The Productive Status of Laurentian French Liaison: Variation across Words and Grammar

Anne-Michelle Tessier[1], Karen Jesney[2], Kaili Vesik[1],
Roger Lo[1], and Marie-Eve Bouchard[1]
[1]University of British Columbia, and [2]Carleton University

1 Introduction

There are competing views in contemporary phonological theory about how to best represent processes that are pervasive, frequent, and phonologically motivated, yet still lexically sensitive (see e.g., among many others, Moore-Cantwell 2016; Pater, Staubs, Jesney & Smith 2012; Zuraw 2000). To what extent can – or should – a process that applies idiosyncratically to different morphemes, words, and even phrases, be represented in a way that allows it to generalize to novel forms?

This paper examines this question by looking at prenominal *liaison* as it is used in contemporary Laurentian French. We present the results of an online production study that compares application of liaison in real vs. nonce nouns, and that considers the effect of nonce nouns' phonological properties and morphosyntactic context on the process. We interpret our results as evidence that liaison behaviour is driven jointly by lexical representations and the abstract grammar, with properties of the real-word lexicon affecting the application of liaison to nonce words. We further show that there is considerable variation in the population in the extent to which speakers apply liaison with real *h-aspiré* words, but that all speakers nonetheless share an understanding of what types of words are more vs. less likely to undergo liaison.

The rest of this paper is structured as follows: Section 2 describes the patterns of French liaison and summarizes a range of theoretical approaches that have been applied to the phenomenon, with a particular eye toward their predictions for nonce words. Section 3 presents the methods of our production study; section 4 describes and quantifies the results. Section 5 discusses the implications of our findings for models of liaison, focusing on the types of representations that are needed in order to account for the range of variation seen in our results.

2 French Liaison Patterns and Analyses

Liaison has been extensively studied in phonological theory. This paper deals with a fairly small range of liaison contexts, but for more comprehensive overviews see Côté (2011); Durand & Lyche (2008); Tranel (1995); and references therein.

2.1 *French pre-nominal liaison*   French liaison is a C~zero alternation that occurs at word boundaries. It is widely understood to be phonotactically motivated by a drive to avoid hiatus and/or onsetless syllables, although neither hiatus nor onsetless syllables are generally prohibited in French. In the prenominal context, liaison triggers (Word 1) are a closed class of functional items; some of these (determiners) are obligatory triggers of liaison, and others (adjectives) are optional triggers. Each W1 trigger is consistently associated with a specific liaison consonant – generally either [n], [z] or [t][1]. Examples are given in (1) for two W1s:

[1] Notably, all plural determiners have [z] as their liaison consonant, and a [z]-suffix also marks plural agreement on prenominal adjectives. Less transparently, most singular determiners which trigger liaison use [n], but there are also numerous singular determiners that do not trigger liaison (including all feminine ones), and those that trigger [t] liaison instead (e.g., *cet*).

the plural definite determiner *les* with a liaison [z], and the pre-nominal adjective *petit* with a liaison [t]. When the following noun (Word 2) begins with a consonant – [n] in the case of *nuage* or [b] in the case of *bébé* – no liaison consonant appears between W1 and W2, but when the noun is vowel-initial – as with *ami* or *ours* – the liaison consonant appears.

(1) a.   Word 1 = plural determiner *les*   b.   Word 1 = masculine adjective *petit*
         les nuages   [le.ny.ˈaʒ]   'the clouds'   *petit nuage*   [pə.tˢi.ny.ˈaʒ]   'little cloud'
         les bébés   [le.be.ˈbe]   'the profs'   *petit bébé*   [pə.tˢi.be.ˈbe]   'little baby'
         les amis   [le.**z**a.ˈmi]   'the friends'   *petit ami*   [pə.tˢi.**t**a.ˈmi]   'little friend'
         les ours   [le.ˈ**z**uʁs]   'the bears'   *petit ours*   [pə.tˢi.ˈ**t**uʁs]   'little bear'

There is also an exceptional class of W2s that block the liaison alternation – that is, vowel-initial W2s that behave as though they begin with consonants. Due to their history and the fact that they are almost all spelled with an unpronounced initial <h>, these liaison-resistant words are traditionally called *h-aspiré* words. We adopt the term here to stay somewhat agnostic about their status in contemporary Laurentian French. Three such h-aspiré words are illustrated below in (2). While almost every h-aspiré word is spelled with initial <h>, the opposite implication does not hold; the majority of orthographically h-initial words behave like typical vowel-initial words and follow the liaison pattern in (1) and are known as *h-muet* words.

(2) a.   Word 1 = determiner *les*   b.   Word 1 = masculine adjective *petit*
         les héros   [le.e.ˈʁo]   'the heroes'   *petit héros*   [pə.tˢi.e.ˈʁo]   'little hero'
         les hiboux   [le.i.ˈbu]   'the owls'   *petit hibou*   [pə.tˢi.i.ˈbu]   'little owl'
         les haricots [le.a.ʁi.ˈko] 'the beans'   *petit haricot* [pə.tˢi.a.ʁi.ˈko] 'little bean'

As a very rough estimate, h-aspiré words make up approximately 5% of the relevant vowel-initial French lexicon: a search of the online Usito dictionary of Laurentian French (Cajolet-Laganière et al. 2022) reveals roughly 3600 masculine[2] nouns which are transcribed in citation form with an initial vowel, and of these about 160 are marked as h-aspiré. This set of h-aspiré masculine nouns does include many technical, formal and recently-borrowed words, but as (2) shows, it also includes many quite common nouns. The lexicon of h-aspiré words and their special pronunciation is to some extent a focus of elementary school education, and so most speakers of French have some metalinguistic awareness of this class of words.

