbia
2075 Wesbrook Place
Vancouver, Canada
V6T 1W5

Date June 4, 1976

ABSTRACT

The present study investigates the perception of vowel nasalization in French. Previous studies had shown that a nasalized vowel in a non-nasalized context was produced with the velopharyngeal port closed for part of the duration of the vowel. Similarly, it was found that an oral vowel in a nasalized context was produced with the velopharyngeal port open for part of the duration of the vowel. Single cycles were spliced out of vowels of both types at times corresponding to various and known velar heights. They were used to construct the stimuli of two listening tests. Each test also contained non-manipulated reference vowels (oral and nasalized) as controls to evaluate the listeners' responses. Native speakers of French were asked to rate, on a 5-point scale, the degree of nasality of each stimulus.

There seems to exist a non-monotonic relationship between curves representing the listeners' judgement of nasality and those representing the timing of velar movement. Results also suggest that a rather large velopharyngeal cross-section may be necessary for a French listener to perceive an open vowel such as [ɑ] as nasalized, whereas a smaller velopharyngeal cross-section may be sufficient to give a close vowel such as [ɛ] and [ɔ] a nasalized quality. It also appears that the listeners perform some kind of time integration of the nasality function in order to judge whether a vowel is nasalized or not.

An acoustic analysis was carried out, using an analysis-by-synthesis method. It shows that different changes occur in the vowel spectrum as the velum lowers e.g., the first and second formant come close together and at least one zero is present when the velopharyngeal port is open. The position and movement of this zero in the spectrum suggest that there is a relation between the position of the zero on the frequency axis on the one hand, and size of the velopharyngeal cross-section and oral vocal tract on the other. The acoustic analysis also suggests that there exists some relation between the presence of a zero in the spectrum of vowels, the corresponding position of the velum and the perception of nasalization. Possible experiments are suggested for further research.

TABLE OF CONTENTS

Chapter Page

ABSTRACT ii
TABLE OF CONTENTS iv
LIST OF TABLES vii
LIST OF FIGURES. viii
ACKNOWLEDGEMENT. ix

1. INTRODUCTION. 1
2. LITERATURE REVIEW 6
   2.1 Introduction 6
   2.2 Direct Methods 6
      2.21 Cinefluorography. 6
      2.22 Fiberscope. 7
   2.3 Indirect Methods 9
      2.31 Aerodynamic Methods 9
      2.32 Electromyography. 11
      2.33 Nasograph 12
      2.34 Ultrasounds 12
   2.4 Results Obtained in the Literature 13
      2.41 Results Obtained by Use of Single Methods. 13
      2.42 Results Obtained by Use of a Direct Method Combined with EMG. 14
      2.43 Results Obtained by Use of Cinefluorography and Aerodynamic Parameters 16

Chapter Page

2.5 Perception of Nasality in the Literature 17

3. MATERIALS AND METHODS 21

3.1 Stimuli 21

3.2 Selection of Stimuli Utterances. 21

3.3 Vowel Editing. 23

3.4 Test Preparation 26

3.5 Subjects 27

3.6 Test Procedure and Equipment 28

4. RESULTS 29

4.1 Data Analysis. 29

4.11 Storage and Sorting of the Data . . . 29

4.12 Normalization 31

4.13 Criteria for Selecting Best Subjects 34

4.2 Analysis of Graphs 42

4.21 General Observations. 42

4.22 Analysis of Individual Vowel Types 42

4.23 General Trends. 44

4.3 Correlation of Results with Data of Benguerel et al. (1975a) 49

4.4 Acoustic Analysis. 56

5. SUMMARY AND DISCUSSION 62

Chapter Page

BIBLIOGRAPHY 69

APPENDIX - Instructions. 72

LIST OF TABLES

Table		Page
I.	Sorted data for Subject #10	30
II.	Normalized scores for Subject #10's responses to all five and to the last four presentations of each stimulus	33
III.	m_4 quartile differences for criterion 2	36
IV.	Stimuli preceding each of the last four presentations of the reference vowels in both tests and their average score (27-subject group) re: 3	45
V.	Subjects' s_4's for the reference vowels in both tests	47
VI.	Comparison of the results of Test 1 (27-subject group) with the data of Benguerel et al. (1975a)	51
VII.	Comparison of the results of Test 2 (27-subject group) with the data of Benguerel et al. (1975a)	52

LIST OF FIGURES

Figure		Page
1a	Block diagram of apparatus for digitization of speech signal	24
1b	Block diagram of apparatus for production of test items.	24
2	Average scores (five responses) in Test 1. . .	38
3	Average scores (five responses) in Test 2. . .	39
4	Average scores (last four responses) in Test 1	40
5	Average scores (last four responses) in Test 2	41
6	Time correspondence between the five ɛ̃ stimuli and the velar height curve of Benguerel et al. (1975a).	50
7	Spectra of stimuli obtained by analysis-by-synthesis.	58

ACKNOWLEDGEMENT

I would like to express my feelings of gratitude to the many people who have had a part in this thesis. In particular, I would like to thank:

- André-Pierre Benguerel, my friend and supervisor, who patiently guided me through the meanders of the experiment and taught me (believe it or not!!) the art of exactitude.

- John Gilbert for serving on my committee.

- All my subjects.

- Ingrid Jeffrey and Carolyne Pederson for going over the manuscript.

- All of the students, faculty and staff of AUS for their support in many forms during the past two years.

- Last, but not least, my parents and Mamie for always being there when I needed them.

Chapter 1

INTRODUCTION

Most human languages use sounds to carry information from a speaker to a listener. Each language possesses its own phonetic inventory which consists of a set of speech sounds. Some languages, for instance, use differences in the patterning of man's upper respiratory system (everything else being equal) to differentiate speech sounds. The two channels via which the air flows in and out of the vocal apparatus (i.e. the nasal and oral cavities) may be used to produce different types of speech sounds. Peterson and Shoup (1966, p. 60) list four different configurations of the respiratory system which may be used to produce speech sounds:

"1. Oral is a configuration of the vocal tract in which there is a velopharyngeal closure.

2. Nasalized is a configuration of the vocal tract in which there is an intermediate velopharyngeal opening and in which there is not an oral closure.

3. Nasoral is a configuration of the vocal tract in which there is a relatively large velopharyngeal opening and in which there is not an oral closure.(...).

4. Nonoral is a configuration of the vocal tract in which there is an oral closure but not a velopharyngeal closure."

The fourth configuration listed by Peterson and Shoup (1966) is most often described as "nasal" in the literature.

Some or all of these configurations are used by various languages of the world in their phonetic inventory. In particular, some languages use nasalized (or nasoral) vowels which are phonemically distinct from their oral counterparts. French is such a language; its phonetic inventory not only contains nasal consonants but also nasalized vowels which are phonemically distinct from oral vowels: for example, the two words "tête": [tɛt] (head) and "teinte": [tɛ̃t] (colour) differ only by their vowels, an oral one in "tête", a nasalized one in "teinte". The main difference between the two vowels is the presence of an open velopharyngeal port in the production of the latter.

The main articulator which controls the opening and closing of the velopharyngeal port is the soft palate or velum. When the velum is lowered, the velopharyngeal port is open and the air can flow through both the nasal and the oral cavity if the mouth is open. The velum, like most other speech articulators, has been studied by speech scientists in an attempt to reach a better understanding of the human communication process. Different methods have been designed and used to study the temporal course of velar opening during speech. They will be reviewed in the next chapter. One of the general findings about velar position concerns the relationship which exists between velar height and the degree of opening of the vowel: open vowels are produced with a lower velum position than close ones (Fant, 1960, p. 139).

Velar movement in connected speech has also been investigated. When vowels and consonants are uttered in various contexts, the velum moves for the production of the different phones and depending on the contexts, different velar activities are observable. Velar movement in French has been studied by different methods (Benguerel 1974, Benguerel et al. 1975a,b). In a study of nasal airflow Benguerel (1974, p. 111) found that

> "There is some similarity between velar coarticulation for nasal consonants and that for nasalized vowels, namely that velar opening gesture and nasal airflow start well ahead of the nasal sound itself."

In an attempt to further investigate this similarity, Benguerel et al. (1975a) used a fiberscope to study velar movement when oral and nasalized vowels are uttered in various consonant environments. They found that, in French, when phonemically oral vowels are uttered before nasal consonants, such as in /ɑnɑ/, the velopharyngeal port starts opening before the onset of the consonant even though it does not completely open until after the onset of the consonant. These same authors (ibid., p. 71) also found that phonemically nasalized vowels are normally produced with an open velopharyngeal port, although in cases where the phonemically nasalized vowels are between two voiceless stops such as in /tɑ̃t/,

> "the opening of the velopharyngeal port starts after the onset of the vowel (30-60 ms) and ends slightly (10-20 ms) before the beginning of the second consonant."

However, when spoken in the above-mentioned contexts as well as other similar ones, the vowels /ɑ/ and /ɑ̃/ are heard by French listeners as being phonemically oral and phonemically nasalized respectively. In other words, phonemically oral vowels may be uttered with the velopharyngeal port partially open during their production and phonemically nasalized vowels may be uttered with the velopharyngeal port partially closed during their production. The following questions thus arise, the answers to which will be the aim of this particular study:
1) With how much velopharyngeal port opening can an oral vowel be produced by a French native speaker to still be perceived as phonemically oral by a French native listener? 2) Alternatively, since the velum position changes rapidly through the vowel, does the listener perform some kind of time integration to identify vowel sounds as being oral or nasalized? 3) If so, is it the "average velar passage which matters; or rather the average value of some auditory function depending on the size of this passage?" (Benguerel et al., 1975a, p. 71).

The first question will be investigated by means of a perception test designed and conducted so that different vowels associated with different velopharyngeal port openings will be presented to French listeners who will be asked to evaluate subjectively the oral or nasal nature of the test stimuli. The reader is referred to Chapter 3 of this thesis for a more detailed description of the procedure. It is necessary to use native speakers of French for the listening task since French, the language investigated, contains nasalized vowels which are phonemically different from oral vowels, and most other languages do not. The French listeners should be able to associate a stimulus with a nasalized or an oral vowel which they use in their speech. Non-native speakers, whose language does not contain phonemically distinct nasalized vowels, would not be expected to perform properly because vowel nasalization does not play an important part in their speech. Therefore, they would not have any intrinsic reference points to establish criteria of nasality and orality.

The results of the perception test may provide some clues concerning the factors which affect the listeners' subjective judgement of nasality. Since the original recordings as well as the velar height measurements of the experiments of Benguerel et al. (1975a) are available, one should be able to correlate the listeners' judgements of oral versus nasalized vowel with a certain degree of velopharyngeal port opening. In particular, one should be able to determine at what point during vowel production in the contexts mentioned above listeners are able to associate a nasalized quality with the vowel they perceive.

Chapter 2

LITERATURE REVIEW

2.1 Introduction

This chapter is a survey of the different techniques and methods used in the study of velar movement during speech. It will also include some of the experimental findings relevant to this investigation.

There are many possibilities for studying velar movement. It is possible to distinguish between direct methods, where one actually observes velar movement, and indirect methods where one uses other parameters to draw conclusions concerning velar activity. Some experimenters have combined two or more methods simultaneously in order to gain more information and to correlate the results obtained by themselves and others using separate methods.

2.2 Direct Methods

2.21 Cinefluorography

This method and its relevant equipment are described in Moll (1960). Essentially, the major components of the equipment are an X-ray tube, an image intensifier tube, a fluoroscopic screen and a camera. Lateral X-rays of the subjects are made at a camera speed which can vary from 24 to 150 frames per second. The distances between structures (e.g.

between velum and pharyngeal wall) are then measured on enlargements of each frame. Such a method presents several problems:

i) Though modern techniques such as the one designed by Fujimura et al. (1968) may substantially reduce the amount of radiations, subjects are generally exposed to a high rate of radiation. This forces the experimenter to limit the size of the corpus since any one subject can only be exposed to a very limited amount of radiation before hazard level is reached.

ii) Data processing of film data is done frame-by-frame which is time consuming. Moll (1960) estimates that 35 frames at most can be measured and analysed in one hour.

iii) It gives only a mid-sagittal plane picture of the subject's head. Therefore, it is very hard to determine whether the velopharyngeal port is actually open.

The velum is a three dimensional organ moving in a three dimensional cavity. The experimenter using X-rays is able to observe and measure only two of these dimensions, and may only guess at what is happening in the third dimension: the cross-section of the velopharyngeal port for instance, is only known by its sagittal plane dimension.

2.22 Fiberscope

The fiberscope was designed by Sawashima and Hirose (Sawashima and Hirose, 1968) at the University of Tokyo, for observation of the larynx. It consists of two bundles of very fine glass fibers; one bundle of fibers carries light from a light source to the organ being studied, while the other bundle of fibers carries the image of the organ to a camera.

Ushijima and Sawashima (1972) used a wide angle fiberscope which was inserted through the nose of the subject so that the nasal surface of the velum could be observed and filmed. After filming calibration sentences for two different positions of the fiberscope, they were able to correlate different measurements of velum position with the amount of opening of the velopharyngeal port.

Using this method, Benguerel et al. (1975a, p. 67) studied the timing of velar height in French. They found that:

> "while the processing of data obtained by this method is relatively slow, and thus not very favorable to the gathering of a large corpus, it is practically free of experimental artifacts; in particular, the fiberscope does not interfere in any way with velar movement since it does not extend beyond the moving part of the palate."

This method never allowed them to determine precisely the size of the velopharyngeal opening. They could set a criterion for velar height corresponding to an opening or a closure of the velopharyngeal port, but could not associate a number with the velopharyngeal opening.

2.3 Indirect Methods

2.31 Aerodynamic Methods

Measurements of aerodynamic parameters provide information about velopharyngeal opening and velar movement. Various devices are used to pick up variations in nasal and oral airflow, and simple or differential pressures. These devices are connected to transducers which enable the experimenter to observe and record the fluctuations of the different parameters during speech. Some conclusions about velar movement can then be drawn. Certain experimenters study velar movement by observing one parameter only. For example, Benguerel (1974) observed nasal airflow only during the production of phonemically oral and nasalized French vowels in different contexts.

However, measurements of the variations of several parameters may provide more information.

Warren and Dubois (1964) using the equation of continuity, the energy equation, and assuming that the area of an orifice can be determined if the differential pressure across the orifice is measured simultaneously with the rate of airflow through it, derived the following equation, which takes into account the several details of turbulent, non-uniform and rotational airflow by introducing a correction factor k:

$$A = \frac{\mathring{V}}{k\sqrt{2\,\frac{\Delta P}{D}}}$$

where A = velopharyngeal orifice area in cm^2

$\mathring{V}$ = nasal orifice airflow in cm^3 per sec

ΔP = differential static pressure at the velopharyngeal orifice in dynes per cm^2

D = density of the air, 0.001 gm per cm^3

k = correction coefficient (0.65)

This formula can be used in an attempt to determine the orifice area in real time. The different parameters of the equation can be determined by using the method designed by Benguerel (1974). Differential static pressure around the orifice can be found using two small tubes connected to a differential pressure transducer and inserted through the nasal meatus of the subject and located respectively just above and just below the velopharyngeal opening. The different parameters can be fed in real time to a computer to determine the cross-section A in real time.

