Ann. Bull. RILP, No. 9, 67-78 (1975)

VELAR HEIGHT AND ITS TIMING IN FRENCH : A FIBERSCOPIC STUDY

A-P. Benguerel, H. Hirose, M. Sawashima and T. Ushijima

Introduction

The purpose of this study is to examine the temporal course of velar movement in French nasal consonants and nasalized vowels by direct observation through a fiberscope.

Information on velopharyngeal opening can be obtained by several experimental methods, but none of these provides exactly the same type of information; in particular, direct observation and filming of velar movement by means of a wide angle fiberscope (the technique used in this study) is the only technique which can presently provide both a measure of velar height and a measure of velopharyngeal cross section. 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.

A previous study by one of the authors (Benguerel, 1974), attempted to investigate the same problems, using nasal airflow patterns, but some inconclusive or ambiguous results had to be resolved using a different method. Another study (Ushijima and Sawashima, 1972), investigated similar problems in Japanese, but differences between the two languages warrant a separate study of French nasal sounds. French has nasal consonants, as most languages, but also nasalized vowels which are phonemically distinct from oral vowels.

Experimental Procedure

In the following, C and V (without any subscript) stand for consonant and vowel respectively, whether oral or nasal(ized). The subscript o restricts each category to oral sounds, while the subscript n restricts C to nasal consonants and V to nasalized vowels.

The corpus consisted of recorded and filmed material, as follows:
a) two sets of isolated sustained vowels separated by a breathing pause;
b) CVC utterances (without frame sentence), where C was /t/ or /s/, and V was /a/, /ã/, /ɛ/, /ɛ̃/, /ɔ/ or /ɔ̃/;
c) V_1CV_2 utterances (without frame sentence), where C was /t/ or /n/, V_1 was /a/, /ɛ/, or /ɔ/, and V_2 was either identical to V_1, or its nasalized counterpart;
d) V_1CV_2 utterances in the frame "répétez/VCV/deux fois"; the same V_1CV_2 were used as for the previous group;
e) two-word sentences in which a vowel (/a/, /ɛ/ or /ɔ/) or its nasalized counterpart was preceded by the four-consonant cluster /-rzfr-/, e. g. "quatorze frangines": /katɔrzfrãʒin/;
f) utterances in which the consonant /n/ was preceded by /kre-/, /kree-/, or /kreee-/, or by /œ̃brua-/, /œ̃bruaa-/ or /œ̃bruaaa-/.

