Aerodynamic, articulatory and acoustic realization of French /ʁ/

Cedric Gendrot[1], Barbara Kuhnert[1], Didier Demolin[1]

Université Sorbonne Nouvelle Paris 3, Laboratoire de Phonétique et Phonologie UMR 7018
cgendrot@univ-paris3.fr, barbara.kuhnert@univ-paris3.fr, ddemolin@univ-paris3.fr

ABSTRACT

The study of realization of French uvular /ʁ/ is usually considered as problematic due to its variability. In this study, an articulatory (EMA) and aerodynamic experiment allowed us to determine its major axes of variation. We show that the degree of constriction between the tongue and the palate is related to the voicing of the consonant and we validate the use of a harmonic-to-noise ratio for the measurement of variation of /ʁ/. Aerodynamic data also show that the variation of Nasal Airflow is not significant but a ratio between Oral Airflow and Intra-Oral Pressure significantly varies according to the different realizations of /ʁ/. Subglottic Pressure only plays a significant role in specific cases. Finally, acoustic analyses of continuous speech based on results found in this study investigate prosodic factors of variation for the production of /ʁ/.

Keywords: /ʁ/, physiology, aerodynamic, acoustic, continuous speech, variation.

1. INTRODUCTION

This study is a preliminary work within the junior researcher financed project SHS2 ANR-13-JSH2, the aim of which is to understand the variability of French standard (uvular) /ʁ/.

Chafcouloff ([1,2]), Fougeron [3], Tranel [15], Walter [16], and Meunier [9] for French have mentioned the variability of /ʁ/, especially in positions such as word initial and word final, thus qualifying them as free variants. Its main axes of variation seem to be the voicing/unvoicing and approximant/frication features, thus leading roughly to four kinds of /ʁ/ : voiced approximant, unvoiced approximant, voiced fricative and unvoiced fricative. This variability seems inherent to the class of rhotics, whatever their place of articulation or the language they appear in [7,8,10].

From a phonological point of view, /ʁ/ might be considered both as a fricative or as an approximant and the specification of its voicing may also be reconsidered [12]. We consider here that the variability of French /ʁ/ is not fully apprehended, since it has been mainly acoustically studied in read speech corpora and sometimes even based on impressionistic judgments. The variability of /ʁ/ might be better explained by its physiological characteristics: if we consider the aerodynamic voicing constraint [11], the fricative or approximant status of /ʁ/ should imply in fine a dichotomy between unvoiced fricative and voiced approximant. Especially since French /ʁ/ is a posterior (uvular) phoneme and that the back cavity of the vocal tract is thus much reduced, oral pressure soon reaches the level of sub-glottal pressure and "puts out" voicing. French /ʁ/ could then be realized as an unvoiced fricative or as a voiced approximant according to the degree of constriction. We hypothesize in this study that that variation of /ʁ/ is mainly explained by its constriction parameter and that it leads to a continuum between unvoiced fricative and voiced approximant; the two others (voiced fricative and unvoiced approximant) being rare realizations due to semantic/prosodic factors mainly, that we will also pinpoint.

A first experiment with Electro-Magnetic Articulograph (EMA) data will show whether the different productions of /ʁ/ that vary in terms of voicing also vary in terms of their tongue-palate constriction. Secondly a an aerodynamic study will investigate the same productions in terms of their resistance airflow (intra-oral pressure- oral airflow ratio) which is related to the degree of constriction, but we will be able to check the influence of subglottic pressure. The aim of these first two experiments is also to provide acoustic measurements that will capture the /ʁ/ variation as much as possible. In a third step we will provide an acoustic study on a large corpora of continuous speech, so as to test the variability of French /ʁ/ in terms of the aforementioned results.

2. METHOD

In this section we develop the procedure used for the three experiments to be presented.