The primary drawback of all these methods which involve measurements of aerodynamic parameters is that instrumentation interferes to some extent with the articulation process. Secondarily, these methods are very prone to artifacts. For instance, a negative nasal airflow may be observed at the end of a vowel preceding a nasal consonant (Benguerel 1974, p. 109). This negative air flow is attributable to an increase of volume of the nasal cavity due to the initiation of the lowering of the velum when the velopharyngeal port is still closed. However it does not indicate an ingressive airflow at the velopharyngeal port.

2.32 Electromyography

Electromyography (EMG) has been used in speech research since the late 1950's. Using this method one may observe activity of the numerous muscles involved in speech production. Studies on velar movement have been carried out using several types of electrodes and various data analysis techniques. Hirose (1971) describes the newest instruments and Kewley-Port (1973) gives a description of the most recent techniques of EMG data processing. Electrodes are inserted into the muscle to be studied under light topical anaesthesia. As the muscle is activated, potentials are produced which are picked up by the electrodes, amplified and recorded. In the case of velar movement, the two muscles of main interest are generally the levator veli palatini and the palatoglossus.

One of the main problems encountered in using EMG is that the potentials produced are very small; moreover the electrodes pick up a certain amount of physiological noise. In order to increase the signal-to-noise ratio and have a clearer picture of the actual muscle activity it is necessary to record and average the muscle potentials obtained for several tokens of the same utterance. This method is therefore time consuming even though a computer is used to average the signals.

2.33 Nasograph

The nasograph is a simple device designed by Ohala (1971) which enables one to monitor the velopharyngeal opening in real time. It consists of a d.c. light source at one end of a small plastic tube. The tube is inserted in the subject's nose. A light sensor is inserted in the same tube so that "it rests approximately over the junction between the hard and soft palate when the light is well down into the pharynx, approximately at the level of the epiglottis or slightly above" (Ohala 1971, p. 2). The amount of light impinging on the light sensor and accordingly the voltage developed by it are taken as an indication of the size of the velopharyngeal opening.

Ohala (1971) mentions that it is difficult to establish a calibration of the relationship between the velopharyngeal port cross-section and the nasograph output voltage because light might impinge on the sensor by reflection as well as directly; in addition, the plastic tube might move during the pronounciation of the utterance. But as he points out (ibid., p. 3): "at any rate it ought to give a fairly reliable indication of relative palatal opening and of the timing of palatal movements."

2.34 Ultrasounds

Movements of the lateral pharyngeal wall also play an important part in the closure of the velopharynx. During speech the lateral pharyngeal wall usually moves medially. The Doppler ultrasound method by which researchers have investigated these movements is described in Minifie et al. (1968) and in Kelsey et al. (1969). It is based on the fact that an ultrasound beam reflected from a moving body undergoes a Doppler frequency shift. Calculations of the corresponding velocity from individual periods of Doppler signal enables the experimenter to calculate the relative displacement of that moving body. Since the ultrasonic transducer is placed externally on the lateral neck wall, this method does not require the insertion of any tube or wire in the vocal tract nor does it pose a radiation hazard to the subject. It is still, however, at the experimental stage.

2.4 Results Obtained in the Literature

2.41 Results Obtained by Use of Single Methods

Moll and Daniloff (1971, p. 681) used cinefluorography to study velar movement during connected speech in English. Their corpus consisted of several combinations of phone sequences. They found that, in English

> "movement toward velar closure is initiated long before the articulatory contact for the non nasal consonant, usually during the preceding nasal".

The major finding of their study was that

> "in sequences in which a nasal consonant is preceded by one or two vowel sounds, the velar-opening gesture for the nasal is initiated near the beginning of primary articulatory movement toward the first vowel in the sequence." (ibid., p. 683)

Ohala (1971) using the nasograph to compare utterances where velum activity occurred, noticed a lowering of the velum for low vowels. His findings also suggest that "the soft palate begins to lower for an upcoming nasal consonant as soon as it can, i.e. as soon as it is no longer required to be closed for an obstruent". (Ohala, 1971, p. 1)

2.42 Results Obtained by Use of a Direct Method Combined with EMG

EMG has been used simultaneously with cinefluorography and with fiberoscopic cinematography to correlate muscle activity with movements of the velum.

Lubker (1968, p. 4) by placing electrodes "in the middle third of the soft palate, at the approximate level of the palatal levator dimple" observed variations in palatal EMG activity during speech which appear to be positively correlated with velopharyngeal opening.

Fritzell (1969, p. 71) found similar results and concluded that

> "variations in velar elevation during speech which also occur during prolonged velopharyngeal closure as observed by researchers using cinefluorography is determined by variations in the degree of velar muscle activity and not by variations in timing and movements of adjacent structures such as hypothesized by Moll and Shriner (1967)".

Ushijima and Hirose (1974) studied EMG activity of the levator palatini in the production of meaningful disyllabic Japanese test words. By correlating their results with Ushijima and Sawashima (1972)'s fiberoptic analysis of the same utterances, they found that: "electromyographically anticipatory nasal coarticulation generally followed the so called "look-ahead" mechanism ..." (ibid., p. 315) and also that "there seems to be no anticipatory lowering of the velum during the vowel segments before a syllable boundary in the /CVV'VN/ environment." (ibid., p. 315) This was not verified by Benguerel et al. (1975a,b) in their study of velar movement during the production of French vowels (both oral and nasalized) in various environments. They obtained their data by recording EMG activity of two velar muscles (levator veli palatini and palatoglossus) and by simultaneously filming velar movement with a fiberscope. They found (ibid., p. 80) that utterances such as /ɑnɑ/ (where an oral vowel precedes a nasal consonant) "show a LP [levator palatini] suppression starting during the first half or third of the initial vowel and ceasing in or shortly after the middle of /n/". Palatoglossus activity in such utterances was found to be very low. At these particular moments the fiberoptic films show a lowering of the velum. By comparing the utterances /tɑt/ and /tɑ̃t/ they found that a lowering of the velum is often accompanied by increased palatoglossus activity in /tɑ̃t/, and that this appeared to correspond to a suppression of levator palatini activity.

These results suggest that a lowering of the velum might be controlled differently for the production of nasalized vowels than for the production of nasal consonants in French.

Bell-Berti and Hirose (1975, p. 69) also combined both EMG and fiberoscopic cinematography and showed that "there is a strong correlation between the size of the increase in EMG potential and the size of change in velar height". Their electrodes were inserted in LP only.

2.43 Results Obtained by Use of Cinefluorography and Aerodynamic Parameters

Lubker and Moll (1965, p. 268) using cinefluorography and measuring simultaneously oral and nasal airflows found that on many occasions, there may exist an "increase in airflow from the nose while the amount of velopharyngeal opening remains constant and velar height decreases". One of their explanations for this phenomenon is that:

> "velopharyngeal orifice area may be increasing even though velopharyngeal distance at the midline does not change; that is, the measures of velopharyngeal distance may not provide an adequate index of velopharyngeal opening". (ibid., p. 268)

Their explanation seems plausible and stresses the fact that direct measurements of velopharyngeal port opening cannot be obtained with a cinefluorographic method.

2.5 Perception of Nasality in the Literature

Since nasality is so closely related to the movement of the velum, dysfunction of that organ will produce noticeable distortions of the speech signal. In particular, cleft palate patients have a strong nasality component in their speech since no velopharyngeal closure can be achieved. Therefore, it is not surprising that much of the research on nasality has been done using cleft palate patients as subjects. Some devices such as the "Voice Systems Nasality Meter" presented by Weatherley-White et al. (1964) have been built to measure nasality, but as Moll (1964, p. 373) has remarked "it appears that listener judgement techniques provide the only direct and logically valid measures of nasal voice quality." In studies on perception of nasality, speech pathologists, particularly those trained in treating voice disorders, are usually used as judges. Only a few of the experiments pertinent to our investigation will be reported here.

Spriestersbach and Powers (1959) used recordings of seven isolated vowels and of connected speech produced by fifty children with cleft palates. Thirty trained speech therapists judged the severity of nasality on a 7-point scale. It was found that

> "severity of nasality in connected speech is related to severity of nasality for each isolated vowel studied... For cleft palate speakers high vowels are in general, more nasal than low vowels. For vowels with fairly comparable tongue height, front vowels are more nasal than back vowels". (ibid., p. 45)

Massengill and Bryson (1967) used a cinefluorographic unit with a television monitor and a synchronized sound track to study the relationship between velopharyngeal closure on the four vowels [i], [æ], [a] and [u] and judgements of nasality for these vowels. Their ten normal subjects attempted to simulate varying degrees of nasality while pronouncing the vowels. Ten students from a speech class were used as judges of the nasality using a three point scale. The author's findings were that "as the opening between the velum and the pharyngeal wall increased, the vowels were rated as being more nasal" (ibid., p. 50). The correlation ranked in the vowel order: [u], [æ], [i] and [a] where [a] showed the least correlation. Shelton et al. (1967) used two high-sensitivity microphones to compare nasal and oral sound pressure levels (SPL) with nasality judgements. One of the microphones was placed directly below the subject's mouth, the other was attached to a probe which was inserted in the subject's naris. The signals picked up by the microphones were amplified and recorded. Sound pressure level measurements of the recorded signals were then made. The authors found that a statistically significant correlation existed between nasal SPL and nasality judgements, but that such a relationship "is not sufficiently great to warrant use of the apparatus in the diagnosis of nasality in individual clients". (ibid., p. 557). During the last decade, Fletcher developed an instrument called TONAR (The Oral-Nasal Acoustic Ratio) which seems to be reliable enough to measure, quantify and display relative nasalization in continuous speech. (see Fletcher, 1970, Fletcher and Bishop, 1970). The latest version called TONAR II (Fletcher 1976, p. 31) "is used to derive a numerical acoustic ratio score expressed in percent "nasalance", which reflects the relative proportion of sound within a specified frequency band [350-650 Hz] emitted from the mouth and nose during speech". In the same study, Fletcher found a high correlation (0.91) between the "nasalance" scores and judgements of nasality made by naive listeners who were presented with recordings of cleft palate speakers reading a phonemically balanced passage containing no nasal sounds. Possible clinical application of TONAR II for self monitoring of velopharyngeal closure is discussed in Fletcher (1972, 1973).

Ali et al. (1971) studied perception of coarticulated nasality. They used CVC and CVVC syllables where the final consonant was either nasal or non-nasal. They spliced out the entire final consonant along with its vowel-consonant transition and presented the stimuli to twenty-two subjects. They found that the subjects could accurately predict whether the segment truncated was nasal or non-nasal and in particular that "consonants which followed the low back /a/ were perceived as nasal with a significantly greater frequency than the vowels /u/, /eI/ and /i/" (ibid., p. 540)

Even though the proposed experiment slightly differs from those outlined above, it is hoped that some of its results will be relatable to them. Although most subjects in this study are not trained to assess a nasal quality to a stimulus, it is expected that their task be relatively easy because of the phonemic oral/nasalized vowel difference in their mother tongue.

Chapter 3

MATERIALS AND METHODS

3.1 Stimuli

The vowel stimuli used for the two perceptual tests were of the following types:

I: 3 (reference) oral vowels: ɑ, ɛ, ɔ

II: 3 (reference) nasalized vowels: ɑ̃, ɛ̃, ɔ̃

III: 15 (test) nasalized vowels: $\tilde{ɑ}_1$, $\tilde{ɑ}_2$, $\tilde{ɑ}_3$, $\tilde{ɑ}_4$, $\tilde{ɑ}_5$, $\tilde{ɛ}_1$, $\tilde{ɛ}_2$, $\tilde{ɛ}_3$, $\tilde{ɛ}_4$, $\tilde{ɛ}_5$, $\tilde{ɔ}_1$, $\tilde{ɔ}_2$, $\tilde{ɔ}_3$, $\tilde{ɔ}_4$, $\tilde{ɔ}_5$

IV: 9 (test) oral vowels: $ɑ_1$, $ɑ_2$, $ɑ_3$, $ɛ_1$, $ɛ_2$, $ɛ_3$, $ɔ_1$, $ɔ_2$, $ɔ_3$

Items of types I, II and III were used to construct Test 1 while items of types I, II and IV were used for Test 2.

The 3 (reference) oral vowels were spliced out of the 3 words /tɑt/, /tɛt/, /tɔt/.

The 3 (reference) nasalized vowels were spliced out of the 3 words /tɑ̃t/, /tɛ̃t/, /tɔ̃t/.

The 15 (test) nasalized vowels were constructed from 5 single periods isolated in each of the 3 words /tɑ̃t/, /tɛ̃t/, /tɔ̃t/ whereas the 9 (test) oral vowels were constructed from 3 single periods isolated in each of the 3 words /ɑnɑ/, /ɛnɛ/, /ɔnɔ/.

3.2 Selection of Stimuli Utterances

The words /tɑ̃t/, /tɛ̃t/, /tɔ̃t/ from which the stimuli of type III were constructed, were taken from the original sound track of the experiments by Benguerel et al. (1975a,b). They had been originally produced in isolation by Benguerel, a native speaker of French from Lausanne, Switzerland. After digitization, these original recordings were observed on a computer scope screen. The reader is referred to Section 3.3 for a more detailed description of the digitization method. The experimenter was able to determine the fundamental frequency for each utterance by counting the number of samples in each cycle of the vowel.

A pilot experiment had shown that fundamental frequency, intensity and duration of the stimuli could influence the listeners' judgement by providing spurious identification cues. It was therefore decided to use stimuli where no variations of these parameters could be used as cues by the listeners. The /ɑnɑ/, /ɛnɛ/, /ɔnɔ/ utterances were found to have a lower fundamental frequency (130 Hz) than the /tɑ̃t/, /tɛ̃t/, /tɔ̃t/ utterances (143 Hz). It was decided to keep the original recordings of /tɑ̃t/, /tɛ̃t/, /tɔ̃t/ to construct vowel stimuli and to re-record a series of /ɑnɑ/, /ɛnɛ/, /ɔnɔ/ utterances. However, since the /tɑ̃t/, /tɛ̃t/, /tɔ̃t/ were too short to construct reference stimuli, the same speaker used in Benguerel et al. (1975a) recorded the 9 utterances /tɑt/, /tɛt/, /tɔt/, /tɑ̃t/, /tɛ̃t/, /tɔ̃t/, /ɑnɑ/, /ɛnɛ/, /ɔnɔ/.