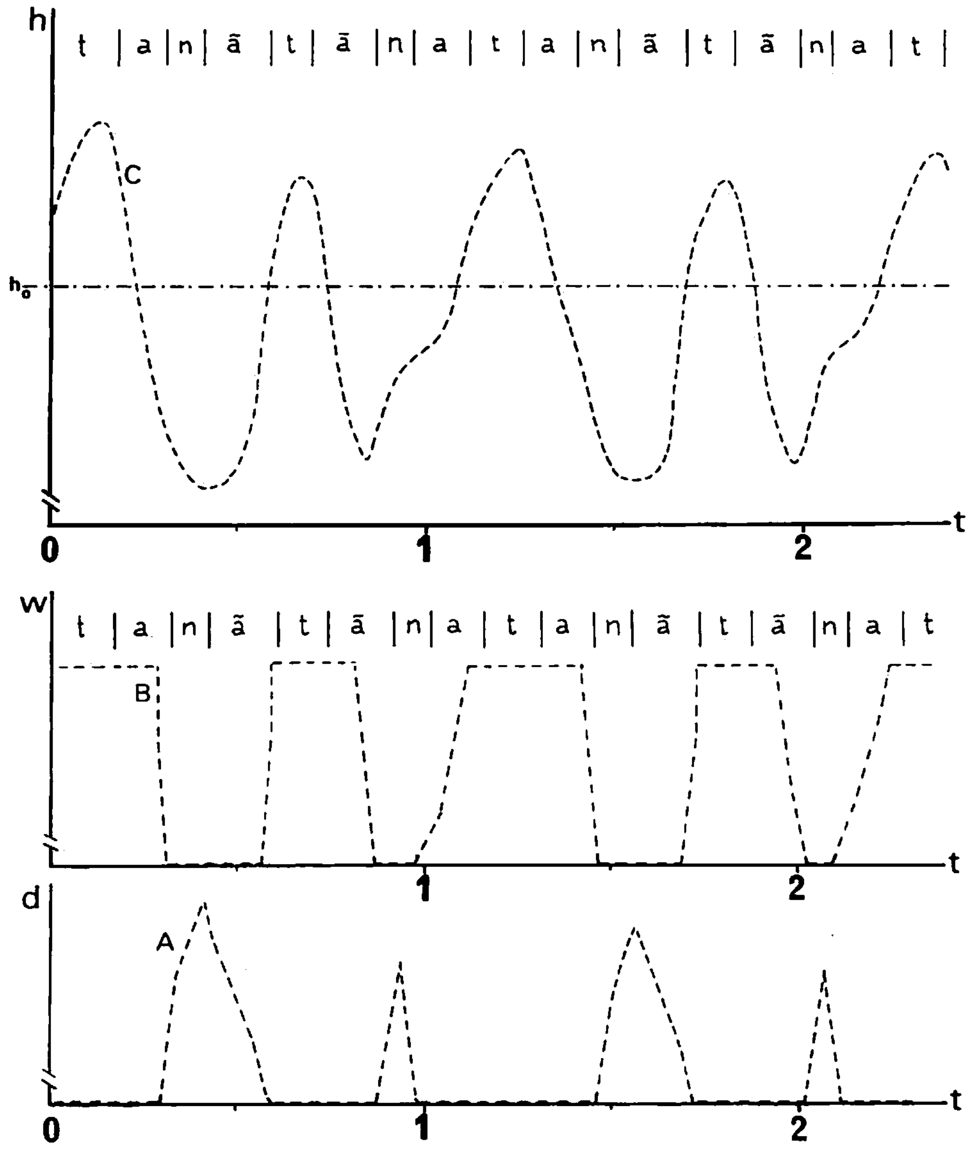

Fig. 1: Velar distance d (curve A), width of velopharyngeal contact w (curve B) and velar height h (curve C); cf. text.

Utterances were randomized by group, separated by a breathing pause, and read at least twice, although in many cases three or four times, with normal intonation and at a comfortable level.

The subject for this study was a male native speaker of French, originally from Lausanne, Switzerland.

A wide angle fiberscope was inserted into the nasal passage of the subject, under light topical anesthesia, and positioned in the inferior nasal meatus (except for the calibration sequences — cf. below — where it was positioned in the middle meatus) and clamped to a headband, in front of the nostril. In this position, the nasal side of the velum — and the upper margin of velopharyngeal contact — could easily be observed, except when it was lowered to its extreme position, as in the case of the rest position.

The distal end of the fiberscope was connected to a movie camera operating at 50 fps. Speech was recorded, via a directional electret condenser microphone, onto one track of a magnetic tape, while synchronization pulses were recorded onto a second track.

Measurements and Results

On all corpus items, velar height was measured in a way similar to that used and described by Ushijima and Sawashima (1972). However, in order to establish first the relation between velar height and velar opening, two "calibration" sequences were recorded, filmed and measured. The utterance /tãnatanã/ repeated three times without pause was filmed first (no. 1) with the fiberscope inserted through the middle meatus with its tip bent downward, the larynx being fully visible when the velum was lowered and the root of the tongue not retracted; the same utterance was filmed a second time (no. 2) with the fiberscope inserted through the inferior meatus with its tip almost horizontal. That last position was kept unchanged for the rest of the film session.

For calibration sequence no. 1, two measurements were taken and plotted in Fig. 1: A) the velopharyngeal distance d, measured on an arbitrary (but consistent) scale, in the sagittal plane (in which the velopharyngeal contact always lasted the longest); B) the width of velopharyngeal contact w, measured on an arbitrary scale perpendicular to the previous one. For calibration sequence no. 2, one measurement only was taken: C) the velar height h, also measured on an arbitrary scale. Curves A, B and C are plotted on Fig. 1, but it should be kept in mind that curves A and B pertain to one utterance and curve C to a repetition of this utterance. Although very similar in timing, the two are not identical; however, they can be correlated by observing velar position at the time of specific articulatory events such as the closure, or the opening of the vocal tract, or the onset of voicing for example. From the three curves A, B and C, it can be safely stated that when $h > h_0$ (cf. figures), the velopharyngeal port is completely closed. For $h < h_0$, there is a definite, however small, opening of the velopharyngeal port, and when $w = 0$, i. e. $d > 0$, there is no longer a velopharyngeal contact, although the velopharyngeal opening may be significant for greater values of h (i. e. higher positions of the velum).

