How do French Cʁ# cluster realizations vary across speech styles?

Yaru WU[1,2], Martine ADDA-DECKER[2], Cécile FOUGERON[2], Cédric GENDROT[2]
[1]*Modèles, Dynamiques, Corpus (MoDyCo), UMR 7114, CNRS,*
[2]*Laboratoire de Phonétique et Phonologie (UMR7018, CNRS-Sorbonne Nouvelle), France*

The realization of consonant clusters tends to vary under the influence of the Sonority Sequencing Principle (SSP, [1][2][3]). French Cʁ# clusters are known to have several variants in order to avoid the violation of the SSP. Nonetheless, the distribution of the variants of Cʁ# clusters in spontaneous speech and the different strategies speakers apply across speech styles with regard to SSP violation are not fully understood. In this study, we will be concerned with obstruent + ʁ clusters, as in the word "quatre" (/katʁ/, *four*), which are the only Cʁ# clusters allowed in French. We investigate the realization of Oʁ# clusters when immediately followed by a word starting with a consonant (i.e. Oʁ#C). These 3-consonant sequences are challenging with respect to SSP (the SSP violation is resolved when the Oʁ# cluster is followed by a vowel). Besides the canonical realization (e.g. [katʁ]), speakers may drop the post-consonantal /ʁ/ (eg. [kat]), insert a schwa (eg. [katʁə]), or even delete /ʁ/ and insert schwa (eg. [katə]). The aim of this study is to better understand the distribution of these forms in continuous speech and how this distribution varies across different speech styles.

Three large French speech corpora were used for our investigations: ESTER (≈100 hours) [4], ETAPE (≈50 hours) [5] and the Nijmegen Corpus of Casual French (NCCFr, ≈40 hours) [6]. Both ESTER and ETAPE are broadcast speech. The ESTER corpus mainly contains formal journalistic news speech, whereas the ETAPE corpus is mostly composed of conversational journalistic speech, including conversations and debates. The NCCFr corpus contains casual conversations between friends.

Speech files were automatically aligned with the help of the LIMSI alignment system. For our study, we introduced specific pronunciation variants (i.e. /ʁ/ and schwa variants) allowing for optional /ʁ/ and schwa in all words ending in Oʁ# [7][8]. The aligned variant is considered reflecting the realized pronunciation. A subset of the automatic alignment (≈30h) is accompanied by a manual alignment carried out by an experienced French phonetician. Cohen's kappa [9] reveals that the two types of alignment have "almost perfect agreement" (kappa > 0.8).

The absence/presence of /ʁ/ and of schwa were determined by comparing the aligned pronunciations (reflecting the speakers' pronunciations) with the canonical pronunciation containing the 2 elements of the clusters without a schwa (eg. /katʁ/), as specified in Lexique380 [10]. A generalized linear mixed model (GLMM) [11] and follow-up post-hoc analyses were carried out to validate our analyses on the effect of speech style.

Figure 1 shows the overall results where variant production rates are pooled using the 3 corpora. The distribution of the variants of Oʁ# per speech style is shown in Figure 2. Overall, speakers tend to either insert schwa (47%) or delete /ʁ/ (33%) when producing the Oʁ#C sequences. Speakers almost never (≈0%) delete /ʁ/ and insert schwa at the same time. Only 20% of the tokens are aligned using the canonical pronunciation ([Oʁ#]), which correspond to 3-consonant sequences violating SSP ([OʁC]). The examination of these cases shows interesting results related to the type of the following consonant: more than 30% of the [Oʁ] pronunciation is followed by [l] (eg. quatre livre [katʁ#livʁ]). As illustrated in Figure 2, the distribution of the variants depends on speech style ($p < 0.001$). /ʁ/ deletion ([kat]) tends to be much more frequent than schwa insertion ([katʁə]) in the casual speech corpus NCCFr, whereas the opposite trend is found for the formal journalistic speech corpus ESTER. The less formal the speech style, the less we observe schwa insertion and the more we observe /ʁ/ deletion.