The utterances mentioned above were recorded on an audio tape using a Revox A 77 tape recorder and an Altec 681 ALO microphone. A pure tone of 143 Hz, monitored with a calibrated frequency counter, was presented via headphones to the speaker who attempted to phonate at the frequency he heard. The speaker also prolonged his /tɑt/, /tɛt/, /tɔt/, /tɑ̃t/, /tɛ̃t/, /tɔ̃t/ utterances.

Narrow band spectrograms of all the re-recorded utterances were made to determine the sequences with the steadiest fundamental frequency. Eventually, twelve words with a fundamental frequency between 142 and 146 Hz were selected to construct the test stimuli.

3.3 Vowel Editing

The twelve selected utterances were digitized and stored on a LINC tape using a PDP-12 computer and a set of programs developed by Lloyd Rice at UCLA. The computer can sample the speech signal at a rate of up to 12,000 samples per second. However, at such a high rate, only approximately 1.5 second of speech can be digitized because of the slow transfer of samples to the LINC tape. Once the core buffer has been filled, samples are lost and a noticeable distortion of the speech signal is produced. To overcome this problem, the audio tape containing the selected utterances was played at half speed. The signal was sent to the computer via a Rockland LP filter with a 48 dB/octave attenuation slope and the cut-off frequency set at 6,000 Hz. The result of this operation was equivalent to the digitization of the utterance at a sample rate of 12,000 samples per second, when played at regular speed (see Fig. 1 a). The volume control of the tape recorder

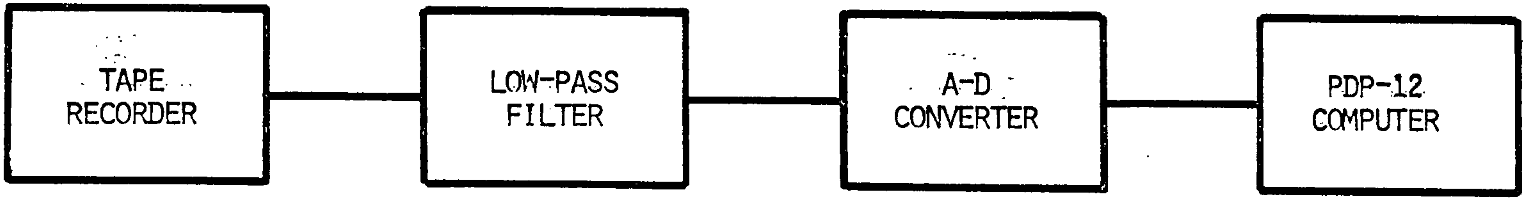

Fig. 1a. Block diagram of apparatus for digitization of speech signal

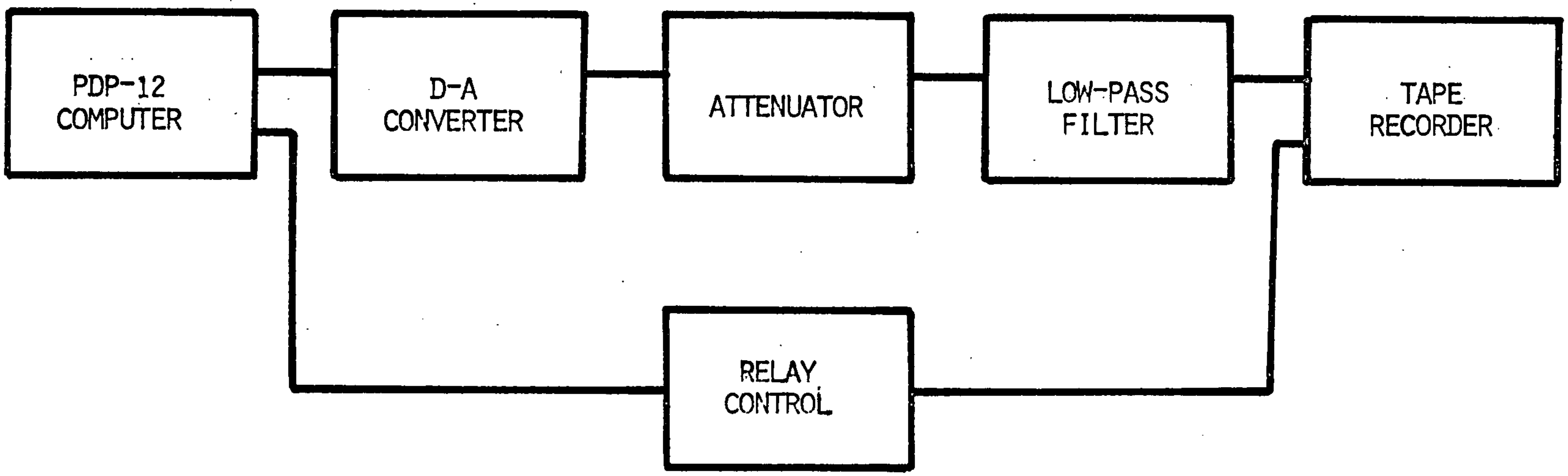

Fig. 1b. Block diagram of apparatus for production of test items

was adjusted for each utterance so that all the stimuli words were digitized at approximately the same peak intensity. The speech signal stored on the tape was then displayed on the oscilloscope screen. A knob enabled the operator to move the signal back and forth on the screen. A program made it possible to isolate parts of the digitized signal and store them on tape.

Single cycles, containing as much as possible the same number of samples, and taken at five different places in the /tɑ̃t/, /tɛ̃t/, /tɔ̃t/ utterances taken from the data of Benguerel et al. (1975a,b) and at three different places in the re-recorded /ɑnɑ/, /ɛnɛ/, /ɔnɔ/ utterances, were isolated and stored on the LINC tape. Intervals between the selected cycles were approximately the same. Edited cycles were always taken between two zero-crossing of the signal wave. Each of the single isolated cycles was then reproduced 40 times and stored on the LINC tape. The signals so produced were long enough to be identified by the listeners. Portions of vowels taken from the re-recorded /tɑt/, /tɛt/, /tɔt/, /tɑ̃t/, /tɛ̃t/, /tɔ̃t/ utterances, having a length of 40 cycles were isolated and stored on the computer tape. The envelopes of the signals so produced had a rectangular shape. They were given a trapezoidal shape to produce a smooth onset and offset of 25 ms each.

The digitized signals were reconverted into analog signals and recorded on an audio tape via a low pass filter which was set at 6,000 Hz to eliminate high frequency digital noise generated by the computer (See Fig. 1b). Mingograms, including a speech power display, were made of the vowel stimuli obtained and differences in vowel intensity were found in comparison with calibrated 1,000 Hz pure tones recorded at several reference levels. It was then decided to correct the intensity of each vowel so that they all be within ±.5 dB of each other.

3.4 Test Preparations

Two tests (Test 1, Test 2) were constructed. Both tests contained the six reference vowel stimuli ɑ,ɛ,ɔ,ɑ̃,ɛ̃,ɔ̃. This was done to evaluate the listeners' subjective judgement of the degree of nasality attributed to unmodified vowels and to help establish a base line for the evaluation of the listeners' responses to the vowels reconstructed from single cycles.

The items of Test 1 consisted of five presentations of each of the six reference vowel stimuli and of the fifteen reconstructed vowels from the original /tɑ̃t/, /tɛ̃t/, /tɔ̃t/ utterances. The test stimuli were recorded in random order on one track of a stereo tape. No two stimuli originating from the same utterance were juxtaposed. Ten "buffer" vowels, selected from the 105 test stimuli were recorded, five at the beginning, and five at the end of the tape.

For Test 2, the test items consisted of five presentations of each of the six reference vowel stimuli and of the nine reconstructed vowels from the re-recorded /ɑnɑ/, /ɛnɛ/, /ɔnɔ/ utterances. A total of 85 vowel stimuli were recorded.

In both tests a silence of five seconds was inserted between the end of one stimulus and the beginning of the next. In the five seconds preceding each stimulus, French numbers from 1 to 115 and from 1 to 85, for Test 1 and Test 2 respectively, were recorded at low intensity level on the other track of the stereo tape. Only items #6 to 110 (Test 1) and #6 to 80 (Test 2) were considered in the analysis.

3.5 Subjects

Twenty-seven native speakers of French were used as subjects. Their ages varied between 20 and 45. They had been in North America for periods ranging from a few days to eighteen years. Most regions of France, including St. Pierre and Miquelon (three subjects) were represented. Two subjects were born in Switzerland, another subject was born in Southern Belgium and one was born in Val d'Aosta in Northern Italy, and was bilingual. One of the French-born subjects (aged 22) had been living in Canada for most of his life but the language spoken at home was French. All but one of the subjects had a working knowledge of English; eight subjects had no knowledge of phonetics, while ten had had phonetic training and fifteen were teachers of the French language.

The subjects reported no history of deafness or hearing loss. Two subjects who were concerned about a possible loss were given a thorough audiological examination. Their hearing was found to be well within limits of normal range.

3.6 Test Procedure and Equipment

The subject was seated in a quiet room with an ambiant noise of about 35 dBA as measured on a B & K model 2203 sound level meter. The test tape was played on a Revox model A 77 tape recorder through a set of GSC headphones model TDH 39-300Z at a level corresponding to 70 - 80 dB SPL as measured through the same headphones with a B & K model 2203 sound level meter, a B & K model 4152 artificial ear and a 6 cm^3 coupler.

Subjects were instructed, in French, to assess the degree of nasality of each vowel stimulus on a 5 point scale (See Appendix for the actual wording of the written instructions). In addition it was pointed out to the subjects that the tests did not contain any specific number of oral or nasalized vowels. Therefore, they were encouraged to note down what they actually heard, and not to assume any particular distribution. For example; if they thought they heard every stimulus as nasal they were instructed to check the column "nasale" all the time.

The first five items were played to them. The tape was stopped and the subjects were asked if they had any questions. The tape was then reset at the first item and formal testing began. Fifteen subjects were presented with Test 1 first and the remaining twelve heard Test 2 first.

Chapter 4

RESULTS

4.1 Data Analysis

4.11 Storage and Sorting of the Data

In order to analyze the results of the perception tests, listeners' responses were assigned a number from 1 to 5 corresponding to the number of the column in which they had registered their answer; for instance, an answer "orale" was assigned a 1, an answer "nasale" a 5. Listeners' responses were then stored on a LINC tape so that they would be available for computation. Responses to items #6 through 110 in Test 1 and to items #6 through 80 in Test 2 were then sorted in matrix form, for each listener separately. An example of such a display is shown in Table I. It was felt that hearing only five "buffer" stimuli might not be sufficient for most listeners and that a better evaluation of their performances might be obtained by considering only their responses to the last four presentations of each stimulus. It was decided to compare their responses to all five presentations of each stimulus, with their responses to the last four presentations of each stimulus. The mean and standard deviation of each listener's responses to the five presentations (m_5, s_5) and to the last four presentations (m_4, s_4) of each stimulus were computed and listed in Table I. Also in each case and for each test, the average standard deviations ($\bar{s}_5$, $\bar{s}_4$) were computed. Table I is an example of the sorted data for subject

Suject#10

Test 1.

Stimuli	Responses	m_5	s_5	m_4	s_4
ɑ	5 1 1 1 1	1.80	1.60	1.00	0.00
ε	1 1 1 1 1	1.00	0.00	1.00	0.00
ɔ	1 1 1 1 1	1.00	0.00	1.00	0.00
ɑ̃	5 2 4 4 5	4.00	1.10	3.75	1.09
ε̃	1 4 4 4 4	3.40	1.20	4.00	0.00
ɔ̃	5 5 5 5 5	5.00	0.00	5.00	0.00
ɑ̃$_{1}$	4 2 1 1 1	1.80	1.17	1.25	0.43
ɑ̃$_{2}$	5 1 1 1 1	1.80	1.60	1.00	0.00
ɑ̃$_{3}$	4 2 4 4 5	3.80	0.98	3.75	1.09
ɑ̃$_{4}$	5 5 5 4 5	4.80	0.40	4.75	0.43
ɑ̃$_{5}$	3 5 5 5 4	4.40	0.80	4.75	0.43
ε̃$_{1}$	5 1 1 1 1	1.80	1.60	1.00	0.00
ε̃$_{2}$	1 5 5 5 5	4.20	1.60	5.00	0.00
ε̃$_{3}$	5 5 5 5 5	5.00	0.00	5.00	0.00
ε̃$_{4}$	5 5 5 5 5	5.00	0.00	5.00	0.00
ε̃$_{5}$	1 1 1 1 1	1.00	0.00	1.00	0.00
ɔ̃$_{1}$	1 1 1 1 1	1.00	0.00	1.00	0.00
ɔ̃$_{2}$	4 4 1 1 1	2.20	1.47	1.75	1.30
ɔ̃$_{3}$	4 2 1 1 2	2.00	1.10	1.50	0.50
ɔ̃$_{4}$	4 5 5 4 2	4.00	1.10	4.00	1.22
ɔ̃$_{5}$	5 5 1 1 1	2.60	1.96	2.00	1.73
			$\bar{s}_5$=0.84		$\bar{s}_4$=0.39

Test 2.

Stimuli	Responses	m_5	s_5	m_4	s_4
ɑ	2 2 2 1 4	2.20	0.98	2.25	1.09
ε	1 1 1 1 1	1.00	0.00	1.00	0.00
ɔ	1 1 1 1 1	1.00	0.00	1.00	0.00
ɑ̃	4 5 5 5 4	4.60	0.49	4.75	0.43
ε̃	5 5 5 4 5	4.80	0.40	4.75	0.43
ɔ̃	5 5 5 5 5	5.00	0.00	5.00	0.00
ɑ$_{1}$	4 4 4 1 2	3.00	1.26	2.75	1.30
ɑ$_{2}$	1 1 1 1 1	1.00	0.00	1.00	1.00
ɑ$_{3}$	2 4 4 1 5	3.20	1.47	3.50	1.50
ε$_{1}$	1 1 1 1 1	1.00	0.00	1.00	0.00
ε$_{2}$	1 1 1 1 1	1.00	0.00	1.00	0.00
ε$_{3}$	1 1 1 1 1	1.00	0.00	1.00	0.00
ɔ$_{1}$	1 1 1 1 1	1.00	0.00	1.00	0.00
ɔ$_{2}$	1 1 1 1 1	1.00	0.00	1.00	0.00
ɔ$_{3}$	1 1 1 1 1	1.00	0.00	1.00	0.00
			$\bar{s}_5$=0.31		$\bar{s}_4$=0.32

Table I. Sorted data for Subject #10

#10; it represents her responses to Test 1 and Test 2 sorted as described above.