2.1. Experiment 1: EMA study

EMA data were recorded (figure 1) at the IPS-Munich on five French speakers by the second author with the the 3 dimensional EMA technique

(AG500 Carstens Medizinelektronik). Our speakers are all French native speakers aged between 29 and 50 years and with no known history of speech or hearing disorders. It is known that the uvula can come in contact with the back of the tongue to enable friction (figure 2) [1]. Unfortunately no sensor was placed on the uvula. Some other data (not presented here) showed some lowering and fronting of the captor placed on the uvula, although not as important as for nasal phonemes. Some X-ray movies were investigated showing that the uvula is able to lower while maintained the nasal cavity closed as seen on figure 2. In our aerodynamic study, we will be able to show whether there is some nasal airflow during the production of /ʁ/.

Figure 1: Illustration of EMA sensors

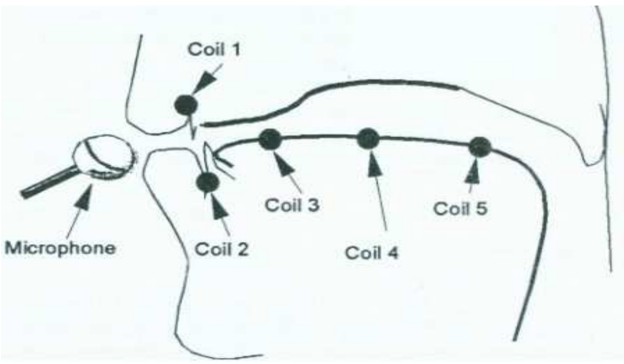
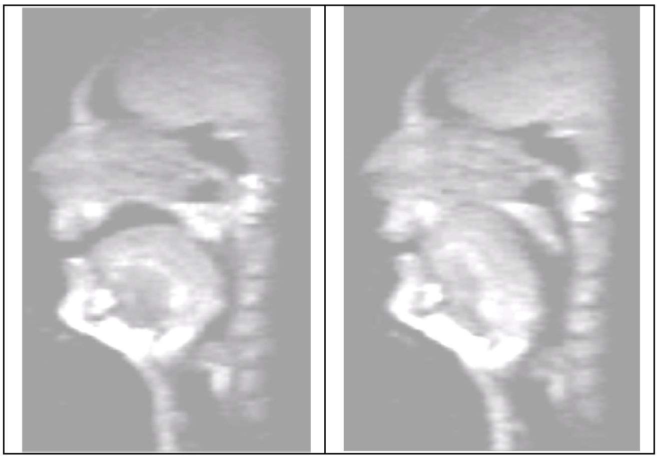


Figure 2: Illustration of the lowering of the uvula without (left) or with (right) opening the nasal cavity for the production of /ʁ/ (left) and /ɲ/ (right)



We will analyze the position of coil 3, 4, and 5 (figure 1) in some specific contexts that are known for their absence of variation as stated in the literature. For example /ʁ/ preceded by an unvoiced obstruent is always unvoiced, while /ʁ/ in an voiced context is mostly (in natural context, not emphasized) realized as voiced [1,3]. We expect the position of the sensor (coil 5 mostly) to be higher for a better constriction in the case of unvoiced context, and lower in the case of voiced context. The horizontal position of the coil is hypothesized to vary mostly with place of articulation but could also be dependent on the constriction parameter.

The corpus consisted in short sentences allowing for all possible clusters in French, repeated 10 times with /ʁ/ in #C/ʁ/V contexts. The vowels were /a/, /i/, and /e/. So as to relate the constriction parameters with acoustic measurements, we calculated in parallel with PRAAT /ʁ/ acoustic duration as well as f0, intensity and Harmonic-to-Noise Ratio (HNR, calculated with default parameters) in the mid acoustic section of the /ʁ/. The HNR measurement [13] represents the degree of acoustic periodicity and is useful to quantify a continuum between voicing and friction: for these two extrema, values can go respectively from 20 to -20 dB.

