Deriving the gradient behavior of French liaison through constraint interaction

Benjamin Storme

Université de Lausanne

Going Romance 2020

The puzzle of French liaison

- In French, some words have a consonant-final variant that only occurs before vowel-initial words (e.g. *grand* [gʁã(t)] 'tall-masc').
- Liaison consonants are challenging for phonological theory because of evidence that they pattern ambiguously between word-final and word-initial consonants (see Côté, 2011 for an overview).

(1) /ti/-affrication in Quebec French (Côté, 2014)

		Example	% affrication
(a)	Word-final consonant	*tren*/t/ *innocents* 'thirty innocent-masc.plur'	36.5%
(b)	Liaison consonant	*gran*/t/ *innocent* 'tall-masc innocent-masc'	66.0%
(c)	Word-initial consonant	/t/*imide* 'shy'	99.2%

- **Research question:** How to account for the ambiguous/gradient behavior of French liaison consonants?

A representational analysis

- In recent works, the gradient behavior of liaison consonants has motivated different *underlying* representations for liaison consonants and non-liaison consonants (e.g. Smolensky and Goldrick, 2016).
- Phonological representations for segments are enriched with specific activity levels that determine how likely segments are to surface.
  - Liaison consonants have a smaller activity level than stable consonants, explaining why they do not always surface.
- Lexical representations of all vowel-initial words (e.g. *ami*) are assumed to contain the different segments available as liaison consonants in French word-initially (e.g. *ami* is represented underlyingly as /{t, n, z}ami/).
- E.g. underlying representation for *petit ami*: /pəti($t_{0.48}$)/ + /{$t_{0.09}$, $z_{0.09}$, $n_{0.09}$}ami/

An alternative analysis based on constraint interaction

- **Goal:** proposing an alternative analysis where the gradient behavior of liaison derives from *constraint interaction* operating on classical phonological and lexical representations.
- **Insight:** Gradient effects in word pronunciation may emerge from interactions with independent paradigmatic properties of words.
- E.g. incomplete voicing neutralization in German *Rad* [ʁaːt] 'wheel' (Winter and Roettger, 2011) as a compromise between
  - word-final voiceless allophone (enforced by phonotactic constraint) and
  - prevocalic voiced allophone (enforced by paradigmatic-uniformity constraint)

Rat [ʁaːt]	$\xrightarrow[\text{shortening}]{\text{phonological}}$	Rad [ʁaːt]	$\xleftarrow[\text{lengthening}]{\text{analogical}}$	Räder [ʁæːdɐ]
'advice'		'wheel'		'wheels'

An alternative analysis based on constraint interaction

- **Proposal:** the ambiguous patterning of liaison consonants in word 1-word 2 sequences is due to the influence of word 1's and word 2's *citation forms*, i.e. words as pronounced in isolation.
  - Uniformity with the citation form of *word 1* favors attachment to *word 2*.
  - Uniformity with the citation form of *word 2* favors attachment to *word 1*.
  - ⇒ These conflicting uniformity requirements of word 1 and word 2 in liaison contexts result in the intermediary status for liaison consonants observed in the literature.

(2) Consonant type and uniformity with citation forms

		Example	Citation forms
(a)	Word-final consonant	*trente innocents*	tʁãt inosã
(b)	Liaison consonant	*grand innocent*	gʁã t inosã
(c)	Word-initial consonant	*grand timide*	gʁã timid

- In Storme (2020), I showed how this approach can derive the intermediary rate of affrication for liaison consonants in Quebec French.

Goals for today

- Providing additional evidence for the gradient behavior of liaison consonants, through an experimental study of *liaison enchaînée* et *non-enchaînée* in Swiss French.
  - E.g. in *petit ami*, liaison [t] is *non-enchaîné* if [t] is prosodically treated as word-final and *enchaînée* if it is treated as word-initial (Plénat, 2008).
- Showing that this gradient behavior can be derived through constraint interaction, using independently motivated output-output faithfulness constraints (i.e. output-variant (OV) faithfulness; Kawahara, 2002).
  - Connected-speech word variants must be similar to the corresponding citation form.

