A Multifactorial Analysis of French Liaison
Mathieu Avanzi1, Sandrine Brognaux2,3
1DTAL, Cambridge University; 2CENTAL - ICTEAM, Université catholique de Louvain
3TCTS, UMONs
mathieu.avanzi@unine.ch; sandrine.brognaux@uclouvain.be

Liaison is the phenomenon whereby a latent final consonant in a word (Word-1) is pronounced as the onset of a following vowel-initial word (Word-2). Since Delattre [1947], liaison is usually described as obligatory, optional or prohibited. Research has shown that the realization of liaison varies according to a certain number of cues [Encrevé 1988; De Jong 1988; Morin & Kaye 1982; Durand et al. 2011], among which: segmental variables (liaison being less common with [k] and [v] than with [z] and [t]), the morphological and grammatical status of the two connected words (the liaison being obligatory between a function Word-1 and a content Word-2, while optional in clusters where two lexical items are joined) and their frequency (words of more frequent usage favor liaison in a greater extent), syntactic and prosodic variables (liaison being realized more often within a Phonological Phrase rather than between two Phonological Phrases, henceforth PP), realizational variables (a slow speaking rate favoring the presence of liaison, as would shorter words), stylistic variables (read speech favors liaison compared with spontaneous speech) and speaker-related variables (such as sex, age, socio-educational level and geographical origin).

A lot of studies have considered some of these predictors simultaneously to account for liaison in French (see Soum-Favaro et al. [2014]), but, to our knowledge, the way prosodic, realizational and speaker-related variables interact has been especially understudied. The aim of this study is to shed new light on the role played by these three classes of factors on the realization of liaison in French through the use of a large manually-annotated corpus. To this aim, we focus on the clusters where the realization of potential liaison has been claimed to manifest great variation, and which can be analyzed as a single PP or two PP, i.e. in clusters including a Noun, followed or preceded by and Adjective (un mauvais(_z_)enfant, des pâtes(_z_)italiennes), an Adjective preceded by an Adverb (très(_z_)inquiet) and a Verb followed by a NP or a PrepP (est(_t_)enceinte, c’était(_z_)en avril). Our study has the advantage of relying on a 13 hour-long corpus extracted from the c-prom-pfc database [Avanzi 2014]. It was phonetically aligned and labeled at different levels (phones, syllables and words, PP, PoS, disfluencies) in Praat [Boersma & Weenink 2014] and designed to study prosodic variation among speaking style (read vs conversational speech) and dialectal variation. The database contains 120 speakers of different age and matched for sex, equally originated from 15 locales in France, Belgium and Switzerland [Durand et al. 2009]. In total, more than 2000 sites of potential liaisons were automatically extracted from the corpus.

Data were analyzed by means of (Generalized) Linear Models [Baayen et al. 2004], with the realization of the liaison (1/0) as a dependent variable, and more than 15 predictors as independent variables, such as speaker’s age, geographical origin, sex, socio-educational level, speaking style, PoS structure of the cluster hosting the liaison, number of the words (singular/plural), syllabic length of the words, nature of the liaison consonant, position of Word-2 within the host PP, position of the Adjective or the Adverb (before or after Word-1), articulation rate and relative word frequency in the corpus. Results show that the geographical origin and sex have no impact on the realization of liaison, neither has the inclusion of the second word within the same PP as Word-1. On the other hand, factors such as socio-educational level, speaker’s age, speaking style, nature of the cluster hosting the liaison and PoS of Word-1, articulation rate, nature of the consonant of liaison and some of their interactions do exert a significant role on liaison realization.
References