2.2. Experiment 2: aerodynamic study

Aerodynamic data were recorded from 5 native French speakers by the third author, using the Physiologia workstation for simultaneous acquisition [14]. The systems were linked to a data collection system equipped with different transducers. Oral airflow (OAF) measurements were taken with a small flexible silicon mask placed against the mouth ; nasal airflow (NAF) by two plastic tubes connected to each nostril by a small silicon olive. Intraoral pressure (PIO) was recorded with a small flexible plastic tube (ID 2 mm) inserted through the nasal cavity into the oro-pharynx. Subglottal pressure (P**SG**) was measured with a needle (ID 2 mm) inserted in the trachea, right under the cricoid cartilage.

Only two speakers are currently analyzed. Since a Rothenberg mask is covering the mouth for the airflow acquisition, the acoustic data can't be analyzed with precision other than f0 and duration. A value of resistance airflow was calculated as the ratio between PIO and AOF in order to evaluate the degree of constriction: if the resistance airflow value is high, then it suggests that the constriction in the back of the mouth is important, and vice-versa. The point of maximum during the realization of /ʁ/ was considered in our analyses.

The corpus consisted in a few sentences repeated three times with /ʁ/ in C/ʁ/V contexts, V/ʁ/C contexts, V/ʁ/V contexts, and V/ʁ/# (sentence final) contexts. The vowels were /a/, /i/, and /e/ and the consonants /p,b,t,d,k,g,f,v,s,z/. Although the corpus is not identical to the EMA experiment, similar phonemic contexts are being analyzed. In experiments 1 and 2, ANOVAs are used for statistical analysis.

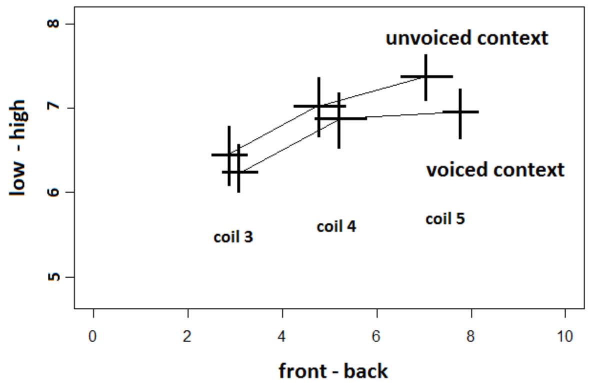
2.3. Experiment 3: acoustic study

The ESTER corpus [4] has been used for the acoustic analysis part: it is a 30 hours broadcast corpus considered as prepared speech rather than read speech, with few sequences of spontaneous speech. Orthographic transcription and rough segmentation have been realized by human transcribers, then phonemic and lexical alignment have been realized automatically by the LIMSI alignment tool [5]. Measurements were automatically realized using PRAAT on acoustic parameters as duration, f0, and HNR measurements of /ʁ/ and compared with its surrounding phonemes. Linear mixed effect models with speaker as a random effect are used for statistical analysis.

3. EMA ANALYSIS

Figure 4 shows that vertical position of coil 5 is significantly higher ($p<0.01$) in unvoiced contexts (unvoiced obstruent preceding /ʁ/) than in voiced contexts (voiced obstruent preceding /ʁ/) for 4 speakers out of 5. This confirms our hypothesis that constriction is more important in unvoiced contexts. As for horizontal position of coil 5, it is more anterior ($p<0.01$) in unvoiced contexts thus revealing a more anterior constriction. Positions of coil 3 and 4 do not vary significantly.

Figure 4: Vertical and horizon position of coil 3, 4, and 5 according to different positions of /ʁ/ (the length of the crosses represents standard errors) for speaker 3.