In Table I the left column represents the test stimuli. Reference vowels in the order: ɑ, ɛ, ɔ, ɑ̃, ɛ̃, ɔ̃ were listed first, followed by the corresponding constructed stimuli. The next five columns list for each stimulus the subject's responses; the last four columns list m_5, s_5, m_4, s_4, for each stimulus. From now on, the test stimuli will be referred to in the forms indicated in Table I. The scores to the last four responses and to all five responses will be referred to as m_4's and m_5's respectively.

Table I suggests that subject #10 might have experienced some difficulty in assessing a degree of nasality to some of the stimuli the first time they were presented. It can also be noted that $\bar{s}_4$ is smaller than $\bar{s}_5$ in Test 1. The same observation was made for most of the subjects and in both tests. This suggests that the subjects' responses to the last four presentations of each stimulus are more representative of their judgement after adaptation to the test items than are their responses to all five presentations.

4.12. Normalization

Some listeners, consistently, did not use the extremes of the five-point scale while responding to presentations of the reference vowels. Therefore, their five/four response scores were distant from the extreme scores (1 or 5) for both oral and nasalized vowels. All the data was thus normalized by mapping the scores for the reference vowels onto 1 and 5 (oral and nasalized) and for the corresponding constructed stimuli by using the following transformation formula:

$$n = 4 \frac{m - a}{b - a} + 1$$

where

a = subject's score for each oral reference vowel

b = subject's score for its nazalized counterpart

m = subject's non-normalized score for one of its constructed counterparts.

n = subject's normalized score for this same constructed vowel

It can be seen that this formula maps the reference vowel scores onto 1 or 5, as desired. Hence, the standard deviations of the normalized scores for the reference vowels are zero.

Table II shows the normalized scores and their standard deviations for subject #10's responses to all five and to the last four presentations of each stimulus. Original scores did not always map onto a number from 1 to 5. In cases where $m < a$, then $n < 1$; where $m > b$, then $n > 5$; where $a = b$, then $n \to \infty$, and where $b < a$ then n in many cases becomes negative. The standard deviations (SD) for the normalized scores of the constructed stimuli are usually larger than the SD's for the non-normalized scores. This is due to the fact that normalized scores were in general larger than the non-normalized ones (because in general $b - a < 4$).

Subject#10

Test 1.

Stimuli	Normalized m_5	SD	Normalized m_4	SD
a	1.00	0.00	1.00	0.00
ε	1.00	0.00	1.00	0.00
ɔ	1.00	0.00	1.00	0.00
ã	5.00	0.00	5.00	0.00
ɛ̃	5.00	0.00	5.00	0.00
ɔ̃	5.00	0.00	5.00	0.00
$\tilde{a}_1$	1.00	2.12	1.36	0.63
$\tilde{a}_2$	1.00	2.91	1.00	0.00
$\tilde{a}_3$	4.64	1.78	5.00	1.59
$\tilde{a}_4$	6.45	0.73	6.45	0.63
$\tilde{a}_5$	5.73	1.45	6.45	0.63
$\tilde{\varepsilon}_1$	2.33	2.67	1.00	0.00
$\tilde{\varepsilon}_2$	6.33	2.67	6.33	0.00
$\tilde{\varepsilon}_3$	7.67	0.00	6.33	0.00
$\tilde{\varepsilon}_4$	7.67	0.00	6.33	0.00
$\tilde{\varepsilon}_5$	1.00	0.00	1.00	0.00
$\tilde{ɔ}_1$	1.00	0.00	1.00	0.00
$\tilde{ɔ}_2$	2.20	1.47	1.75	1.30
$\tilde{ɔ}_3$	2.00	1.10	1.50	0.50
$\tilde{ɔ}_4$	4.00	1.10	4.00	1.22
$\tilde{ɔ}_5$	2.60	1.96	2.00	1.73

Test 2.

Stimuli	Normalized m_5	SD	Normalized m_4	SD
a	1.00	0.00	1.00	0.00
ε	1.00	0.00	1.00	0.00
ɔ	1.00	0.00	1.00	0.00
ã	5.00	0.00	5.00	0.00
ɛ̃	5.00	0.00	5.00	0.00
ɔ̃	5.00	0.00	5.00	0.00
a_1	2.33	2.11	1.80	2.08
a_2	-1.00	0.00	-1.00	0.00
a_3	2.67	2.45	3.00	2.40
ε_1	1.00	0.00	1.00	0.00
ε_2	1.00	0.00	1.00	0.00
ε_3	1.00	0.00	1.00	0.00
$ɔ_1$	1.00	0.00	1.00	0.00
$ɔ_2$	1.00	0.00	1.00	0.00
$ɔ_3$	1.00	0.00	1.00	0.00

Table II. Normalized scores for Subject #10's responses to all five and to the last four presentations of each stimulus type

4.13 Criteria for Selecting the Best Subjects

Each subject's responses were tabulated as in Tables I and II, and examined individually. Some of the listeners appeared to have performed more consistently than others. It was then decided to select the best subjects (according to some consistency criterion) and to compare their results with those of the entire population. Using the responses to the last four presentations, two criteria, based on ability to differentiate between oral and nasalized vowels were established:

Criterion 1: If a subject made the oral/nasalized distinction for the six pairs of reference vowels in both tests by having $m_4 < 3$ for all oral reference vowels and on $m_4 > 3$ for all nasalized reference vowels, then he was deemed to have passed Criterion 1. This subject's performance for the other stimuli could be expected to be adequate since he/she could use the reference vowels as guides in his/her judgement.

Criterion 2: For each reference vowel and for each test, the distribution of the 27 subjects' m_4 's were established. For each pair of reference vowels the difference between the lower quartile value for the nasalized vowel and the upper quartile for the oral vowel was determined and listed in Table III. These differences ranged from 1.25 for ε/ε̃ in Test 2 to 2.75 for ɔ/ɔ̃ in Test 2. Similarly, calculated differences were determined for score intervals other than the quartile interval but no change in rank order (between vowel pairs) was observed. (Possible implications of these values will be discussed in section 4.23). For each subject and for each pair of reference vowels the differences between the m_4 of each nasalized vowel and the m_4 of its oral counterpart were calculated. If the six values so obtained for each subject were larger or equal to the corresponding values listed in Table III, the subject was deemed to have passed Criterion 2. Subjects passing this criterion were felt to have made a definite discrimination between oral and nasalized reference vowels.

Two other possible criteria based on the subject's consistency in responding were considered: 1) selecting a subject if he/she had responded at least 3 times on the same side of 3 for the different stimuli; 2) selecting a subject if his/her average standard deviations $\bar{s}_4$'s in both tests were smaller than the median $\bar{s}_4$. These two criteria, however, were felt to be

Reference vowel pairs	m_4 quartile differences
Test 1.	
ɑ/ɑ̃	1.75
ɛ/ɛ̃	1.50
ɔ/ɔ̃	2.00
Test 2.	
ɑ/ɑ̃	1.75
ɛ/ɛ̃	1.25
ɔ/ɔ̃	2.75

Table III. m_4 quartile differences for criterion 2 too dependent on the subject's strategy and on the nature of the constructed stimuli. They were thought to be less efficient than the others and were thus not used. Nine subjects were then grouped together and their scores as a group were compared with the scores of the entire population as a group. Both the five-response scores and the last-four-response scores were considered separately. In each case, the non-normalized scores were averaged separately for the 9-subject group and for the 27-subject group and the standard deviations were found. The normalized scores of the 9-subject group were also averaged and their standard deviations computed. Since normalization of certain subjects' scores resulted in some very large, very small or negative numbers as mentioned in section 4.12, the averages and standard deviations of the normalized scores for the 27-subject group were not computed. Not surprisingly, the standard deviations of the 27-subject group were found to be larger than those of the 9-subject group. The results of the above computations are shown in Fig. 2 to Fig. 5.

Each figure summarizes the results of one test. On each graph the triangles and dash-dot line represents the average scores for the 27-subject group. The circles and dash line on the one hand and the squares and solid line on the other hand represent the average non-normalized and average normalized scores, respectively, for the 9-subject group.

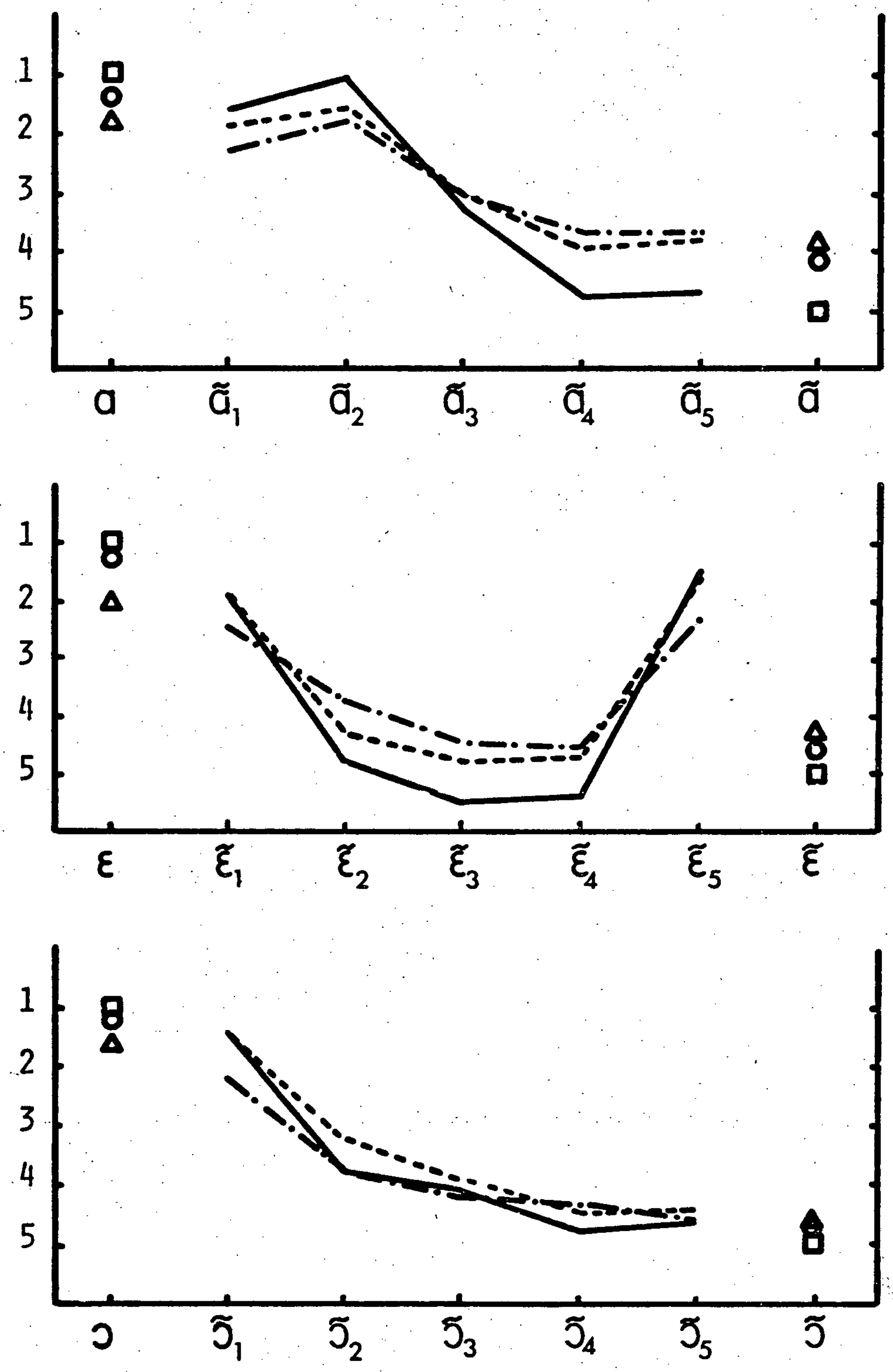

Fig. 2. Average scores (five responses) in Test 1 for:

△ —·—· 27-subject group, non-normalized

○ - - - - 9-subject group, non-normalized

□ ——— 9-subject group, normalized

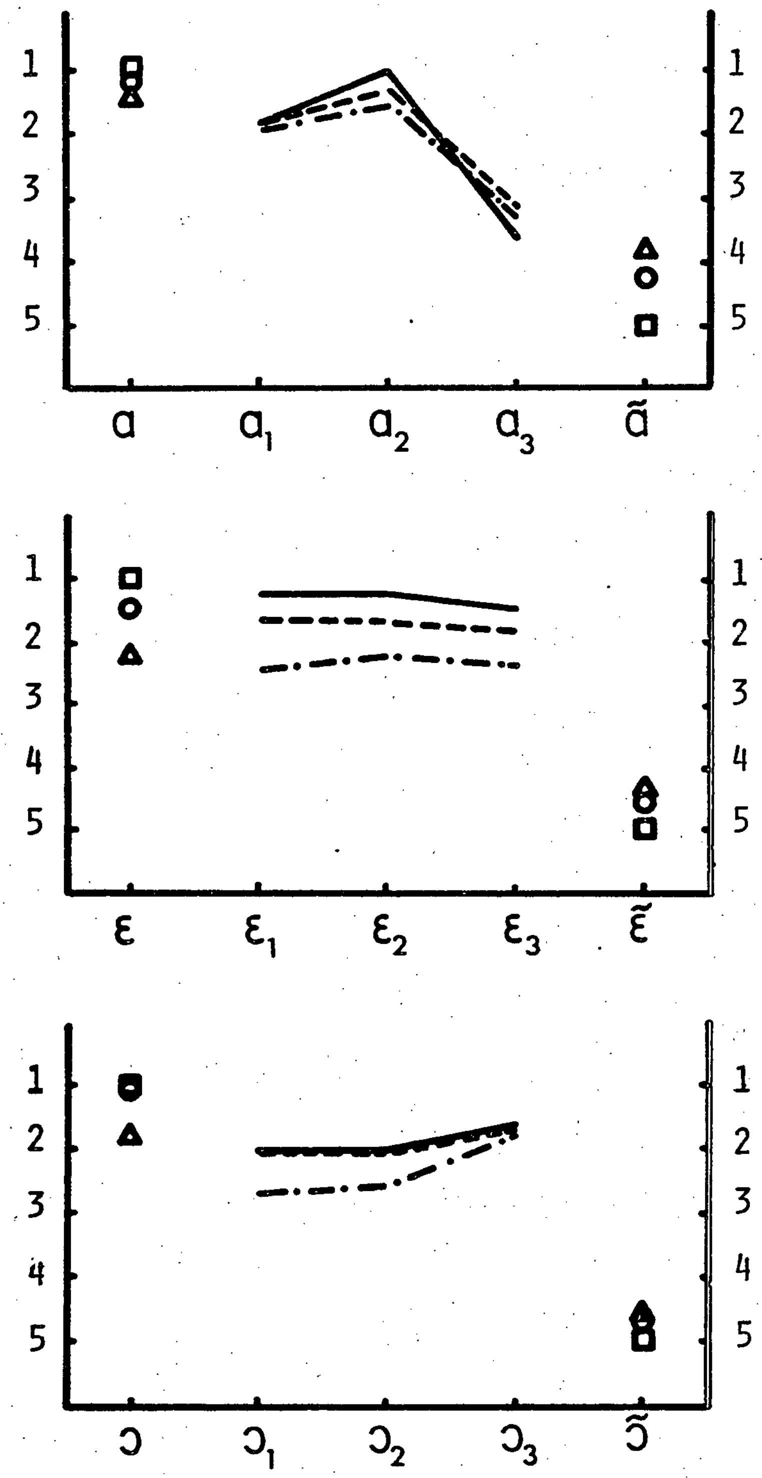

Fig. 3. Average scores (five responses) in Test 2 for:

△ —·—·- 27-subject group, non-normalized

○ – – – – 9-subject group, non-normalized

□ ———— 9-subject group, normalized

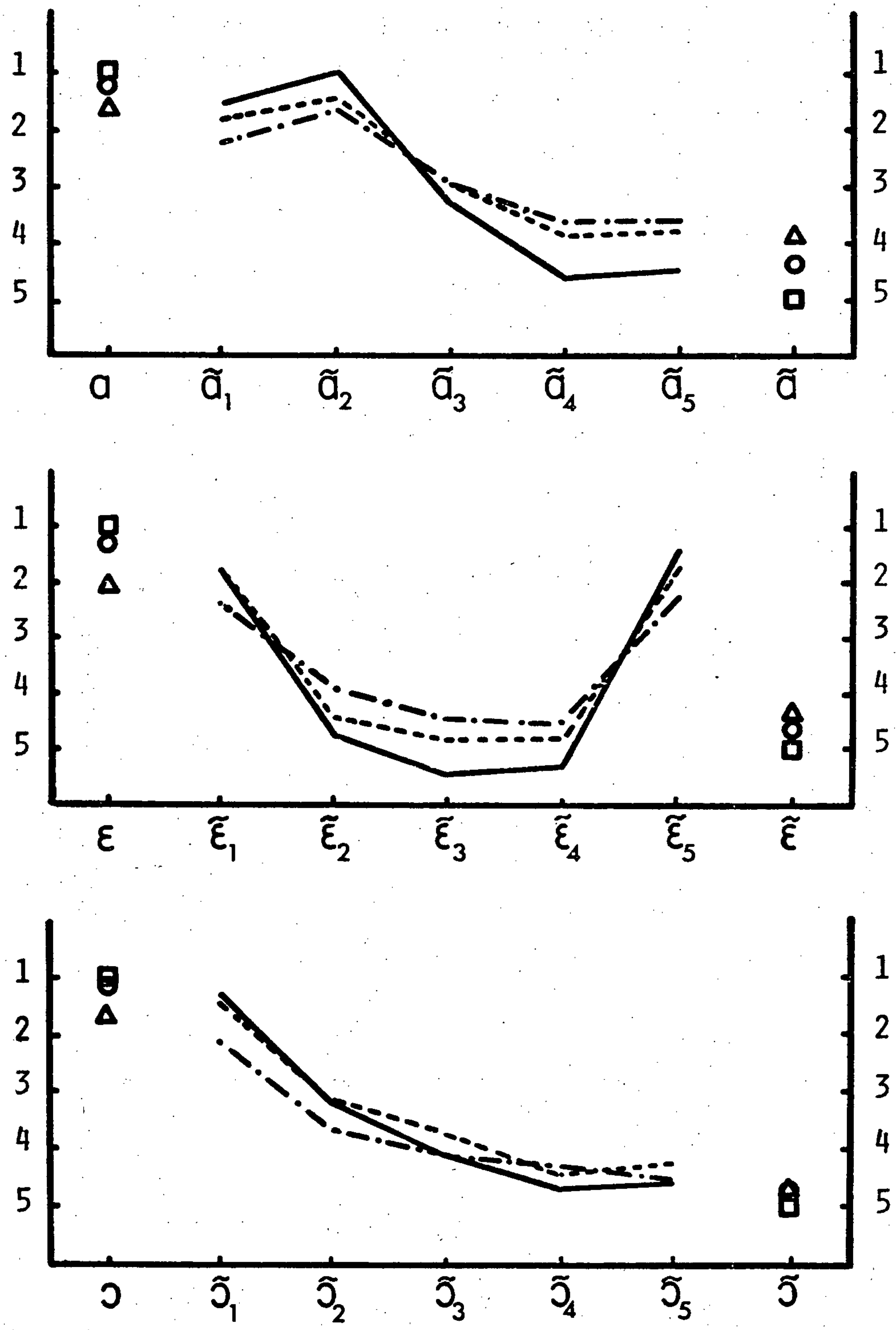

Fig. 4. Average scores (last four responses) in Test 1 for:

△ —·—· 27-subject group, non-normalized

○ - - - - 9-subject group, non-normalized

□ ——— 9-subject group, normalized

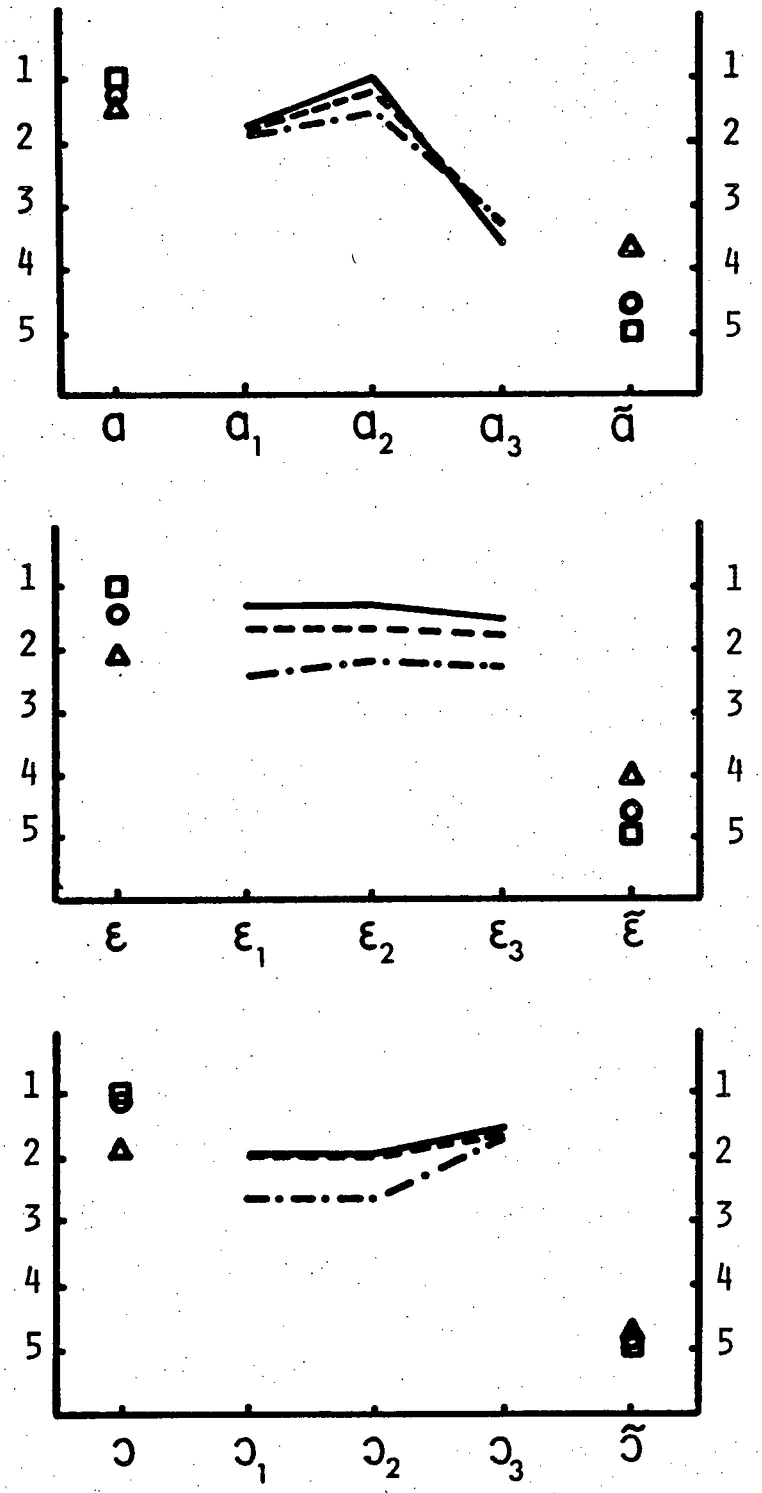

Fig. 5. Average scores (last four responses) in Test 2 for:

△ —·—· 27-subject group, non-normalized

○ – – – – 9-subject group, non-normalized

□ ——— 9-subject group, normalized

4.2 Analysis of Graphs

4.21 General Observations

Whether, in the data presented in Fig. 2 to Fig. 5, one considers the scores averaged on five or four responses, for the 9-subject group or for the 27-subject group, non-normalized or normalized (for the 9-subject group), the curves obtained have basically the same shape. The main difference is that the standard deviations, although not plotted, are always greater for five responses than for four; they are also greater for the 27-subject group than for the 9-subject group. Average scores for the 9-subject group reach more extreme values than do those for the 27-subject group and as one should expect, the average scores for the normalized data also reach more extreme values than for the non-normalized ones, sometimes even greater than 5, e.g. for $\tilde{\varepsilon}_3$ and $\tilde{\varepsilon}_4$ on Fig. 2 and Fig. 4.

4.22 Analysis of Individual Vowel Types

Since, as shown above, the average score curves were basically the same for the 27-subject group as for the 9-subject group, it was decided to use the curves representing the average scores of the whole population in analyzing the subjects' responses to individual vowel types. This was done in an attempt to obtain more general results and thus, give more weight to the observations. The same comments will hold for any curve, mutatis mutandis. If an average score is less than 3, it will be postulated that the listeners judged the vowel as "oral";

if an average score is more than 3, it will be postulated that the listeners judged the vowel as "nasalized", and if an average score is close to 3, it will be said that the listeners assessed an "in between" quality to the vowel. Considering the subjects' average scores for the stimuli constructed from /tãt/ as displayed on Fig. 2 and Fig. 4, one may notice that the average scores for $\tilde{a}_1$ and $\tilde{a}_2$ are less than 3 and that the average scores for $\tilde{a}_4$ and $\tilde{a}_5$ are more than 3. The average scores for $\tilde{a}_3$ are close to 3. The boundary between "oral" and "nasalized" appears to be around $\tilde{a}_3$. It can also be noticed that average scores for $\tilde{a}_2$ are usually less than those for $\tilde{a}_1$; this would suggest that subjects had a tendency to judge $\tilde{a}_2$ as more "oral" than $\tilde{a}_1$. These items were thoroughly rechecked to rule out a possible inversion of the two items in the preparation of the listening test tape. No explanation can be found to explain this observation.

For stimuli constructed from /tɛ̃t/, one notices (Fig. 2 and Fig. 4) steadily increasing average scores from $\tilde{\varepsilon}_1$ to $\tilde{\varepsilon}_4$, the boundary between "oral" and "nasalized" occurs between $\tilde{\varepsilon}_1$ and $\tilde{\varepsilon}_2$. Average scores for $\tilde{\varepsilon}_5$ are much smaller than those for $\tilde{\varepsilon}_4$ and would suggest that $\tilde{\varepsilon}_5$ had been judged as "oral" by the listeners. A discussion of this point will be presented in Section 4.3.

For stimuli constructed from /tɔ̃t/, the graphs on Fig. 2 and Fig. 4 show an increase in average scores from $\tilde{ɔ}_1$ to $\tilde{ɔ}_5$. The boundary between "oral" and "nasalized" occurs between $\tilde{ɔ}_1$ and $\tilde{ɔ}_2$.

Considering the subjects' average scores to stimuli constructed from /ɑnɑ/ displayed in Fig. 3 and Fig. 5, one notices that the average score for $ɑ_1$ and $ɑ_2$ are smaller than those for $ɑ_3$. The average scores for $ɑ_1$ and $ɑ_2$ are less than 3 and the average scores for $ɑ_3$ are slightly more than 3.

The graphs on Fig. 3 and Fig. 5 also show that the average scores for stimuli constructed from /ɛnɛ/ and /ɔnɔ/ are all less than 3, and that the average scores for $ɔ_3$ are slightly less than those for $ɔ_1$ and $ɔ_2$. So, the listeners judged stimuli constructed from /ɑnɑ/, /ɛnɛ/ and /ɔnɔ/ as being mostly "oral", with the exception of $ɑ_3$ which was judged "in between". None of these stimuli were perceived as "nasalized" as was the case for stimuli constructed from /tɑ̃t/, /tɛ̃t/ and /tɔ̃t/, but average scores indicate that some stimuli were judged more "oral" than others.

4.23 General Trends

Table IV shows the stimuli immediately preceding each of the last four presentations of the reference vowels in both tests and their average scores with respect to 3 in parentheses for the 27-subject group.

It will be noted that those stimuli preceding presentations of ɛ and ɑ̃ have average scores more than or close to 3 in Test 1 and less than or close to 3 in Test 2. In order to investigate a possible influence of the preceding stimulus on the listeners' judgement for these particular reference vowels, Test 1.

Reference vowel	Preceding stimulus			
ɑ	ɔ̃$_2$ (≈3)	ɛ̃$_5$ (<3)	ɔ̃$_4$ (>3)	ɛ̃$_1$ (<3)
ɛ	ɑ̃$_3$ (≈3)	ɔ̃$_5$ (>3)	ɑ̃$_3$ (≈3)	ɔ̃$_5$ (>3)
ɔ	ɛ̃$_2$ (>3)	ɛ̃$_5$ (<3)	ɛ̃$_2$ (>3)	ɛ̃$_1$ (<3)
ɑ̃	ɔ̃$_3$ (>3)	ɛ̃$_2$ (>3)	ɛ̃$_4$ (>3)	ɔ̃$_5$ (>3)
ɛ̃	ɔ̃$_1$ (<3)	ɑ̃$_1$ (<3)	ɔ̃$_3$ (>3)	ɔ̃$_4$ (>3)
ɔ̃	ɛ̃$_1$ (<3)	ɛ̃$_3$ (>3)	ɛ̃$_1$ (<3)	ɛ̃$_2$ (>3)

Test 2.