Although several frames show considerable blur, particularly at those times when the velum is lowering or rising very rapidly, the accuracy of the measurements of h on the chosen arbitrary scale is ± 1 mm (on a picture enlarged 30 times with respect to the original film, and on which the range of velar movement is greater than 10 cm). Repeatability of measurement was randomly tested and found to satisfy this tolerance.

The time relation between the speech signal and each film frame was established by means of oscillograms displaying the speech signal and the recorded pulses synchronous with each aperture of the camera shutter. The accuracy of this time relation measurement can be estimated to be at worst ± 5 ms.

a) Analysis of isolated vowels

The vowels recorded are: /i/, /e/, /ɛ/, /a/, /ɑ/, /ɔ/, /o/, /u/, /y/, /ø/, /œ/, /ɛ̃/, /ɑ̃/, /ɔ̃/, /œ̃/. For oral vowels, the time course of velar movement rises to a maximum velar height near the middle of the vowel and decreases toward the end of the vowel. Velar height is found to correlate positively with the degree of opening of the vowel, as observed by Fant (Fant, 1960, p. 139), but not with vowel height, as reported by Moll (1962). However, Moll's data contain only four vowels (/i/, /u/, /æ/ and /ɑ/) and his results can thus be interpreted in terms of height as well as in terms of jaw opening. Greatest vowel height is found for /ɔ/, /œ/, /ø/, /y/, and /u/ in this order, while least velar height is found for the open vowels /a/ and /ɑ/; for these two vowels, there is a very slight opening of the velopharyngeal port. For nasalized vowels, the velar movement time pattern is different: the velum rises to a maximum height, approaching complete closure 100 to 200 ms before vowel onset (some kind of speech preparatory position), then decreases before reaching in inflection point (or sometimes a local maximum) in the middle portion of the vowel, then

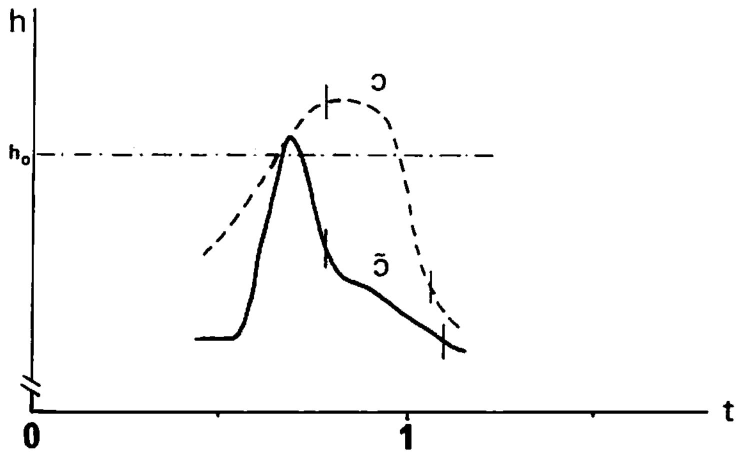

Fig. 2: Velar height for two isolated vowels.

decreases further down to complete opening in the last part of the vowel or after the end of phonation. This can be observed in Fig. 2. Maximum velar height (in the middle of the vowel) is markedly higher for /ɔ̃/ than for the other three nasalized vowels.

Observation of velar height in VCV's, CVC's and longer utterances supports the above observations; in particular, ceteris paribus, velar height for /ɔ/ and /ɔ̃/ is always greater than for other vowels.

b) Analysis of CVC utterances

For all of these utterances, the velar movement pattern shows two maxima — one on each consonant, the first one always markedly higher than the second one — separated by a mimimum on the vowel, whether oral or nasalized. There is no difference in maximum velar height between t's and s's when their position is the same, but for /tV_ot/ utterances, lowering of the velum starts hardly before the onset of the vowel whereas for both /tV_nt/, /sV_os/ and /sV_ns/ utterances, it starts during the first consonant. In all /CV_nC/ cases, however, 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.