Methods

Stimuli

- 48 Adjective-Noun sequences were selected, in four different conditions:

	Example
Final consonant	*magnifi*[k] *hôtel*
Liaison 1 consonant	*gran*[t] *hommage*
Liaison 2 consonant	*be*[l] *appartement*
Initial consonant	*joli* [s]*ourire*

- Liaison 1 and Liaison 2 adjectives differ with respect to their similarity with the masculine's citation form:
  - Liaison 1 adjectives can be analyzed as citation form of masculine adjective + epenthetic consonant (e.g. *grand* [gʁã+t]).
  - Liaison 2 adjectives uses the feminine form in liaison contexts and cannot be analyzed as masculine + epenthetic consonant (e.g. *bel* [bɛl]).

Prediction

- The paradigmatic-uniformity analysis predicts different behaviors for Liaison 1 and Liaison 2, with Liaison 2 consonants behaving like final consonants.

Stimuli

- There are six adjectives for each condition.

	Adjectives
Final consonant	*énorme, jeune, large, magnifique, meilleur, superbe*
Liaison 1 consonant	*faux, grand, gros, mauvais, parfait, petit*
Liaison 2 consonant	*ancien, bel, bon, prochain, vieil*
Initial consonant (V-final Adj)	*affreux, charmant, gentil, joli, long, vrai*

- Each adjective is represented in two Noun-Adj sequences, varying in the strength of the collocation, e.g.:
  - *petit ami* (more frequent collocation; $\log P(N|Adj) = -1.24$)
  - *petit anneau* (less frequent collocation; $\log P(N|Adj) = -3.94$)

Note on Adj-Noun collocation

- Liaison consonant is more likely to be pronounced as $P(N|Adj)$ increases (Kilbourn-Ceron, 2017).
- The corpus of subtitles OpenSubtitles was used to compute these frequencies (thanks to Aris Xanthos for his help with this part of the project).

Stimuli

- To probe the word-initial/word-final status of consonant, each Adj-Noun sequences was pronounced by a French native speaker (myself) with a hesitation (*euh* [ø]) occurring between Adj and Noun.

	Prononciation	
	Attachment to word 1	Attachment to word 2
Final C	*magnifi*[k] *euh hôtel*	*magnifi euh* [k]*hôtel*
Liaison 1 C	*grand*[t] *euh hommage*	*grand euh* [t]*hommage*
Liaison 2 C	*be*[l] *euh appartement*	*be euh* [l]*appartement*
Initial C	*joli*[s] *euh ourire*	*joli euh* [s]*ourire*

Table: Experimental items

- In the liaison condition, attachment to word 1 corresponds to a liaison *non-enchaînée* and attachment to word 2 corresponds to a *liaison enchaînée* (Plénat, 2008).

Task

- 23 Swiss French participants took part in an online study, using the LimeSurvey platform (LimeSurvey, 2012).
- For each Adj-Noun sequence, participants were presented with the two pronunciations one after the other with a 1 sec inter-stimulus interval.
- All sequences were presented in pseudo-randomized order.
- Participants had to choose which of the two pronunciations they prefer (which one sounds more natural to them).

Predictions

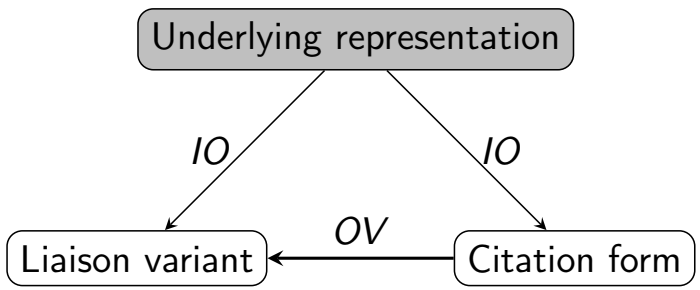
- Final C and Liaison 2 should strongly favor a word-final attachment.
- Liaison 1 C should show an intermediary rate of attachment to Adj and N.
- Initial C should strongly favor a word-initial attachment.

Statistical analysis: Bayesian logistic regression

- The data were analyzed using a Bayesian hierarchical logistic regression.
- The model included Consonant (with 4 levels: initial, liaison 1, liaison 2, final) as a fixed effect.
- It also included the maximal random-effects structure justified by the study's design (Barr et al., 2013):
  - a by-participant random intercept,
  - a by-participant random slope for Consonant,
  - and a by-sequence random intercept.