Reference vowel	Preceding stimulus			
ɑ	ɛ$_2$ (<3)	ɛ$_1$ (<3)	ɛ (<3)	ɔ (<3)
ɛ	ɑ$_2$ (<3)	ɑ$_3$ (≈3)	ɑ$_2$ (<3)	ɑ$_1$ (<3)
ɔ	ɛ$_3$ (<3)	ɛ$_1$ (<3)	ɑ̃ (>3)	ɑ$_3$ (≈3)
ɑ̃	ɛ$_3$ (<3)	ɔ$_2$ (<3)	ɔ$_3$ (<3)	ɔ$_3$ (<3)
ɛ̃	ɔ$_3$ (<3)	ɔ$_1$ (<3)	ɑ$_3$ (≈3)	ɔ$_2$ (<3)
ɔ̃	ɑ$_2$ (<3)	ɛ$_3$ (<3)	ɑ (<3)	ɛ (<3)

Table IV. Stimuli preceding each of the last four presentations of the reference vowels in both tests and their average scores (27-subject group) re.:3

t-tests were run to compare the scores of these vowels in Test 1 and Test 2. The t-tests show no statistically significant differences between the scores. Thus it appears that the preceding stimulus had no significant influence on the listeners' responses to the last four presentations of ε and ã in either test. It may be concluded therefore, that this assumption does also apply to any other test item. Table V lists the subjects' standard deviations (s_4) for the reference vowels in both tests (s_4's of the 9-subject group were listed on the left of the table). Considering each group of subjects separately, a series of t-tests was run to compare the distribution of the s_4's of each reference vowel with the distribution of the s_4's of all six reference vowels, for each test separately. The only significant difference was in the 27-subject group in Test 2 where the s_4's for ɔ̃ were found to be statistically significantly different from the six reference vowels as a whole ($p < .01$). Table V shows that, for this group, s_4's for ɔ̃ in Test 2 are generally smaller than the s_4's for the other reference vowels and that there are more s_4 's equal to zero for ɔ̃ . This suggests that ɔ̃ was more easily identified than the other reference vowels in Test 2.

As noted in section 4.13, the pair ε/ε̃ in Test 2 had an m_4 quartile difference of 1.25 whereas the pair ɔ/ɔ̃ had a difference of 2.75. This shows that the listeners chose more extreme values in responding to ε and ε̃ . This may be attributed to three possible factors: 1) ε̃ might be perceptually less nasalized than ɔ̃, i.e. a listener might, in

Test 1.

	27-subject group																										
	9-subject group																										
ɑ̆	0.00	0.83	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50	0.43	0.00	0.43	0.00	0.00	0.83	0.00	0.83	0.00	0.00	0.83	0.00	0.00	0.43	0.71	1.58	0.83
ɛ	0.00	0.87	0.00	0.00	0.00	0.50	0.00	0.43	0.00	0.50	0.43	0.43	1.09	0.43	1.30	0.50	0.50	0.00	0.43	0.00	1.50	0.87	0.00	0.43	1.22	1.50	0.71
ɔ	0.43	0.00	0.00	0.00	0.00	0.00	0.00	0.50	0.00	0.43	0.83	0.71	0.71	0.00	0.00	0.00	0.00	0.71	0.00	0.00	0.00	0.87	0.43	0.50	0.50	0.00	0.43
ɑ̃	0.50	0.50	1.09	0.00	0.00	0.83	1.73	0.00	0.43	0.43	0.00	0.00	0.50	0.50	0.43	0.50	0.00	0.43	0.43	0.43	0.50	0.43	0.00	0.71	0.83	1.12	0.43
ɛ̃	0.00	0.83	0.00	0.43	0.00	0.43	0.00	0.43	0.00	0.50	0.00	0.00	0,43	0.00	0.00	0.00	0.50	0.00	0.00	0.87	0.00	1.30	0.50	0.50	0.83	0.83	0.50
ɔ̃	0.00	1.09	0.00	0.00	0.00	0.50	0.00	0.43	0.00	0.00	0.00	0.43	0.00	0.00	0.00	0.00	0.00	0.43	0.00	0.00	0.43	0.00	0.00	0.83	1.30	1.12	0.00

Test 2.

ɑ	0.00	0.50	1.09	0.00	0.00	0.00	0.00	0.50	0.00	0.00	0.87	0.43	0.00	0.43	0.00	0.50	0.00	0.00	0.43	1.30	0.50	0.00	0.43	0.43	0.43	0.83	1.09
ɛ	0.00	0.00	0.00	0.00	0.00	0.43	0.00	0.87	1.22	0.00	0.50	0.43	1.12	0.00	0.00	0.43	0.43	0.43	0.43	1.30	0.00	1.00	0.00	0.83	0.87	0.50	0.50
ɔ	0.00	0.00	0.00	0.00	0.00	0.43	0.00	0.43	0.00	0.43	0.50	0.83	0.83	0.43	0.00	0.00	0.43	0.83	0.43	0.00	0.43	1.00	0.43	0.43	0.71	0.50	0.50
ɑ̃	0.50	0.83	0.43	0.00	0.00	0.43	0.71	0.00	0.00	0.00	0.87	0.00	1.22	0.00	0.43	1.30	0.43	0.43	0.43	1.73	0.00	0.00	0.00	1.09	0.87	0.83	0.00
ɛ̃	0.00	0.50	0.43	0.00	0.00	0.43	0.00	0.50	0.00	0.00	0.00	0.50	0.83	0.00	0.00	0.50	0.00	0.43	0.43	0.00	1.22	1.22	0.50	0.43	0.83	0.83	0.50
ɔ̃	0.00	0.43	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.43	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.71	1.09	0.00

Table V. Subjects' s_4 for the reference vowels in both tests general, attribute less nasalization to an [ɛ̃] than to an [ɔ̃]. It was noticed, for instance, in the computation of the m_4 quartile differences for each pair of reference vowels (listed in Table III) that, in Test 2, the upper quartile scores for ɛ and ɔ were apart by 0.25 only but that the lower quartile score for ɛ̃ was 1.25 less than that for ɔ̃. This would indicate that the listeners made some difference between the ɛ̃ and ɔ̃, by assessing a lower degree of nasalization to ɛ̃. Thus, one should expect the m_4 quartile difference for the ɛ/ɛ̃ pair to be smaller than that for the ɔ/ɔ̃ pair. 2) The fact that ɔ̃ was more easily identified as nasalized than the other nasalized reference vowels may be due to the particular speaker used in this experiment. After having completed the task, many listeners reported that they "recognized" the ɔ̃ more easily than other stimuli. 3) On the other hand, it could be that in general ɔ̃ is more easily identified as nasalized than ɛ̃ or ɑ̃, regardless of the speaker. One way to verify this point would be to repeat the experiment with several speakers and to observe whether there is any vowel-speaker interaction as far as perceived nasalization is concerned.

Another series of t-tests compared the subjects' s_4's for each constructed vowel with the s_4's for the reference vowels. No significant difference was found. One might have expected to find a difference which reflected the fact that the constructed stimuli were harder to judge. Results of the t-tests, however, suggest that the constructed stimuli were not more difficult to judge than the reference vowels.

No pattern could be found concerning a possible test order effect. However, of the 9 subjects who passed criteria 1 and 2, 7 had been presented with Test 1 first.

4.3 Correlation of Results with the data of Benguerel et al. (1975a)

The data of Benguerel et al. (1975a) being available, comparisons were made between results of their study and those of the present one in an attempt to correlate them. Benguerel et al. (1975a) had plotted the time course of velar height for the different utterances used in their study. Their Fig. 3 (ibid., p. 71) is an example of such a curve for two different tokens of /tɔt/ and for one of /tɔ̃t/. This figure shows in particular that there exists a strong token-to-token repeatability. Because of this, Benguerel et al. (1975a) had plotted the time course of velar height for one only of each token type of their data. It so happens that a plotting of velar height in time had been done by Benguerel et al. (1975a) for the token of /tɛ̃t/ which was used in this study (see Fig. 6). This unpublished particular plot was thus used to correlate the results of the present study with theirs. The other unpublished plots of Benguerel et al. (1975a) were also used because of the high token-to-token repeatability.

Each one of the constructed stimuli was linearily projected onto the plots of Benguerel et al. (1975a). One was then able to determine the presence or absence of an opening of the velopharyngeal port for each one of the constructed

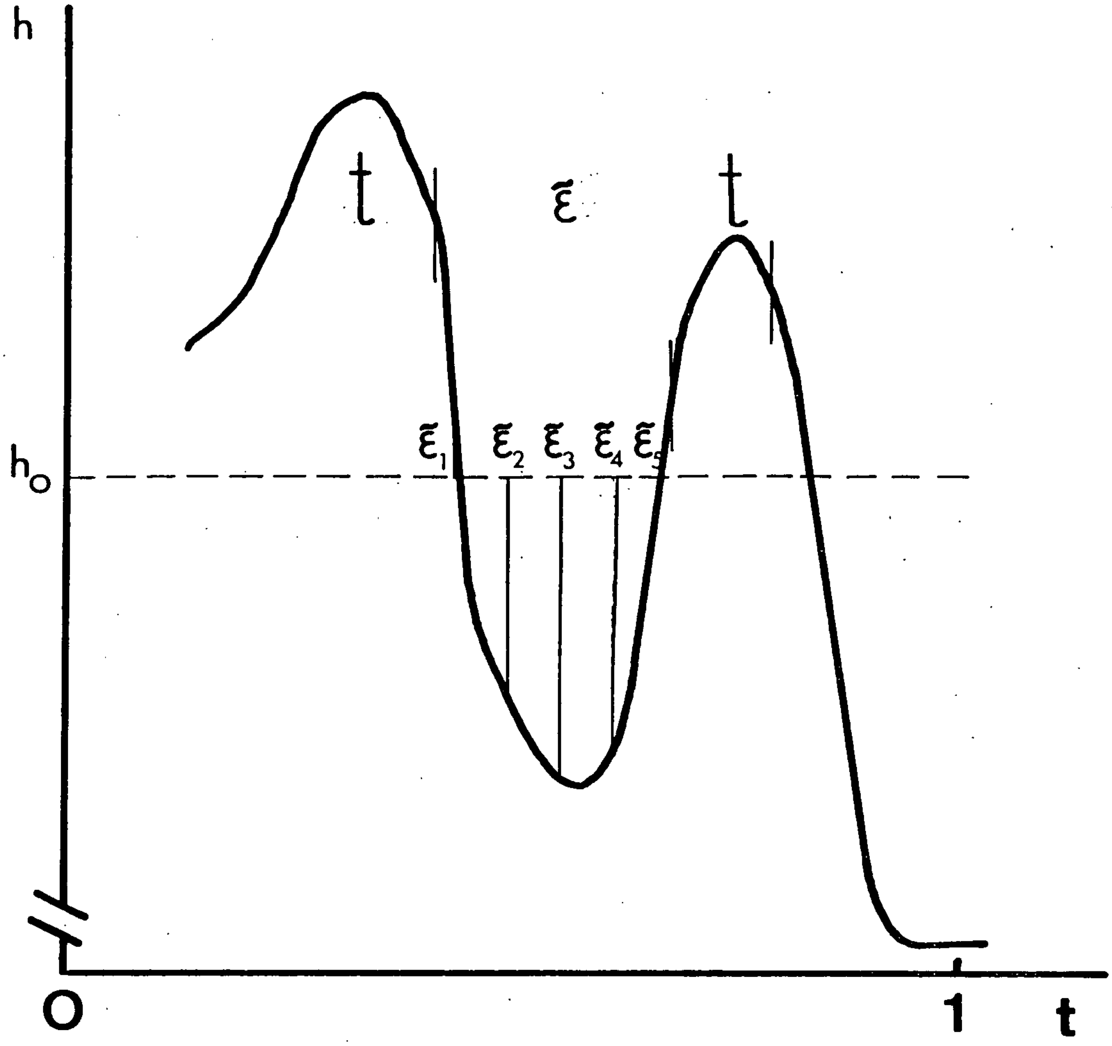

Fig. 6. Time correspondence between the five ɛ̃ stimuli and the velar height curve of Benguerel et al. (1975a)

Stimuli	Subjects' judgement	Velopharyngeal port
$\tilde{\alpha}_1$	Oral	Closed
$\tilde{\alpha}_2$	Oral	Open
$\tilde{\alpha}_3$	In-between	Open
$\tilde{\alpha}_4$	Nasalized	Open
$\tilde{\alpha}_5$	Nasalized	Closed
$\tilde{\varepsilon}_1$	Oral	Closed
$\tilde{\varepsilon}_2$	Nasalized	Open
$\tilde{\varepsilon}_3$	Nasalized	Open
$\tilde{\varepsilon}_4$	Nasalized	Open
$\tilde{\varepsilon}_5$	Oral	Closed
$\tilde{ɔ}_1$	Oral	Closed
$\tilde{ɔ}_2$	Nasalized	Open
$\tilde{ɔ}_3$	Nasalized	Open
$\tilde{ɔ}_4$	Nasalized	Open
$\tilde{ɔ}_5$	Nasalized	Open

Table VI. Comparison of the results of Test 1 (27-subject group) with the data of Benguerel et al. (1975a)

Stimuli	Subjects' judgement	Velopharyngeal port
ɑ1	Oral	Closed
ɑ2	Oral	Closed
ɑ3	In-between	Open
ɛ1	Oral	Closed
ɛ2	Oral	Closed
ɛ3	Oral	Closed
ɔ1	Oral	Closed
ɔ2	Oral	Closed
ɔ3	Oral	Closed

Table VII. Comparison of the results of Test 2 (27-subject group) with the data of Benguerel et al. (1975a) vowel stimuli by comparing the velar height (h) for a particular stimulus with the lower limit (h_0) where an opening of the velopharyngeal port starts. Tables VI and VII list these observations.

It should be recalled that the velum is at its lowest height for the production of open oral vowels such as /ɑ/. It was observed previously that $\tilde{\alpha}_1$ and $\tilde{\alpha}_2$ had average scores less than 3; that average scores for $\tilde{\alpha}_3$ were close to 3 and that $\tilde{\alpha}_4$ and $\tilde{\alpha}_5$ had average scores more than 3. To these stimuli correspond different positions of the velum in Benguerel et al. (1975a)'s plot of the time course of velar height for one token of /tɑ̃t/. As shown in Table VI, to $\tilde{\alpha}_1$ corresponds a high position of the velum, the velopharyngeal port is closed; to $\tilde{\alpha}_2$, $\tilde{\alpha}_3$ and $\tilde{\alpha}_4$ correspond a low position of the velum, the velopharyngeal port is open; to $\tilde{\alpha}_5$ corresponds a position of the velum slightly above the height which corresponds to an opening of the velopharyngeal port. The finding that some stimuli were identified as "oral" when the velopharyngeal was open or conversely, were identified as "nasalized" when the velopharyngeal port was supposedly closed might be explained by the fact that we are plotting our stimuli on a velar height curve obtained for another token of the utterance. Although, as was shown by Benguerel et al. (1975a, p. 71), the token-to-token repeatability was fairly high, the possibility of a slight difference in the timing of the movement, especially towards the end of the vowel and close to the consonant, may not be excluded.