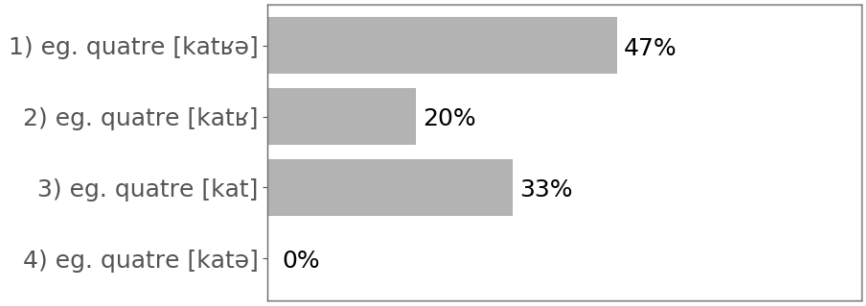

Figure 1. Distribution of the variants of the /Oʁ#/ cluster when followed by /#C/ over all corpora 1) eg. quatre [katʁə], 2) quatre [katʁ], 3) quatre [kat] 4) quatre [katə] indicate whether the word is realized 1) with epenthetic schwa, 2) canonically, 3) without word-final /ʁ/ and 4) without word-final /ʁ/ and with epenthetic schwa.

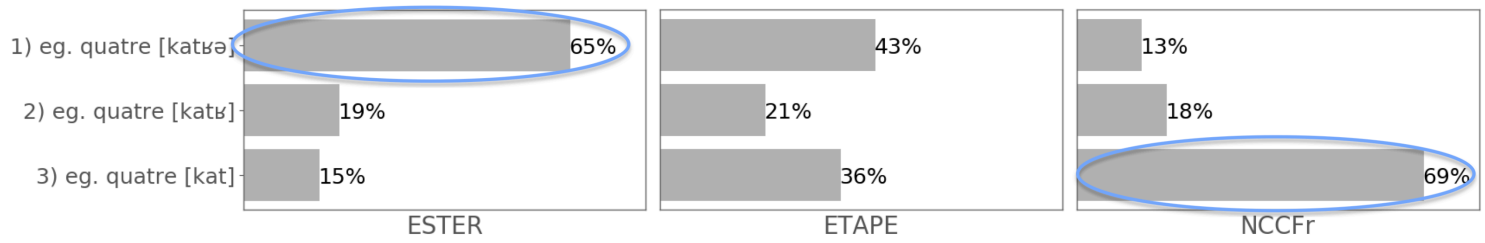

Figure 2. Distribution of the variants of /Oʁ#/ cluster when followed by /#C/ for each speech style.

[1] Clements, G. N. (1988). The sonority cycle and syllable organization. *Phonologica*, 63-76.
[2] Clements, G. N. (1990). The role of the sonority cycle in core syllabification. Papers in *laboratory phonology*, 1, 283-333.
[3] Côté, M. H. (2004). Consonant cluster simplification in Québec French. *Probus*, 16(2), 151-201.
[3] Laks, B. (1977). Contribution empirique à l'analyse socio-différentielle de la chute de /r/ dans les groupes consonantiques finals. *Langue française*, (34), 109-125.
[4] Galliano, S., Geoffrois, E., Gravier, G., Bonastre, J. F., Mostefa, D., & Choukri, K. (2006). Corpus description of the ESTER evaluation campaign for the rich transcription of French broadcast news. In *Proceedings of LREC* (Vol. 6, pp. 315-320).
[5] Gravier, G., Adda, G., Paulson, N., Carré, M., Giraudel, A., & Galibert, O. (2012). The ETAPE corpus for the evaluation of speech-based TV content processing in the French language. In *LREC-Eighth international conference on Language Resources and Evaluation*.
[6] Torreira, F., Adda-Decker, M., & Ernestus, M. (2010). The Nijmegen corpus of casual French. Speech Communication, 52(3), 201-212.
[7] Gauvain, J. L., Lamel, L., & Adda, G. (2002). The LIMSI broadcast news transcription system. *Speech communication*, 37(1), 89-108.
[8] Adda-Decker M., Lamel L. (2000). The use of lexica in automatic speech recognition. In *Lexicon Development for Speech and Language Processing*, F. van Eynde & D. Gibbon (Eds.), Kluwer Academic Publishers.
[9] Gamer, M., Lemon, J., Fellows, I., & Singh, P. (2012). irr: Various coefficients of interrater reliability and agreement. R package version 0.84, 137.
[10] New, B., Brysbaert, M., Veronis, J., & Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied psycholinguistics*, 28(4), 661-677.
[11] Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. arXiv preprint arXiv:1406.5823.