Grammatical analysis: MaxEnt fit via Bayesian logistic regression

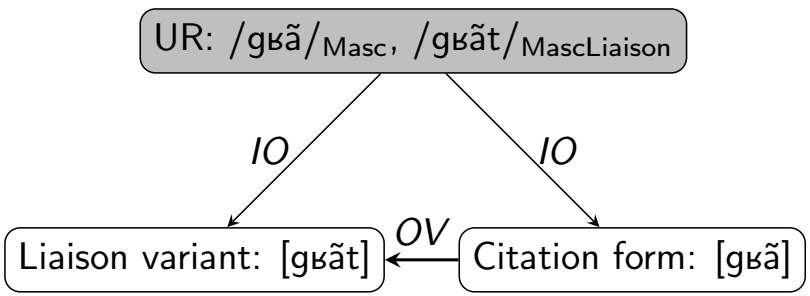
- The statistical analysis was supplemented with a MaxEnt constraint-based analysis (Hayes and Wilson, 2008).
- The full analysis would include both IO faithulness and OV (output-variant) faithfulness constraints (Kawahara, 2002; Myers and Padgett, 2014; Steriade, 1997).
  - Through OV faithfulness, connected-speech variants are penalized if they are not similar enough to the corresponding citation form.

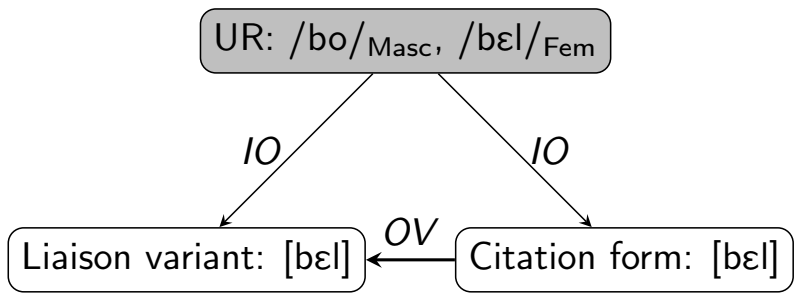


- The present analysis is simplified and focuses on OV faithfulness, with 4 constraints:
  - The constraints penalize consonant epenthesis (Dep) and consonant deletion (Max).
  - The constraints are relativized to a morphological domain (Adj, Noun) to allow variable preference for attachment to Adj or N.

Grammatical analysis: MaxEnt fit via Bayesian logistic regression

- How Liaison 1 and Liaison 2 are distinguished in a liaison context:
  - Liaison 1 is based on the masculine with an epenthetic consonant.
  - Liaison 2 is suppletive and uses the feminine form in liaison context.





Grammatical analysis: MaxEnt fit via Bayesian logistic regression

- Final C

[maʃifik], [otɛl]	Dep_{OV}(C)/Adj	Max_{OV}(C)/Adj	Dep_{OV}(C)/N	Max_{OV}(C)/N
[maʃifik#ø#otɛl]	0	0	0	0
[maʃifi#ø#kotɛl]	0	1	1	0

- Liaison 2 (citation form = corresponding feminine form)

[bɛl], [apaʁtmã]	Dep_{OV}(C)/Adj	Max_{OV}(C)/Adj	Dep_{OV}(C)/N	Max_{OV}(C)/N
[bɛl#ø#apaʁtmã]	0	0	0	0
[bɛ#ø#lapaʁtmã]	0	1	1	0

- Liaison 1 (citation form = masculine form without epenthetic consonant)

[gʁã], [ɔmaʒ]	Dep_{OV}(C)/Adj	Max_{OV}(C)/Adj	Dep_{OV}(C)/N	Max_{OV}(C)/N
[gʁãt#ø#ɔmaʒ]	1	0	0	0
[gʁã#ø#tɔmaʒ]	0	0	1	0

- Initial C

[ʒoli], [suʁiʁ]	Dep_{OV}(C)/Adj	Max_{OV}(C)/Adj	Dep_{OV}(C)/N	Max_{OV}(C)/N
[ʒolis#ø#uʁiʁ]	1	0	0	1
[ʒoli#ø#suʁiʁ]	0	0	0	0