Fig. 3 shows the time course of velar height for two different tokens of /tɔt/ and for one of /tɔ̃t/. In addition to the facts noted above, one can observe among other things the high and typical token-to-token repeatability.

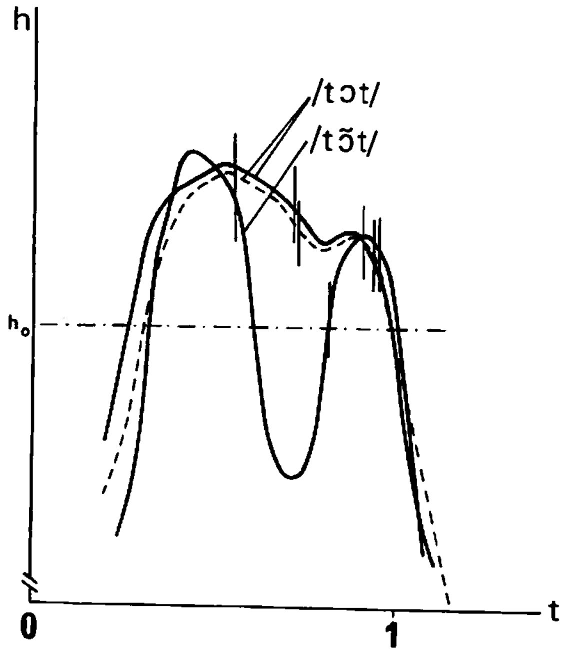

Fig. 3: Velar height for two tokens of /tɔt/ and for /tɔ̃t/.

c) Analysis of /VCV/ sequences in isolation

This group of utterances shows a slightly less good token-to-token repeatability than the other groups, mainly in the time dimension. This is not too surprising however, in view of the fact that timing is usually more crucial for consonants than for vowels. Since utterances in this group contain only one consonant, one can expect less precise overall timing.

$/V_oCV_o/$ sequences all show greater velar height on the initial vowel than on the final one. This seems to parallel a similar tendency on consonants, as mentioned earlier, and as observed on the calibration data where the same sequence (/tãnatanã/) was repeated up to four times in a row.

For all $/V_oC_nV_n/$ sequences, one can observe two distinct velum lowering gestures, the first one for C_n, starting near (sometimes before) the beginning of V_o, the second one for V_n, starting during the second half of C_n. Fig. 4 illustrates this point.

For all $/V_oC_nV_o/$ sequences, the velopharyngeal port remains open until after the end of C_n, however for several tokens involving /ɛ/'s and /ɔ/'s, it opens only after (up to 30 ms) the onset of C_n. Although surprising, this seems quite possible since the vocal folds can vibrate with simultaneous oral and velar closures as long as the oral pressure build-up is not too high.

/ana/ sequences show that in the final /a/, the velopharyngeal port is completely closed for less than 20% of its duration. This correlates with the finding of lowest velar height in open vowels. For all $/V_oC_nV_o/$ sequences without exception (and this holds also when they are in a frame), the second vowel has more velum lowering than the first one. This is in disagreement with claims for Hindi by Kelkar (1968) and observations on English by Ohala (1970).