Looking at Fig. 6 and Table VI, one can see that to the $\tilde{\varepsilon}_1$ and $\tilde{\varepsilon}_5$ stimuli, corresponds a velar height slightly greater than the limit height at which the velopharyngeal port is open. These vowels have an average score lower than 3. To the vowels $\tilde{\varepsilon}_2$, $\tilde{\varepsilon}_3$, $\tilde{\varepsilon}_4$ which had average scores greater than 3 corresponds a low velar position and an open velopharyngeal port. This suggests that the limit of velar height corresponding to an opening of the velopharyngeal port is an important factor in subjectively determining the "oral"/"nasalized" quality of the vowel [ε] in French in these experimental conditions. To $\tilde{ɔ}_1$ which had an average score less than 3 corresponds a velar height greater than the one necessary for an opening of the velopharyngeal port. All the other vowels constructed from /tɔ̃t/, i.e. $\tilde{ɔ}_2$, $\tilde{ɔ}_3$, $\tilde{ɔ}_4$, $\tilde{ɔ}_5$, have an opening of the velopharyngeal port, and an average score greater than 3. The same observation made for $\tilde{a}_4$ and $\tilde{a}_5$ might hold here to explain the fact that $\tilde{ɔ}_5$ was always heard as more "nasalized" than $\tilde{ɔ}_2$ for instance, when the velar position for $\tilde{ɔ}_5$ on the plot of Benguerel et al. (1975a) is higher than for $\tilde{ɔ}_2$. These results seem to agree with Massengill and Bryson (1967, p. 51)'s finding that

> "in some instances even though the subject produced complete velopharyngeal closure (closure indicated in the cinefluorographic analysis), the vowels produced were nevertheless judged as being nasal. On the other hand, some of the vowels were rated as being normal when there was as much as 8 mm opening between the velum and the pharyngeal wall".

Neither fiberoptic nor cinefluorographic methods enable the experimenters to directly observe, and measure, movements of the pharyngeal wall. The observable position of the velum might indicate a closure of the velopharyngeal port while in fact some air may be escaping along the sides of the pharyngeal walls and resonating in the nasal cavity. In particular, Minifie et al. (1970, p. 594) by using the Doppler ultrasound method showed that

> "the dynamic articulatory gestures within the pharynx vary systematically as a function of vowel height. The greatest inward movement of the LPW [lateral pharyngeal wall] occurs during the production of low vowels, thereby offering the greatest pharyngeal construction. There is very little movement of the LPW during the production of high vowels."

As can be seen in Table VII, to α_1 and α_2, which have an average score less than 3 correspond a high position of the velum and a closure of the velopharyngeal port. To α_3 which has a score close to 3 corresponds a low velar height which suggests an opening of the velopharngeal port. This stimulus was judged as being neither "oral" nor "nasalized" but rather as "in between" like $\tilde{\alpha}_3$ (constructed from /tãt/). This might be explained by the fact that Test 2 consisted of items which were judged by most people as being "oral". Although the nasalized reference vowels were all judged as "nasalized" by most listeners, α_3 was not judged "nasalized" but rather "in between". The average score, close to 3, suggests that the listeners noticed a slight difference in α_3, but the opening of the velopharyngeal port must not have been large enough to enable the listeners to assess a "nasalized" quality to the stimulus. This finding also seems to match Ali et al. (1971)'s observation that listeners could predict the nasal quality of a consonant if it followed a vowel /ɑ/. Our stimulus α_3 was constructed from a single cycle spliced out close to the beginning of the consonant /n/ in /ɑnɑ/. At that point, one may assume that the cue(s) which enabled Ali et al. (1971)'s listeners to predict the nasal quality of the following consonant, might also be present in the acoustic signal of α_3 and influence our listeners towards assessing an "in between" quality to α_3. All the stimuli constructed from /ɛnɛ/ and /ɔnɔ/ which correspond to high velum position and closed velopharyngeal port, were assessed as "oral" by the listeners.

4.4 Acoustic Analysis

An acoustic analysis of each of the constructed stimuli of Test 1 and of the oral reference vowels was carried out using an analysis-by-synthesis technique. The goal was to observe the variations in the acoustic parameters in vowel spectra for different positions of the velum.

To different positions of the velum correspond different cross-sections of the velopharyngeal port opening, thus different nasal coupling. House and Stevens (1956), among others, have studied variations in the spectra of vowels as nasal coupling increased. They used an electrical analog model and found that some of the effects in increasing nasal coupling on vowel spectra are:

1) a broadening and flattening of spectral peaks with most prominent changes in the region of the first formant;

2) a reduction of the third formant and even elimination of this peak;

3) the presence of an anti-resonance in the range 700-1800 Hz, its frequency increasing as the amount of coupling increases;

4) a double peak in the region of the third formant, when damping in the nasal tract is reduced.

The method used to examine variations in vowel spectra consisted of visually adjusting a theoretical acoustical spectrum based on Fant's (1960) model, to fit the spectra obtained using amplitude sections of the vowels of Test 1. The results, which are given in Figure 7, are only tentative and should be examined keeping in mind the conditions under which they were obtained. The use of a larger and faster computer would permit almost instant display of the synthesized spectral envelopes after changes have been made. The values given here, however, will be helpful as a preliminary study for further research using more sophisticated equipment.

Considering the values obtained for ɛ and for the stimuli constructed from /tɛ̃t/, one may see that:

1) the formant values obtained for ɛ are close to those of Peterson and Barney (1952, p. 183);

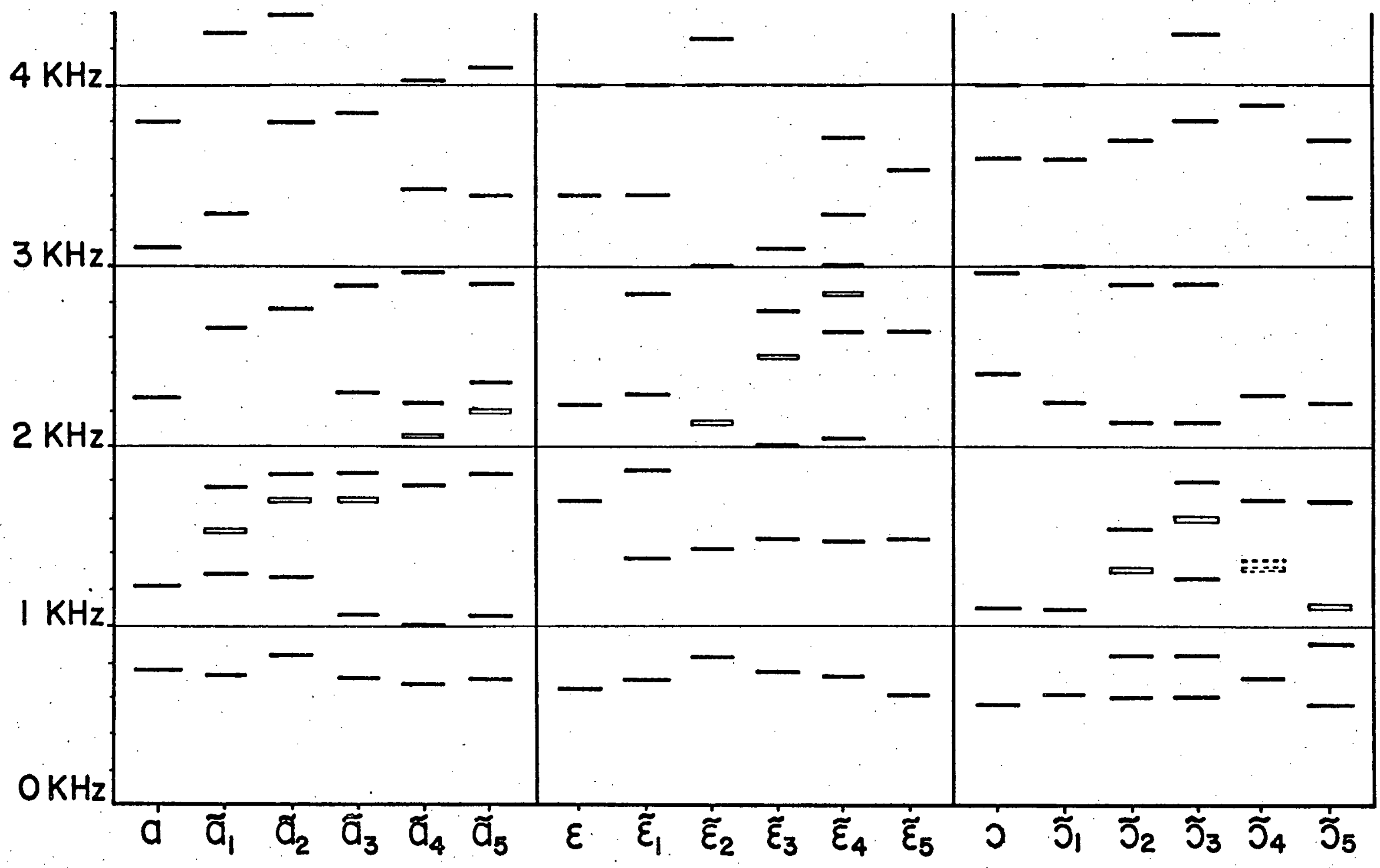

Fig. 7. Spectra of stimuli obtained by analysis-by-synthesis

2) the spectra of ε, $\tilde{\varepsilon}_1$ and $\tilde{\varepsilon}_5$ were fitted by a theoretical spectrum which contained no zero, in fact, the spectra for these stimuli were found to be relatively easy to fit; for these stimuli the corresponding velum position was raised and the velopharyngeal port closed; not surprisingly, they were heard as "oral" by the listeners;

3) the spectra of $\tilde{\varepsilon}_2$, $\tilde{\varepsilon}_3$, $\tilde{\varepsilon}_4$ seem to require at least one zero each; for these stimuli the velum was lowered and the port open; one would expect the spectrum of $\tilde{\varepsilon}_4$ to be similar to that of $\tilde{\varepsilon}_2$ since velar height was approximately the same; however this was not the case as the frequency of the zero of $\tilde{\varepsilon}_4$ is markedly higher than that of $\tilde{\varepsilon}_2$.

Looking at the values of the parameters for ɔ and for the vowels constructed from /tɔ̃t/, one may notice that:

1) the spectra for ɔ and for $\tilde{ɔ}_1$ are quite similar and the values of formant frequencies are close to those given for the vowel [ɔ] in Peterson and Barney (1952, p. 183); for these two stimuli the corresponding velum position was raised and the velopharyngeal port closed; both were judged "oral" by the listeners;

2) the spectra of $\tilde{ɔ}_2$, $\tilde{ɔ}_3$ and $\tilde{ɔ}_5$ seem to require at least one zero each, whose frequency seem to vary inversely with velar height;

3) the spectrum of $\tilde{ɔ}_4$ does not seem to require any zero; this spectrum was found to be relatively easy to fit; this may be due to the small number of measurable harmonic amplitudes on the amplitude section for this stimulus; it is more

likely, however, that a zero corresponding to the one present in $\tilde{ɔ}_3$ has moved slightly down in frequency and is so close to a nearby pole that they practically cancel each other out; thus, no peak or trough would be observable in the spectral envelope; the approximated position of this hypothesized zero-pole pair is shown in broken lines in Fig. 7.

4) F1 and F2 come close together in $\tilde{ɔ}_2$, when the velopharyngeal port is open; this results in a broadening of the peak on the vowel spectrum.

Considering the values obtained for the spectra of ɑ and of the stimuli constructed from /tɑ̃t/ one can see that:

1) the spectrum of ɑ does not require any zero and the frequency of the formants are similar to those given by Peterson and Barney (1952, p. 183) for the corresponding vowel;

2) the spectra of the vowels constructed from /tɑ̃t/ seem to require a zero in the theoretical spectra in order to obtain a satisfactory match; the frequency of that zero increases steadily from $\tilde{ɑ}_1$ to $\tilde{ɑ}_5$, whereas the velum lowers until $\tilde{ɑ}_3$ and starts rising thereafter; this may be due to a change in the configuration of the oral or pharyngeal cavities, associated with velar movement;

3) comparing the spectra of $\tilde{ɑ}_1$ and $\tilde{ɑ}_2$ on the one hand and those of $\tilde{ɑ}_3$, $\tilde{ɑ}_4$ and $\tilde{ɑ}_5$ on the other hand, it seems that an added resonance occurs at approximately 2300 Hz in the spectrum of the vowels of the latter group.

Considering the results obtained for the stimuli constructed from /tɛ̃t/ on the one hand and for those constructed from /tɑ̃t/ and /tɔ̃t/ on the other hand, one may see that the frequency of the zeros (when they exist) is in general higher for the /tɛ̃t/ stimuli than for the /tɔ̃t/ and the /tɑ̃t/ stimuli. This could be related to the fact that ɔ̃ and ɑ̃ are both more open than ɛ̃, although this needs further investigation.

Fig. 7 also shows that, the frequency of the zero increases more rapidly from $\tilde{\varepsilon}_2$ to $\tilde{\varepsilon}_4$ than it does from $\tilde{ɔ}_2$ to $\tilde{ɔ}_4$ or from $\tilde{ɑ}_2$ to $\tilde{ɑ}_4$. This is not surprising if one considers that the oral vocal tract impedance for ɛ̃ is higher than that for ɔ̃ and ɑ̃. A given nasal coupling will thus have a greater effect (percentage wise) on ɛ̃ than on ɔ̃ or ɑ̃ (see House and Stevens, 1956 p. 223).

A close examination of Fig. 7 shows that there exists a correlation between the position of zeros in a vowel spectrum and the position of the velum. However, it was found that two velar height values which were approximately equal (such as in $\tilde{\varepsilon}_4$ and $\tilde{\varepsilon}_2$) did not correspond to the same zero frequencies. This suggests that the frequency of a zero is not in a one-to-one relation with velar height.

Chapter 5

SUMMARY AND DISCUSSION

The main goal of this study was to try to answer the following questions:

1) What is the maximum amount of velopharyngeal port opening allowable for a vowel, (produced by a French speaker), such that it will continue to be perceived as phonemically oral by a French listener?

2) Since it was found by Benguerel et al. (1975a) that the velum position in an utterance such as /tɑ̃t/ is changing rapidly through the vowel, and since a portion of the phonemically nasalized vowel /ɑ̃/ is produced with a closed velopharyngeal port, does the listener perform some kind of time integration to decide whether this vowel is oral or nasalized?

3) What kind of integration is performed? Is it related directly to the average velar cross-section or to some more complex auditory function of the velopharyngeal cross-section?

Two listening tests were prepared and presented to French listeners. Single cycles, spliced out of French vowels and corresponding to known velar height were used to construct the test stimuli. Each test also contained 6 reference vowels (oral and nasalized) as controls to evaluate the listeners' responses.