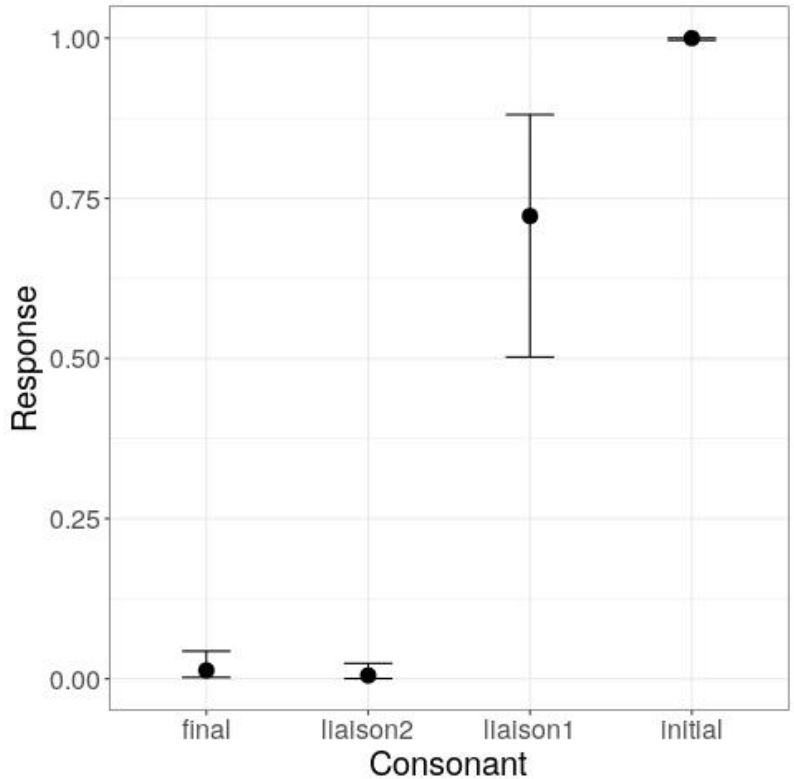
Note on the statistical analyses

- A Bayesian approach was adopted (rather than a frequentist approach) for inferring the parameters of both analyses:
  - Statistical analysis: Bayesian hierarchical logistic regression using `brms` (Bürkner, 2017)
  - Grammatical analysis: Bayesian logistic regression using `rjags` (Plummer, 2016)
- This choice was motivated by the fact that Bayesian inference yields outcomes that are intuitive and easy to interpret (i.e. $P(H|D)$).
- Also, Bayesian approaches virtually always converge to accurate values of the parameters (Liddell and Kruschke, 2018), making it easier to control for individual-specific or item-specific effects.