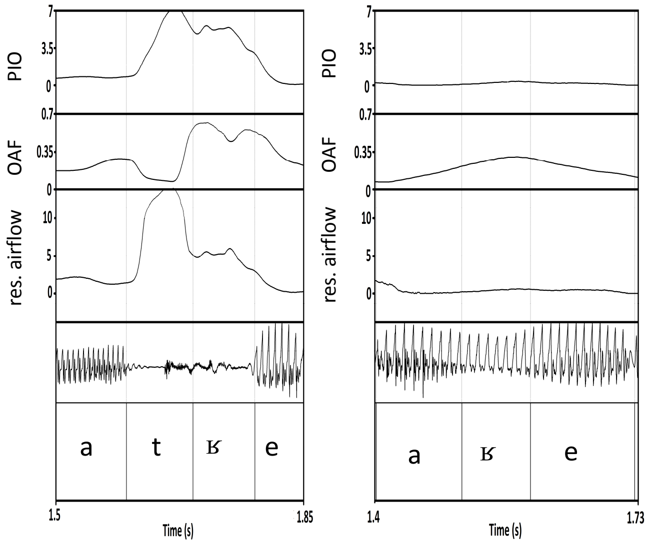


An acoustic analysis of /ʁ/ in the same contexts reveals that /ʁ/ in unvoiced context is significantly longer ($p<0.0001$), with smaller intensity ($p<0.0001$) and a lower HNR ($p<0.0001$) value than /ʁ/ in voiced context.

4. AERODYNAMIC ANALYSIS

The nasal airflow has been detected close to zero (below 0.01 dm^3/s, not higher than other oral phonemes) while nasal phonemes reach between 0.03 dm^3/s and 0.05 dm^3/s.

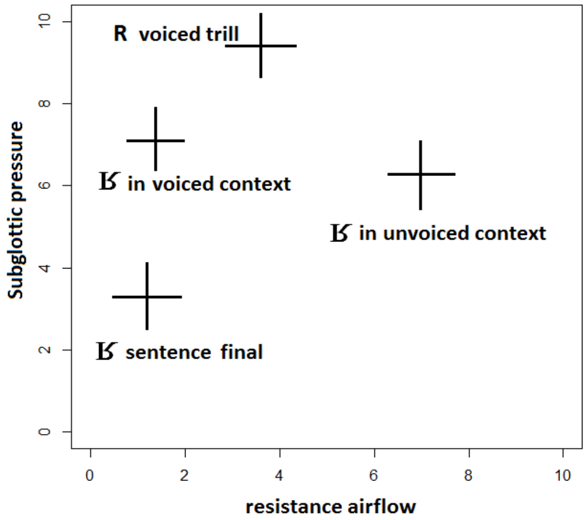
Figure 5: aerodynamic results for speaker 1: PIO and OAF and resistance airflow for two /ʁ/ variants : in unvoiced context (left) and voiced context (right).



Intra-oral pressure was found to be significantly higher ($p<0.0001$) for /ʁ/ in the vicinity of unvoiced obstruents than in the vicinity of vowels and voiced obstruents (cf. figure 5). Oral airflow is not significantly different for /ʁ/ between unvoiced and voiced contexts ($p>0.05$). The resistance airflow as described in the method section logically indicates a higher value for /ʁ/ in the vicinity of unvoiced obstruents while significantly lower for /ʁ/ in the vicinity of vowels and voiced obstruents. Recall that acoustic measurements are made difficult because of the Rothenberg mask. We only measured duration and f0, while HNR were too erratic to be meaningful and we find, identically to the EMA experiment, that /ʁ/s in unvoiced context are characterized by voicing (f0 detection) and a longer duration ($p<0.001$).

One of the speakers voluntarily realized uvular voiced trills in some sentences. These are characterized by a higher subglottal pressure (with striations) and an intermediate resistance airflow (cf. figure 6) leading in a voiced fricative. At the end of the sentences, both speakers produced /ʁ/ with a strongly decreasing subglottal pressure but with a low resistance airflow leading in an unvoiced approximant. When considering subglottal pressure, we notice that it doesn't rise or lower for /ʁ/ except in the case of emphasis (where it rises) or end of sentences (where it lowers). These variations allow to explain the two other /ʁ/ variants mentioned in the literature: respectively voiced fricative and unvoiced approximant as shown in figure 6.