Test 1 consisted of 5 presentations of the 6 reference vowels ɑ, ɛ, ɔ, ɑ̃, ɛ̃, ɔ̃ and of 15 stimuli constructed from cycles isolated at 5 different places in the vowel of each of the 3 utterances /tɑ̃t/, /tɛ̃t/, /tɔ̃t/.

Test 2 consisted of 5 repetitions of the 6 reference vowels and of 9 stimuli constructed from cycles isolated at 3 different places in the first vowel of the 3 utterances /ɑnɑ/, /ɛnɛ/, /ɔnɔ/. All the utterances were produced by a male native speaker of French. Listeners recorded their answers to each stimulus on a 5-point scale, 1 corresponding to "oral" and 5 to "nasal". Subjects' responses were sorted into tables according to stimulus type and order of presentation (see Tables I and II). The subjects who performed best were selected according to two consistency criteria.

Subjects' scores as a group were compared to those of the whole population. It was found that for each vowel type and for each test, curves representing nasality judgement (see Fig. 2 to Fig. 5) have basically the same shape, whether the scores were averaged on five or on four responses, whether for the 9-subject group or for the 27-subject group, and whether the scores were normalized or not.

By examining the time correspondence between the stimuli of this study and the unpublished graphs of Benguerel et al. (1975a) which plot the time course of velar height for the same utterances (e.g. Fig. 6), it was possible to associate a position of the velum (thus the presence or absence of an opening of the velopharyngeal port) with each one of the stimuli.

In table VI, it was seen that the transitions from "oral" to "nasal" judgements occurs earlier (on a 5-point scale) for the items constructed from /tɛ̃t/ and /tɔ̃t/ than for those constructed from /tɑ̃t/. For example, the transition for ɛ̃ and ɔ̃ is apparently made between $\tilde{\varepsilon}_1$ and $\tilde{\varepsilon}_2$ or between $\tilde{ɔ}_1$ and $\tilde{ɔ}_2$ whereas for ɑ̃ it is made at $\tilde{ɑ}_3$. This suggests that a rather large velopharyngeal cross-section may be necessary for a French listener to perceive the vowel [ɑ] as nasalized, whereas a smaller cross-section may be sufficient to give the vowels [ɛ] and [ɔ] a nasalized quality. These results agree with those of House and Stevens (1956). In their study of nasalization of vowels using an analog model, they found (ibid., p. 223-230) that "as the average area of coupling increases, the 'more nasal' responses to /i/ and /u/ start sooner and increase faster than do such responses to /ɛ/ and /ɔ/, and the nasal responses to /ɑ/ are the last to be manifested." In other words, a smaller coupling would be necessary to give a nasalized quality to /ɛ/ and /ɔ/ than to /ɑ/. Moreover, the well-known fact that the oral vowel [ɑ] is produced with very low position of the velum might also explain this phenomenon. The observation that $ɑ_3$, constructed from /ɑnɑ/, was judged as "in between" when the velum was at a position slightly below that where an opening of the velopharyngeal port starts, could be used as a counter-argument. However, as was mentioned in Section 4.22,the fact that most items of Test 2 were identified as "oral" might have influenced the listeners towards judging α_3 as "in between". In the pre-test instructions given to the listeners, they were told not to assume any particular distribution of the stimuli and were encouraged to note down what they heard even if they thought everything was "oral". It is possible that the nature of α_3 was such that it sounded "different" to the listeners simply because of the high number of stimuli heard as "oral" in Test 2.

The results of this study show that in general, and as expected, subjects judged stimuli as "oral" when the velum was raised and the velopharyngeal port closed, while they rated the stimuli as "nasalized" whenever the velum was lowered and the velopharyngeal port open. The curves of Fig. 2 to Fig. 5 also show that stimuli are rated as more "nasalized" as the velum is in a lower position. A relationship appears to exist between the curves representing the listeners' judgement of nasality and those representing the timing of velar movement. It seems that this relationship may not be monotonic, since it was observed that $\tilde{\alpha}_2$ was perceived as more "oral" than $\tilde{\alpha}_1$, when in fact the velum was at a lower position for $\tilde{\alpha}_2$ than for $\tilde{\alpha}_1$.

In general, $\tilde{\alpha}$ was rated as less "nasalized" than ɛ̃ or ɔ̃. In order to investigate this finding the time course of velar movement was determined for the reference nasalized vowels. The unpublished plots of Benguerel et al. (1975a) for the time course of velar height in the 3 utterances /tɑ̃t/, /tɛ̃t/, /tɔ̃t/ were used. Considering the high token-to-token repeatability, and assuming an approximate linearity in the timing of velar movement, it was found that a relatively large portion of ɑ̃ (about 30%) had been produced with the velopharyngeal port not yet open, whereas, the velum was lowered completely for the duration of ɔ̃, suggesting a relatively large cross-section throughout. This finding could thus explain why ɑ̃ was rated as less "nasalized" than ɔ̃. Since the fiberoptic method does not enable one to actually quantify the velopharyngeal port cross-section, further studies need to be conducted to correlate the results of the perceptual experiment presented here with values of the cross-section of the velopharyngeal port. Warren and Dubois (1964) have presented a pressure-flow technique whereby one is able to quantify the cross-section, using aerodynamic parameters. Section 2.31 of this thesis describes the method and a possible experimental set up using the equipment of Benguerel (1974). An experiment could be conducted to determine the time course of velopharyngeal port opening during the production of utterances similar to those used in this study. One would then be able to associate a relative value of port opening cross-section to each stimulus of the present study.

It appears that listeners do perform some kind of time integration when they judge a vowel as "nasalized" in French. The fact that the stimuli constructed from single cycles isolated at different places in the nasalized vowel were heard as "oral", "nasalized" or "in-between" by the listeners (depending on the place in the vowel where the cycles had been isolated) and that the reference nasalized vowels as a whole were heard as "nasalized" strongly supports this hypothesis. It appears that the percentage of a nasalized vowel during which the velopharyngeal port is closed plays an important part in the listener's judgement of relative degree of nasality. For example, it was suggested that the reference vowel ɑ̃ was judged less "nasalized" than ɔ̃ because a relatively large portion of ɑ̃ was produced with a closed velopharyngeal port, whereas ɔ̃ which was more often identified as "nasalized" was produced with a low position of the velum and an open velopharyngeal port throughout. Further studies could be done to investigate this finding. Vowels could be spliced out at different places from the utterances /tɑ̃t/, /tɛ̃t/, /tɔ̃t/, thus containing different proportions of vowel produced with closed velopharyngeal port. Listeners would then have to evaluate the degree of nasality of these stimuli. One would expect the task to be rather difficult for the listeners, because of the minimal differences between stimuli. As to the question of whether this time integration is related directly to the average velar passage or to some more complex function, this could be investigated by running an experiment such as the one described above where one would correlate the time course of (quantified) velopharyngeal opening with the nasality judgements obtained in the same way as in the present study.

The acoustic analysis reported in the last part of this study suggests that there exists some relation between the presence of zeros in the spectrum of vowels, the corresponding position of the velum and the perception of nasalization. A flattening of the first peak in the spectrum may also occur as the velum lowers. The position of the zeros however does not seem to vary systematically with the movement of the velum. Zeros, when they were present, were also found to have a higher frequency in the spectrum of ɛ̃ than in the spectra of ɑ̃ and ɔ̃ suggesting a connection between zero frequency and degree of opening of the vowel. More studies need to be done in order to investigate further the acoustics of nasalization.

The ultimate goal would be to correlate results obtained in studying the production, the acoustics and the perception of nasalization. The various possible experiments suggested in this study will be hopefully carried out at a later date. They will somewhat clarify some of the mysteries of the human speech process.

BIBLIOGRAPHY

Ali, L., Gallagher, T., Goldstein, J. and Daniloff, R. (1971). "Perception of Coarticulated Nasality," J. Acoust. Soc. of Amer., 49, 538-540.

Bell-Berti, F. and Hirose, H. (1975). "Palatal Activity in Voicing Distinctions: A simultaneous Fiberoptic and Electromyographic Study," J. of Phonetics, 3, 69-74.

Benguerel, A-P. (1974). "Nasal Airflow Patterns and Velar Coarticulation in French," Speech Communication Seminar Proceedings, Vol. II, Almquist & Wiksell International, Stockholm, 105-112.

Benguerel, A-P., Hirose, H., Sawashima, M. and Ushijima, T. (1975a). "Velar Height and its Timing in French: A Fiberoptic Study," Ann. Bull., Res. Inst. of Logopedics and Phoniatrics (Univ. of Tokyo), No. 9, 67-78.

Benguerel, A-P., Hirose, H., Sawashima, M. and Ushijima, T. (1975b). "Electromyographic Study of the Velum in French," Ann. Bull., Res. Inst. of Logopedics and Phoniatrics (Univ. of Tokyo), No. 9, 79-90.

Fant, G. (1960). Acoustic Theory of Speech Production, The Hague, Mouton.

Fletcher, S.G. (1970). "Theory and Instrumentation for Quantitative Measurement of Nasality," Cleft Palate J., 7, 601-609.

Fletcher, S.G. (1972). "Contigencies for Bioelectronic Modification of Nasality," J. Speech and Hearing Dis., 37, 329-346.

Fletcher, S.G. (1973). "Perceptual Skills in Clinical Management of Nasality," Folia Phoniat., 25, 137-145.

Fletcher, S.G. (1976). ""Nasalance" Vs. Listener Judgements of Nasality," Cleft Palate J., 13, 31-44.

Fletcher, S.G. and Bishop, M.E. (1970). "Measurement of nasality with TONAR," Cleft Palate J., 7, 610-621.

Fritzell, B. (1969). "The Velopharyngeal Muscles in Speech on EMG and Cineradiographic Study," Acta Otolaryngologica Suppl., 250.

Fujimura, Q., Ishida, H. and Kiritani, S. (1968). "Computer Controlled Dynamic Cineradiography," Ann. Bull., Res. Inst. of Logopedics and Phoniatrics (Univ. of Tokyo), No. 2, 6-10.

Hirose, H. (1971). "Electromyography of the Articulatory Muscles in Speech: Current Instrumentation and Technique," Haskins Lab. Status Report on Speech Res. SR-25/26, 73-86.

House, A.S. and Stevens, K.N. (1956). "Analog Studies of the Nasalization of Vowels," J. Speech Hearing Dis., 21, 218-232.

Kelsey, C.A., Minifie, F.D. and Hixon, T.J., (1969). "Applications of Ultrasound in Speech Research," J. Speech and Hearing Res., 12, 564-575.

Kewley-Port, D. (1973). "Computer Processing of EMG Signals at Haskins Laboratories," Haskins Lab. Status Report on Speech Res. SR-33, 173-183.

Lubker, J.F. (1968). "An Electromyographic-Cinefluorographic Investigation of Velar Function During Normal Speech Production," Cleft Palate J., 5, 1-17.

Lubker, J.F. and Moll, K.L. (1965). "Simultaneous Oral-Nasal Airflow Measurements and Cinefluorographic Observations During Speech Production," Cleft Palate J., 2, 257-272.

Massengill, R. and Bryson, M. (1967). "A study of Velopharyngeal Function as Related to Perceived Nasality of Vowels, Utilizing a Cinefluorographic Television Monitor," Folia Phoniat., 19, 45-52.

Minifie, F.O., Hixon, T.J., Kelsey, C.A. and Woodhouse, R.J. (1970). "Lateral Pharyngeal Wall Movement During Speech Production," J. Speech and Hearing Res., 13, 584-594.

Minifie, F.D., Kelsy, C.A., and Hixon, T.J. (1968). "Measurement of Vocal Fold Motion Using an Ultrasonic Doppler Velocity Monitor," J. Acoust. Soc. of America, 43, 1165-1169.

Moll, K.L. (1960). "Cinefluorographic Techniques in Speech Research," J. of Speech and Hearing Res., 3, 227-241.

Moll, K.L. (1964). "Objective Measures of Nasality - a Letter to the Editor," Cleft Palate J., 1, 371-374.

Moll, K.L. and Daniloff, R. (1971). "Investigation of the Timing of Velar Movement during Speech," J. Acoust. Soc. of America, 50, 678-684.

Moll, K.L. and Shriner, T.H. (1967). "Preliminary Investigation of a New Concept of Velar Activity During Speech," Cleft Palate J., 4, 58-69.

Ohala, J.J. (1971). "Monitoring Soft Palate Movements in Speech," Univ. of California at Berkeley Phonology Lab., Project on Linguistic Analysis Report, No. 13, J01-J015.

Peterson, G.E. and Barney, H.L. (1952). "Control Methods Used in a Study of the Vowels," J. Acoust. Soc. of America, 24, 175-184.

Peterson, G.E. and Shoup, J.E. (1966). "A Physiological Theory of Phonetics," J. Speech and Hearing Res., 9, 5-67.

Sawashima, M. and Hirose, H. (1968). "New Laryngoscopic Technique by Use of Fiberoptics," J. Acoust. Soc. of America, 43, 168-169.

Shelton, R.L., Knox, A.W., Arndt, W.B. and Elbert, M. (1967), "The Relationship Between Nasality Score Value and Oral and Nasal Sound Pressure Level," J. Speech and Hearing Res., 10, 549-557.

Spriestersbach, D.C. and Powers, G.R. (1959). "Nasality in Isolated Vowels and Connected Speech of Cleft Palate Speakers," J. Speech and Hearing Res., 2, 40-45.

Ushijima, T. and Hirose, H. (1974). "EMG Study of the Velum During Speech," J. of Phonetics, 2, 315-326.

Ushijima, T. and Sawashima, M. (1972). "Fiberscopic Observation of Velar Movements During Speech," Ann. Bull., Res. Inst. of Logopedics and Phoniatrics (Univ. of Tokyo), No. 6, 25-38.

Warren, D.W. and Dubois, A.B. (1964). "A Pressure-Flow Technique for Measuring Velopharyngeal Orifice Area During Continuous Speech," Cleft Palate J., 1, 52-71.

Weatherley-White, R.C.A., Dersh, W.C., and Anderson, R.M. (1964). "Objective Measurement of Nasality in Cleft Palate Patients: A Preliminary Report," Cleft Palate J., 1, 120-124.

APPENDIX

Instructions

Vous allez entendre une série de voyelles.
Votre tâche sera d'évaluer le degré de nasalité de chacune des voyelles.
Si la voyelle vous semble orale, comme dans les mots: bal, bel, bol, cochez la case de gauche.
Si la voyelle vous semble nasale, comme dans les mots: pan, pin, pont, cochez la case de droite.
Si la voyelle vous semble intermédiaire, utilisez celle des trois autres cases qui correspond le mieux à votre jugement; par exemple, si la voyelle vous semble plus orale que nasale vous utilisez la seconde case, etc. .

MERCI