Results

Logistic regression: results

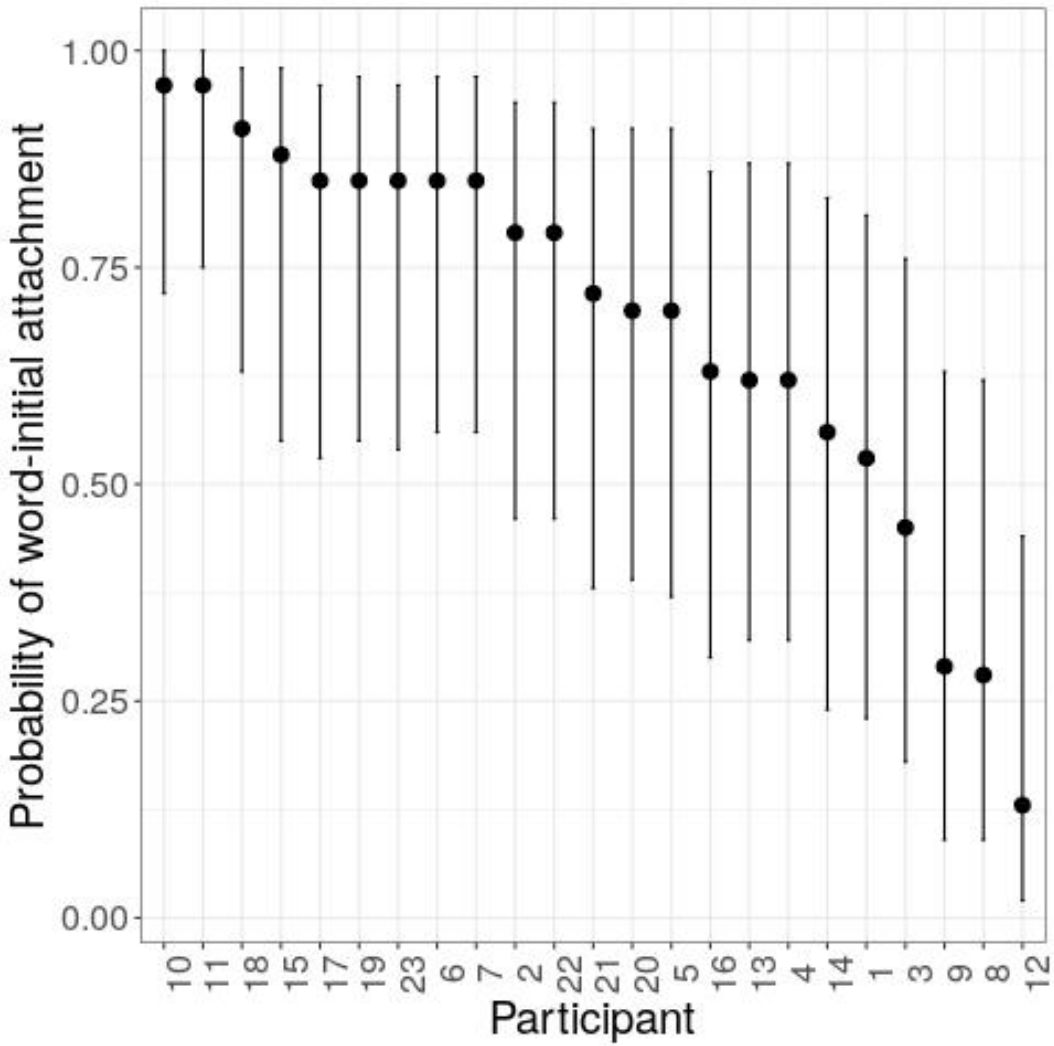
- The figure shows the posterior probability of attachment to word 2 as a function of Consonant (mean and 95% credibility interval).



- The predictions are borne out: only Liaison 1 behaves ambiguously between word-final and word-initial consonant (Liaison 2 behaves like word-final consonants).