Figure 6: summary of aerodynamic results for speaker 1: PSG vs. resistance airflow for four /ʁ/ variants. (the length of the crosses represents standard errors)

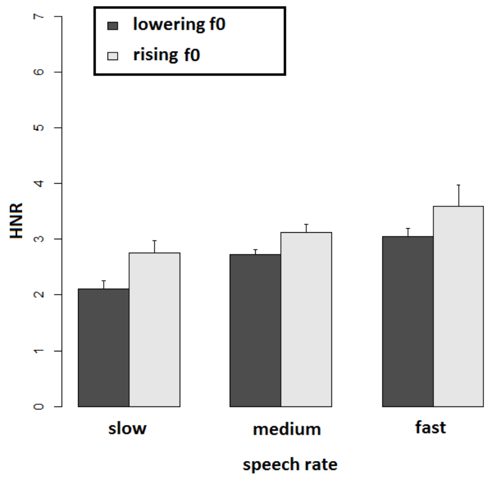


5. ACOUSTIC ANALYSIS

In order to further analyze the results of the previous experiments, we first isolated word-initial positions, and then word-final positions before pauses in a corpus of broadcast speech.

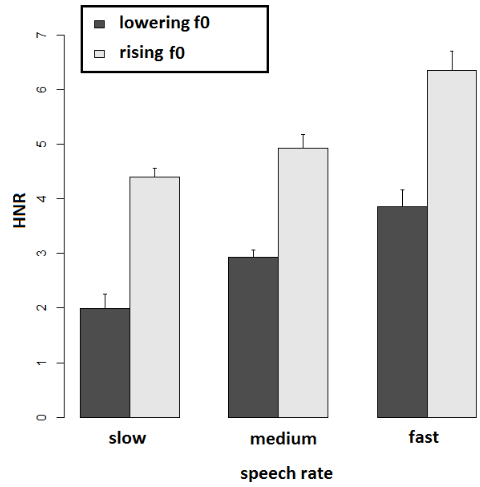
6500 /ʁ/ in word-initial positions (figure 7) analyzed here are only intervocalic. The f0 factor investigated here is the difference between the vowel following and the vowel preceding /ʁ/: we estimate that the presence of a semantic accent can be validated by this increase in f0. We also tested the speech rate calculated as the number of phonemes per second (<11, 11> <16, >16). Figure 7 shows that HNR is higher when the f0 difference is positive ($p<0.05$) and when speech rate is higher ($p<0.05$), thus indicating a more voiced /ʁ/ in these conditions.

Figure 7: HNR as a function of speech rate and f0 difference: word initial /ʁ/ preceded and followed by a vowel.



6300 /ʁ/ in word-final positions (figure 8) analyzed here are preceded by a vowel and followed by a pause [6]. The f0 difference is measured on the preceding vowel only and is aimed at detecting intonational phrase final (when f0 is rising) versus sentence final /ʁ/ (when f0 is falling). Figure 8 shows that HNR is lower when the slope on the preceding vowel is negative ($p<0.01$) and when speech rate is lower ($p<0.05$), thus indicating a less voiced /ʁ/ in these conditions: a results which is confirms the results of experiment 2.

Figure 8: HNR as a function of speech rate and preceding f0 contour: word final /ʁ/ preceded by a vowel and followed by a pause



6. DISCUSSION AND CONCLUSION

In this study we showed that the realization of unvoiced /ʁ/ is characterized by a higher constriction of the back of the tongue (then realized as a unvoiced fricative) as opposed to a voiced /ʁ/ (then realized as a voiced approximant). These realizations are accompanied by longer duration and higher intensity and a lower harmonic-to-noise ratio.

The aerodynamic experiment allowed us to confirm the EMA data but also to investigate other possible realizations, i.e. approximants realized with low subglottal pressure and fricatives with high subglottal pressure. The former are realized as unvoiced approximants while the latter are realized as voiced trills.