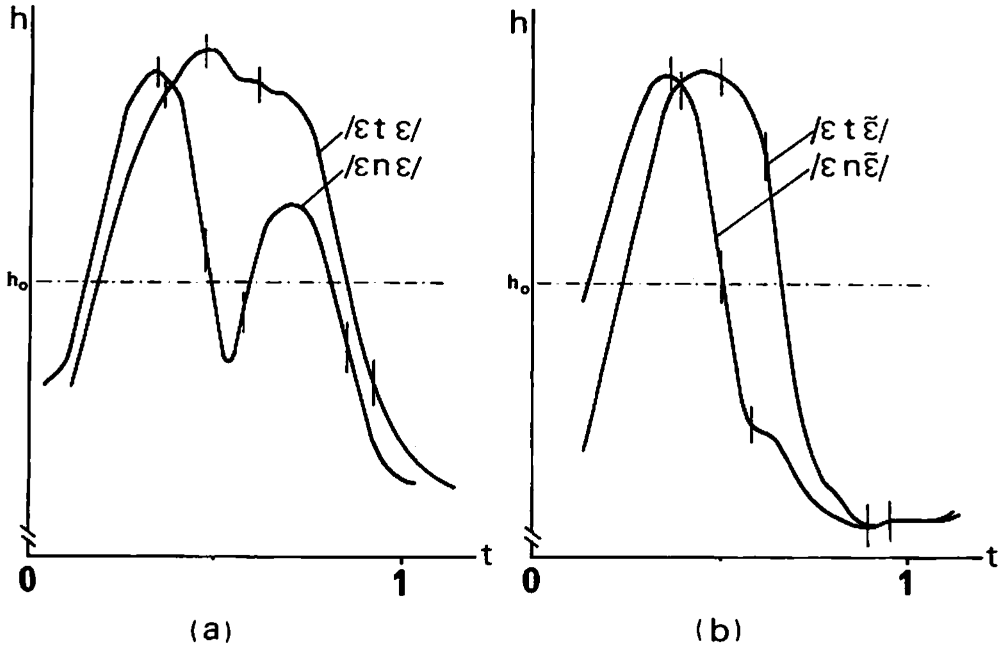

Fig. 4: Velar height for a) /ɛtɛ/ and /ɛnɛ/
b) /ɛtɛ̃/ and /ɛnɛ̃/.

d) Analysis of /VCV/ sequences in frame

This group of data shows a very high degree of token-to-token repeatability. The two gestures observed in isolation for $/V_oC_nV_n/$ are also visible here but in a different way which is schematized in Fig. 5. In this case (i. e. with frame), the second gesture cannot be one of further lowering, due to the constraint of the upcoming /d/; nevertheless, it maintains the velopharyngeal port completely open as long as possible, and after an extremely rapid rise, completes the velar closure in practical synchronization with the oral closure of /d/. This synchronized double closure is normal in most dialects of the northern half of France and in Belgium and Switzerland. In dialects of the South of France, on the other hand, oral closure always occurs first, introducing an extra nasal segment for the time interval in which there is oral closure but no velar closure yet.

For all $/V_oC_nV_o/$ sequences, the velum is almost fully lowered at the beginning of C_n (unlike in the case of $/V_oCV_o/$'s without frame) as well as at the end of C_n (like in the case of $/V_oCV_o/$'s without frame). This means that, acoustically speaking, both V_o's have substantial nasalization, and yet they are phonemically non-nasalized. This point deserves further investigation.

All $/V_oC_oV_o/$ sequences in this group demonstrate very clearly that velar height reaches a local maximum during each oral consonant and a local minimum during each intervening vowel. For the other sequences of the group, the same still holds for oral sounds except for oral vowels next to a nasal sound. Note that all sequences in this group have a /CVCVCVVCVCVCV/ structure.

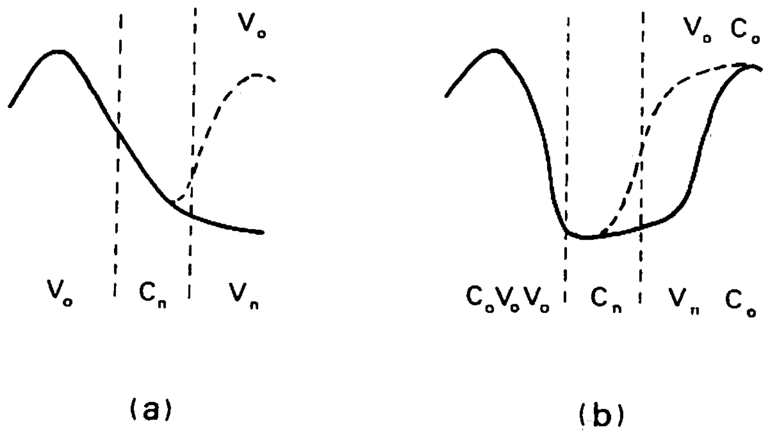

Fig. 5: Schematized velar height patterns for a) /VCV/'s in isolation
b) /VCV/'s in a frame sentence