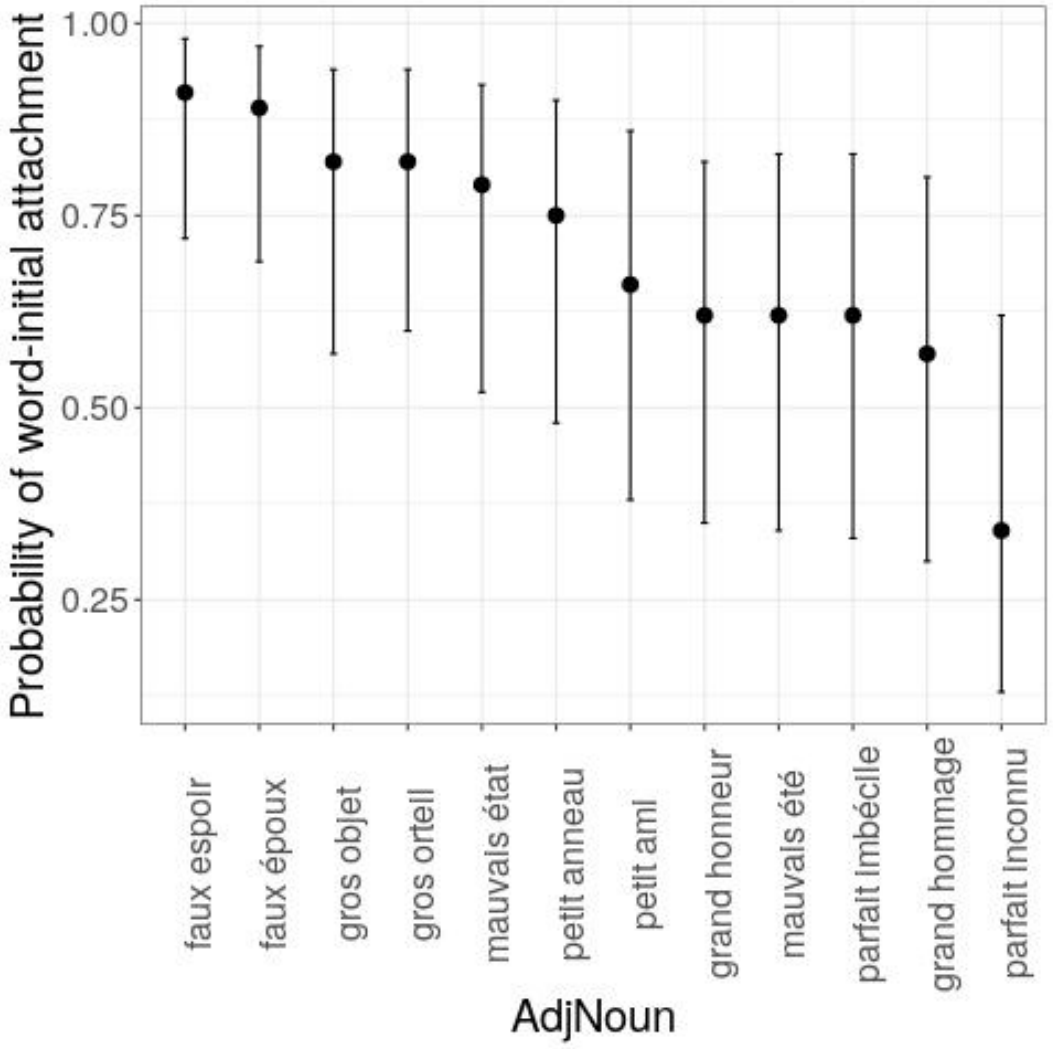
Logistic regression: individual variation for Liaison 1

- We find different behaviors for Liaison 1 at the individual level, with some participants favoring a word-initial attachment and others a word-final attachment.
- But there is generally variation at the individual level.



Logistic regression: item variation for Liaison 1

- We find different behaviors for Liaison 1 as a function of the Adj-Noun sequence.
- However none of them are treated as categorically word-initial or word-final ⇒ Variation at the level of Adj-N sequences.

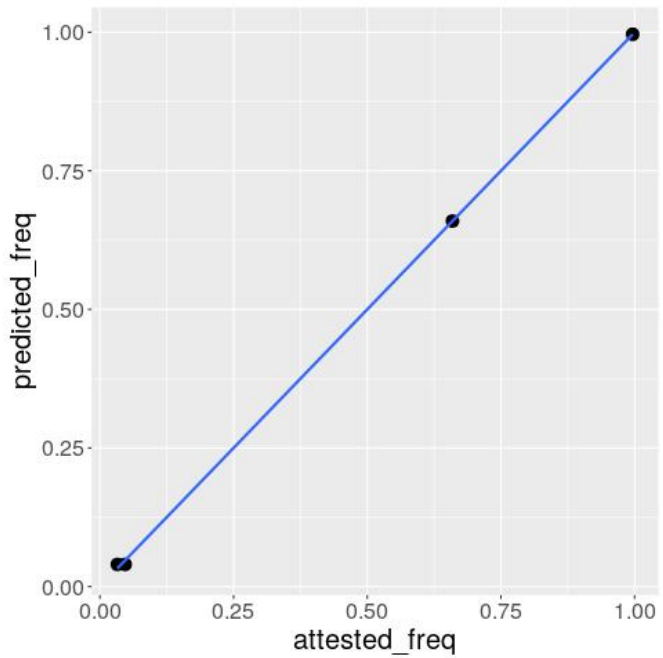


Grammatical analysis

- For the grammatical analysis, the data were aggregated across participants and Adj-N sequences (I model the average individual and the average Adj-N sequence).
- The posterior weights of the grammar are shown in the Table.
  - General preference for word-initial attachment is captured in the relative weights of Dep_{OV}(C)/Adj and Dep_{OV}(C)/Noun.
- The grammar provides a perfect fit to the data, as shown on the right.