Finally the acoustic study enabled us to show two unknown sources of variation for the production of /ʁ/ in French: speech rate and f0.

If we get back to the phonological characterization of /ʁ/, we may conclude from these results that /ʁ/ is to be considered as a fricative when realized in its full form (longer with a complete articulatory gesture), and its approximant realization is seemingly a shorter/reduced form.

7. REFERENCES

[1] Chafcouloff, M. 1980. Les caractéristiques acoustiques de [j, |, w, l, r] en français. In *Travaux de l'Institut de Phonétique d'Aix*, 1-52.

[2] Chafcouloff, M. 1983. A propos des indices de distinction /1-R/ en français. *Speech Communication.* 2:137–139.

[3] Fougeron, C. 2007. Word boundaries and contrast neutralization in the case of enchaînement dans French. In Cole, J. & Hualde J.I. (eds) *Papers in Laboratory Phonology IX: Change in Phonology*, Berlin: Mouton de Gruyter, 609-642

[4] Galliano S., Geoffrois E., Mostefa D., Choukri K., Bonastre J.-P., Gravier G. 2005. The ESTER Phase II evaluation campaign for the rich transcription of French broadcast news, *Proceedings of Eurospeech-Interspeech*, Lisbonne, 1149-1152.

[5] Gauvain, J.L., Lamel, L. and Adda, G. 2002. The Limsi Broadcast News Transcription System, *Speech Communication*, 37(1-2): 89-108.

[6] Gendrot, C. and Gerdes, K. 2011. "Prosodic hierarchy and spectral realization of vowels in French", In : Yoo, H-Y & Delais-Roussarie, E. (eds), Actes d'Interfaces Discours Prosodie, http://makino.linguist.jussieu.fr/idp09/docs/IDP_actes/Articles/gendrot.pdf, 191-205.

[7] Lawson, E., Stuart-Smith, J., Scobbie, J. Articulatory Insights into Language Variation and Change: Preliminary Findings from an Ultrasound Study of Derhoticization in Scottish English. *University of Pennsylvania Working Papers in Linguistics.* 14(2), 2008.

[8] Maddieson, I. 1984. *Patterns of Sounds*. Cambridge: Cambridge University Press.

[9] Meunier, C. 1994. Les groupes de consonnes : problématique de la segmentation et variabilité acoustique. Thèse de Doctorat. Université d'Aix-en-Provence.

[10] Mielke, J, A. Baker, and D. Archangeli. 2010. Variability and homogeneity in American English allophony and retraction. Variation, Detail, and Representation. *LabPhon 10*. Berlin: Mouton de Gruyter. 699-719,.

[11] Ohala, J. J. 1983. The origin of sound patterns in vocal tract constraints. In: MacNeilage, Peter (ed.): *The production of speech*. New York: 189-216.

[12] Russell Webb, E. 2004. Explanatory adequacy in phonology: a deductive approach to /R/. *Linguistica Atlantica*, 25: 77–94.

[13] Severin, F., Bozkurt, B., Dutoit, T. 2005. HNR extraction in voiced speech, oriented towards voice quality analysis. Proceedings of EUSIPCO'05, Antalya,Turkey.

[14] Teston, B. et Galindo, B., (1995), "A diagnostic and rehabilitation aid workstation for speech and voice pathologies", Proceedings of Eurospeech 4, Madrid, 1883-1886.

[15] Tranel, B. 1987. *The Sounds of French: An Introduction*. Cambridge: Cambridge University Press.

[16] Walter, H. 1988. *Le français dans tous les sens*. Paris : Robert Laffont.

8. ACKNOWLEDGMENTS

This work is funded by the ANR young researchers program : SHS2 ANR-13-JSH2-0005 (REPER).
It is also supported by the French Investissements d'Avenir - Labex EFL program (ANR-10-LABX-0083)