e) Analysis of /-rzfrV- / sequences

This group also shows a high token-to-token repeatability. Each utterance, similarly to those of the previous group, has local maxima corresponding to tne closure periods of stop consonants (or fricatives in a few cases). At the end of the first /r/(at the beginning of the /rzfr/ cluster) the velar height curve always has a local minimum, or at least a noticeable bend. At the end of the second /r/, thus of the cluster, velar height is nearly the same when V_o follows. However, when the cluster is followed by V_n the velum starts its downward course during /z / (the second consonant of the cluster) and by the end of the cluster, it has reached its maximum downward speed. The velopharyngeal port starts opening at the end of /f/ or during the second /r/. The upward velar movement following V_n is statistically significantly faster than the lowering movement that precedes it. Contextual asymmetry however, prevents drawing any conclusion from this last observation. Fig. 6 illustrates velar height change for the two sequences "quatorze fraudeuses" - /katɔrzfrodøz/ and "quatorze frondeuses" - /katɔrzfrɔ̃døz/.

f) Analysis of the /-V(V(V))C_n-/ sequences

Each utterance in this group contained a sequence /-V(V(V))C_n-/, i. e. a given vowel (/e/ or /a/) occurred one, two or three times in succession without cessation of phonation, preceding the consonant /n/. A t-test showed that vowel durations for /e/, /ee/ and /eee/ were all significantly different beyond the 0. 001 level. The average durations and their standard deviations are given in the following table:

	$\bar{x}$ (ms)	σ (ms)
— e —	99. 4	6. 1
— ee —	278. 5	27. 9
— eee —	345. 3	48. 9

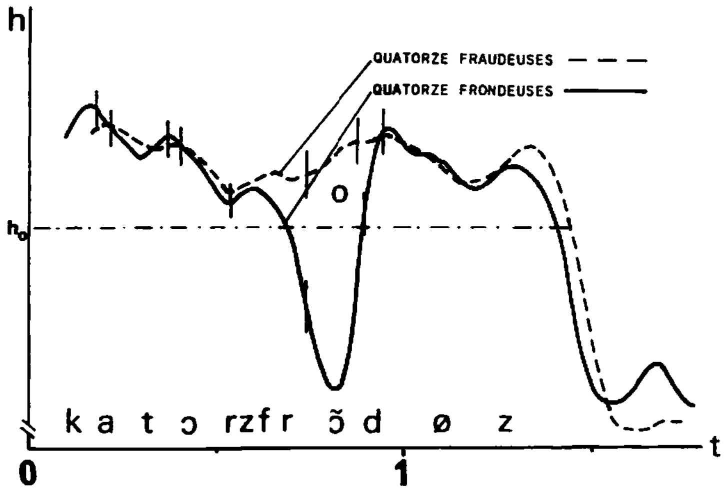

Fig. 6: Velar height for the two phrases "quatorze fraudeuses" and "quatorze frondeuses".

There was not a sufficient number of tokens containing /a/, /aa/ and /aaa/ for statistical testing, but they appeared to vary as consistently and in a similar way to the /e/'s.

The utterances involving only one repetition of the vowel show, although not always clearly, a velar gesture toward a target position for the vowel. The subsequent velar descent for the nasal consonant starts around the beginning of the prenasal vowel.

For the utterances involving /ee/ and /eee/, the velar movement toward the vowel target is very obvious, as can be seen in Fig. 7. The subsequent velar descent for the nasal consonant starts noticeably earlier than for the single vowel case, but no significant difference can be found between double vowel and triple vowel cases.

For the vowel /a/, a similar observation can be made but it is not as obvious: because of the lower intrinsic velar height for /a/, the velar movement pattern is steeper and in general no plateau is visible. This is illustrated in Fig. 8.

In all cases examined, velar lowering ends during (or at the end of) the nasal consonant, which is always followed by an oral vowel. Conversely the nasal consonant always ends when the velum is near its lowest position.