Constraint	Mean	95% CI
Dep_{OV}(C)/Adj	2.87	[2.39, 3.38]
Max_{OV}(C)/Adj	1	
Dep_{OV}(C)/Noun	2.21	[1.80, 2.65]
Max_{OV}(C)/Noun	3.33	[1.34, 6.45]

Table: Posterior distribution of the constraint weights (mean and 95% CI)



Conclusion

Conclusion

- It is possible to derive the gradient behavior of liaison consonants (Liaison 1 in particular) without massively enriching phonological and lexical representations.
- The analysis uses independently motivated mechanisms, in particular correspondence between connected-speech variants and citations forms.
- The analysis is predictive: if the liaison form cannot be analyzed as involving epenthesis but only as involving suppletion (use of the feminine form), then it is predicted not to behave gradiently (but as a word-final consonant).
- This prediction is borne out in the experiment and crucially relies on paradigmatic uniformity.

Références I

- Barr, Dale J. et al. (2013). "Random effects structure for confirmatory hypothesis testing: Keep it maximal". In: *Journal of Memory and Language* 68.3, pp. 255–278. doi: 10.1016/j.jml.2012.11.001.
- Bürkner, Paul-Christian (2017). "brms: An R Package for Bayesian Multilevel Models Using Stan". In: *Journal of Statistical Software* 80.1, pp. 1–28. doi: 10.18637/jss.v080.i01.
- Côté, Marie-Hélène (2011). "French liaison". In: *The Blackwell Companion to Phonology*. Ed. by Marc van Oostendorp et al. Chicester, United Kingdom: John Wiley and Sons Ltd.
- – (2014). "Liaison et assibilation en français laurentien". In: *La liaison: approches contemporaines*. Ed. by Christiane Soum-Favaro, Annelise Coquillon, and Jean-Pierre Chevrot. Berne: Peter Lang, pp. 9–32. doi: 10.3726/978-3-0352-0204-5.
- Hayes, Bruce and Colin Wilson (2008). "A Maximum Entropy Model of Phonotactics and Phonotactic Learning". In: *Linguistic Inquiry* 39 (3), pp. 379–440. doi: 10.1162/ling.2008.39.3.379.

Références II

- Kawahara, Shigeto (2002). "Similarity among variants: Output-variant correspondence". Bachelor's thesis. International Christian University (Tokyo, Japan).
- Kilbourn-Ceron, Oriana (2017). "Speech production planning affects phonological variability: a case study in French liaison". In: *Proceedings of the Annual Meetings on Phonology*. Ed. by Karen Jesney et al. doi: http://dx.doi.org/10.3765/amp.v4i0.4004.
- Liddell, Torrin M and John K Kruschke (2018). "Analyzing ordinal data with metric models: What could possibly go wrong?" In: *Journal of Experimental Social Psychology* 79, pp. 328–348. doi: 10.1016/j.jesp.2018.08.009.
- LimeSurvey (2012). *LimeSurvey: An Open Source survey tool*. url: http://www.limesurvey.org.
- Myers, Scott and Jaye Padgett (2014). "Domain generalisation in artificial language learning". In: *Phonology* 31.3, pp. 399–433. doi: 10.1017/S0952675714000207.
- Plénat, Marc (2008). "La liaison'obligatoire'avec et sans enchaînement". In: *Congrès Mondial de Linguistique Française*. EDP Sciences, p. 149.

Références III

- Plummer, Martyn (2016). *rjags: Bayesian Graphical Models using MCMC*. R package version 4-5. url: https://CRAN.R-project.org/package=rjags.
- Smolensky, Paul and Matthew Goldrick (2016). "Gradient symbolic representations in grammar: The case of French liaison". In: *Ms. Johns Hopkins University and Northwestern University. Available as ROA* 1286.
- Steriade, Donca (1997). *Phonetics in phonology: the case of laryngeal neutralization*. Manuscript.
- Storme, Benjamin (2020). "Gradient behavior without gradient underlying representations: the case of French liaison". In: *Proceedings of the 2019 Annual Meeting on Phonology*. doi: 10.3765/amp.v8i0.4650.
- Winter, Bodo and Timo Roettger (2011). "The nature of incomplete neutralization in German: Implications for laboratory phonology". In: *Grazer Linguistische Studien* 76, pp. 55–74.