

# Initial Sibilant-Obstruent Clusters in Romance: Explanations for Increasing and Decreasing Phonotactic Complexity

Elissa Pustka

How can we explain that there are initial sibilant-obstruent clusters in some languages but not in others? In Romance, for instance, you can find the Italian *scuola* 'school' (< Latin *schola(m)*) and the corresponding adjective *scolastico* with initial [sk], French *école* (< Vulgar Latin *ischola*) without, but *scolaire* with [sk] and Spanish *escuela* and *escolar*, both without initial cluster. The historical evolution from Latin to the Romance languages shows that these variations and changes on the one hand originate from language contact (amongst others with Latin itself via etymological spellings) and on the other hand, from internal processes of vowel reduction which are apparently not stopped by phonotactic constraints.

The paper analyses these dynamics in phonotactic complexity with a detailed analysis of French and an outlook on other Romance languages. The empirical basis is a lexicographical study of the digital dictionary *Petit Robert* for the diachronic perspective, and for the synchronic perspective a corpus study of two sub-corpora of the international research programme (*Inter-Phonologie du Français Contemporain* (I)PFC (cf. Detey et al. 2016). The research is guided by the following questions: Which words allow initial sibilant-obstruent clusters and which do not, and how can this lexical variation be explained by semantics, socio-stylistics and frequency in language use?

The *Petit Robert* with 300,000 entries contains only 163 words beginning with a sibilant-obstruent cluster. Of these words, all excepted one (162) are loans: learned loan words from Latin and Greek such as *structure* (first record in 1528) and *scolaire* (1807) as well as loan words from Germanic languages like English and German, e.g. *spray* (1884) and *strudel* (20<sup>th</sup> century). The only language-internal innovation is *scrogneugneu* (1884), a euphemistic deformation of the interjection *Sacré nom de Dieu!* 'Holy name of God!'. The involved segments are mainly [stʁ] (96 occurrences) and [skʁ] (33 occurrences), alongside [spʁ], [spj] and [skj]. In contrast, clusters with an initial [ʃ] are quite rare: they appear in only two very uncommon lexemes with [ʃpʁ] (*schproum* and *sprechgesang*) and a single one with [ʃtʁ] (*strudel* from Austrian German).

The in-depth analysis of manuscripts, however, reveals a more complex evolution: the first words with initial sibilant-obstruent clusters appear in the 12<sup>th</sup> century, e.g. *splendeur* 'splendour' in 1120 (cf. Samson 2010). In former loans from Germanic like *éperon* 'spurs' (< Frankish \**sporo*) and in inherited vocabulary like French *école* (< Vulgar Latin *ischola* vs. Classic Latin *SCHOLAM*), the cluster has still been broken by a prosthetic vowel and then lost its *s impurum*. Thus, the contemporary spelling hides that the French words with initial sibilant-obstruent clusters were subject to variation. In fact, Medieval manuscripts show variation between the spelling <sC> and <esC> until the 12<sup>th</sup> century. It is assumed that the phonic realization depended on the semantic field of the word and the social class of the speaker: complex syllable structures seem to have appeared since the 14<sup>th</sup> century, mainly in the speech of educated speakers in erudite contexts whereas the prosthesis was common amongst non-educated speakers until the 19<sup>th</sup> century, when the beginning of compulsory schooling and reading and writing capacities can be dated (cf. Rheinfelder 1937, Sampson 2010).

The comparison with other Romance languages also shows considerable variation in increase and loss of phonotactic complexity. Italian continuously allows initial sibilant-obstruent clusters, apart from some variations like *iscuola* alongside *scuola*. The clusters even appear in new word formations like *strarico* 'superrich' with vowel elision from *extra-* (cf. Sampson 2010, Heinz 2010). In Spanish, however, the process of epenthesis is still productive, e.g. in proper names such as *Estrasburgo* for *Strasbourg*. At the same time, we find vowel reduction processes in the highlands of Latin America which lead to

complex syllable structures, e.g. *estrella* [s.treʒa] (cf. de Crignis 2016). Portuguese presents a similar case: The European standard and the Brazilian varieties present forms with an epenthetic vowel (e.g. *escola* EP [iʃ.k'ɔ.lɐ]/BP [is.k'ɔ.lə], *escolar* EP [iʃ.ku.l'ar]/BP [is.ko.l'ar]). In the informal speech of Lisbon, in contrast, the graphic vowel is not yet realized phonically: *escola* [ʃk'ɔ.lɐ] and *escolar* [ʃku.l'ar] (cf. Ashby et al. 2012).

The analysis of contemporary French corpora shows how these variations and changes emerge in language use. This is particularly interesting in the case of the [ʃ] obstruent-liquid clusters which are quasi inexistent in the French lexicon (only three marginal types; see above). The PFC-sub-corpora of Aveyron and Paris (13h20 of spontaneous speech from 40 speakers; cf. Pustka 2007) document how these clusters appear as a result of schwa elision and assimilation of voice: A very common context for this are verbal phrases with the first person singular pronoun *je* and some frequent verbs with grammatical or pragmatic functions as *croire* 'to believe' or *trouver* 'to find', e.g. *j(e) crois* [ʃkʁwa], *j(e) trouve* [ʃtʁuv]. In Parisian French, the emergence of these forms can be traced back to a natural weakening process (that has originated at least at the beginning of the 20<sup>th</sup> century; cf. Grammont 1914), in Southern French, we interpret these constructions that we find mainly in the speech of the younger generation as loans (cf. Pustka 2007). This data shows once again the importance of language contact for the emergence of new phonotactic patterns. Still, the question remains if the innovation has been a purely internal process (where segmental reduction has been possible despite of syllable structure constraints) or if it has only been possible under the prerequisite of linguistic contact.

## References

- Ashby, Simone et al. (2012): "A Rule Based Pronunciation Generator and Regional Accent Databank for Portuguese", in: *Proceedings of Interspeech 2012*, 1886-1887.
- Detey, Sylvain et al. (eds.): *Varieties of Spoken French*, Oxford: OUP.
- Grammont, Maurice (1914): *Traité pratique de prononciation française*, Paris: Delagrave.
- Heinz, Matthias (2010): *Zentrum und Peripherie in romanischen Silbentypsystemen. Studien zur Materialität, Typologie und Diachronie komplexer Silbenstrukturen in romanischen Sprachen unter besonderer Berücksichtigung der Lehnwortphonologie* (unpublished habilitation thesis, Tübingen).
- Rheinfelder, Hans (1937): *Altfranzösische Grammatik*, vol. 1: *Lautlehre*, München: Hueber.
- Robert, Paul/Rey-Debove (2008): *Le nouveau Petit Robert: dictionnaire alphabétique et analogique de la langue française (version électronique)*, Paris: Dictionnaires Le Robert.
- Sampson, Rodney (2010): *Vowel Prosthesis in Romance. A Diachronic Study*, Oxford: O.U.P.