Summary of Results and Discussion

The above results show that for this speaker of French:
1) stop consonants are produced with a closed velopharyngeal port, the velum generally reaching its highest position at the time of oral release;

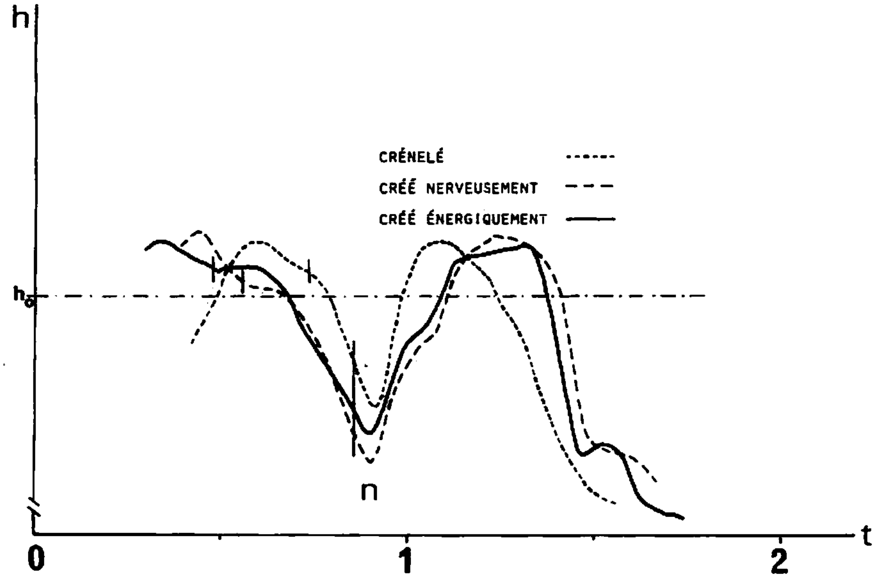

Fig. 7: Velar height for the three phrases "crénelé", "créé nerveusement" and "créé énergiquement". The first segmentation mark indicates the transition /r/ to /e/.

2) fricative consonants are normally produced with a closed velopharyngeal port, the velum generally reaching its highest position during the consonant; occasionally, the velopharyngeal port may start opening before the end of the consonant;
3) nasal consonants are produced with the velopharyngeal port at least partially open; in a few cases however, opening may be delayed (up to 20 ms) with respect to the oral closure;
4) (phonemically) oral vowels are normally (i. e. in the environment of non-nasal sounds) produced with a closed velopharyngeal port; exceptions to this are some instances of open vowels (/a/, /ɑ/) in isolation, and many vowels in final position; when the environment is not non-nasal, the situation is more complex and will be discussed below.
5) (phonemically) nasalized vowels are normally produced with an open velopharyngeal port, subject to contextual constraints; these are most severe in a /CV_nC/ context where C is a stop consonant: the oral release-velum lowering timing at the first transition must be achieved within narrow tolerances and the oral closure-velar closure synchronization at the second transition with even greater precision.

The data suggests that in several contexts, velum rising is faster than lowering, however, it might not be due to mechano-neuro-muscular limitations, but simply to less stringent requirements for opening than for closing. Suitable segment sequences in other languages might support this point.

One problem touched upon but unresolved in this study is: how much velum lowering is tolerable for a phonemically oral vowel to still be perceived as (phonemically) oral? In most relevant cases, the velum position

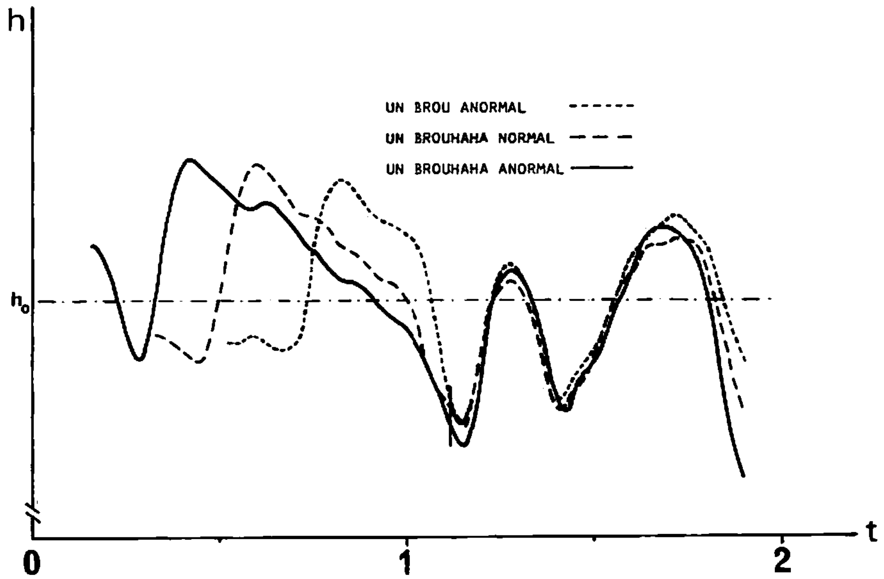

Fig. 8: Velar height for the three phrases "un brou anormal", "un brouhaha normal" and "un brouhaha anormal".

is changing rapidly through the vowel. One might thus wonder whether some kind of time integration is performed by the listener, namely: would it be the average velar passage which matters, or rather the average value of some auditory function depending on the size of this passage? If so, it seems that some quantitative determination of the velar passage cross-section (and most likely that of the oral passage too) will have a better chance of success in solving this problem.

The purpose of the last two groups of utterances under investigation (e) and f)) was to study anticipatory coarticulation effects. Group e) confirms and supplements the results obtained by airflow measurements on several subjects (Benguerel, 1974). This previous study had shown nasal airflow to start at the earliest during the third consonant of the cluster. In the present data, it can be seen that velar lowering, which starts ahead of velopharyngeal port opening, is initiated as early as the second consonant of the cluster. Electromyographic data do indicate an even earlier muscular activity, or suppression of activity, as the case may be for the levator palatini. (Benguerel et al., 1975)

Group f) is inconclusive about testing the look-ahead strategy (Henke, 1966); it would contradict it if it could be assumed that oral vowels are assigned a feature value 0 (rather than + or –) for velum position (cf. Benguerel and Cowan, 1974) since one would expect downward velar movement to be initiated earlier for /–VVVC$_n$–/ sequences than for /–VVC$_n$–/ sequences. As noted earlier, (phonemically) oral vowels can tolerate some amount of velopharyngeal opening (say S_{max}) but it should be obvious that if velum lowering is initiated too early, S_{max} will be exceeded before the onset of the nasal consonant and a (phonemically) nasalized vowel will be produced instead of an oral one. It thus seems that oral vowels must be assigned the feature value + at the model input level, the value of S_{max} remaining to be determined (see above), or that a limit must be put to the maximum "distance" of the look-ahead procedure. This limit would most likely be feature dependent and would thus render the model much less attractive. One way of testing the strategy for velar movement would be to use segments that are known not to involve the velum, e.g. [ʔ], [h], [ɦ] or pharyngeal sounds. Unfortunately none of these exist in French.

It has been suggested that the lowering of the velum in apparent anticipation of the following nasal consonant occurs as soon as it possibly can. Unless one can define independently "how soon it possibly can", the definition and the feature value assignment become circular.

Further Research

Several problems have been suggested in this paper. There is a need for extending this investigation to more subjects, to other languages (in particular those with phonemically nasalized vowels, such as Polish and Portuguese). It appears that inasmuch as it can provide similar information, it is desirable to use a method more adaptable for real-time processing such as the nasograph (Ohala, 1970, 1971) or measurements derived from the aerodynamic parameters.

References

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

Benguerel, A-P. and H. A. Cowan (1974), "Coarticulation of Upper Lip Protrusion in French", Phonetica Vol. 30, 41-55.

Benguerel, A-P., H. Hirose, M. Sawashima and T. Ushijima (1975), "Electromyographic Study of the Velum in French", Ann. Bull. RILP (this issue).

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

Henke, W. (1966), "Dynamic Articulatory Model of Speech Production using Computer Simulation", Ph. D. Dissertation, MIT, Cambridge.

Kelkar, A. (1970), "Hindi-Urdu Phonology", (mentioned in Ohala (1970)).

Moll, K. L. (1962), "Velopharyngeal Closure on Vowels", J. of Speech and Hearing Res. 5, 30-37.

Ohala, J. J. (1970), "And now, a nasograph!", University of California at Berkeley Phonology Laboratory, Monthly Internal Memorandum, December.

Ohala, J. J. (1971), "Monitoring Soft Palate Movements in Speech", University of California at Berkeley Phonology Laboratory, Project on Linguistic Analysis Report No. 13, JO1-JO15.