There are various grammatical properties of a W2 that make it more likely to act as an h-aspiré word and resist liaison. Here we focus on two such properties: shorter words are more likely to resist liaison than longer ones, and W2s beginning with certain initial vowels, particularly [u], are more likely to resist liaison. The tendency for h-aspiré words to be shorter has been discussed fairly frequently in the literature – see especially the appendix of Zuraw & Hayes (2017) – while the influence of initial vowel quality has been considerably less discussed. We note that in the Usito dictionary there are only 14 words which are spelled with an initial <h> and whose first vowel is [u], and all 14 are in fact h-aspiré. This represents 14/47 (29.8%) of all the [u]-initial masculine nouns, regardless of spelling, making the rate of [#u…] liaison resistance much higher than the roughly 5% found in the general vowel-initial lexicon. This particular lexical quirk appears to derive from two unrelated historical developments, one involving cycles of loss and reintroduction of /h/ into the French segmental inventory and the other involving a set of vowel shifts affecting /u/ and other neighbouring vowels; these historical developments are discussed in some detail by Pope (1934)[3].

As a final caveat: liaison and related processes can occur in many contexts beyond the juncture of a noun and a preceding determiner or adjective, and there are numerous complications. These complications include other types of exceptionality (e.g., *beau/belle*-type alternations, and the behavior of numerals in phrase-final position, as discussed by e.g., Faust 2016; Smith 2015), and a large degree of optionality influenced by factors including register, frequency of words or collocations, morpho-syntactic boundary strength, and speaker

---

[2] Our dictionary search focuses on masculine nouns because in the case of singular W2s, it is only masculine determiners and adjectives that trigger liaison. For additional quantitative discussion of liaison in the Usito dictionary, see Jesney & Tessier, in progress.

[3] Thanks especially to Michael Becker for pointing us to this data and the Pope (1934) source, and to Lev Blumenfeld for related discussion.

education (see Booij & de Jong 1987; Bybee 2001; Encrevé 1988; Laks 2009; among many others). While we set all of these aside, it is in part this set of complications that has given rise to a wide range of liaison analyses, which we turn to next.

2.2 *Representations of French liaison* Given the ink already spilled in analyzing liaison within the phonological literature, we confine ourselves here to only one issue. Our goal is not to describe the grammatical system that drives liaison, nor is it to choose between analyses based on attested and unattested liaison patterns in French dictionary or corpus data. Rather, here we simply aim to group existing types of liaison accounts according to which word they attribute liaison consonants to – i.e., whether they attribute liaison consonants to W1s and/or W2s in the lexicon.

In view of French orthography, the default assumption has often been that liaison consonants are stored uniquely on W1, and this too is the classical generative account, going back at least to Milner (1967) and Schane (1968). On the other hand, evidence and errors from child learners of French suggests that they may initially treat liaison consonants as associated with W2 (*cf.* Buerkin-Pontrelli et al. 2017; Chevrot et al. 2009). In fact, the full range of logical possibilities has been proposed, as summarized in (3), drawing heavily on the survey in Côté (2011).

(3) Representation of liaison consonants in the lexicon

Position of liaison consonant	W1 representation	V-initial W2 representation	Sample references – see especially Côté (2011)
all on W1	/lez/	/ami/	Schane (1968); Selkirk & Vergnaud (1973)
	/{le, lez}/	/ami/	de Jong (1994); Plénat 2008; Smith (2015)
	/le(z)/	/ami/	esp. Tranel (1995)
all on W2	/le/	/zami/ /(tami, nami, ami)/	Ternes (1977); see also Chevrot et al. (2007, 2009)
on neither	/le/	/ami/	Côté (2005); Durand & Lyche (2008); Morin & Kaye (1982)
	/le + z_{pl}/	/ami/	esp. Morin & Kaye (1982)
on both	/$lez_{0.75}$/	/{$t_{0.1}$ $z_{0.7}$ $n_{0.2}$} ami/	Smolensky & Goldrick (2016); Smolensky, Rosen & Goldrick (2020); Tessier & Jesney (2021)

2.3 *Predictions for nonce words* The central concern of this paper is how adult French speakers apply liaison in nonce word contexts. When a speaker observes a novel W2 noun, and sees that it surfaces as V-initial in a non-liaison context, how do they expect it to be realized after a W1 liaison trigger? Will nonce W2s all undergo liaison? Will they sometimes resist liaison, akin to h-aspiré words? To our considerable surprise, it appears that this question has not previously been probed experimentally.

Given the representations of liaison consonants in (3), and assuming that the W1 is a known real word, we see three main possibilities, each with their own predictions for nonce word behaviour.

(4) Predictions for nonce word behaviour

	W1 representation	Real V-initial W2 realization	Real h-aspiré W2 realization	Nonce W2 prediction
A	/lez/	/lez + ami/ → [lezami]	/lez + *ibu/ → [leibu]	[lezatʁ]
B	/le/	/le + *zami/ → [lezami]	/le + ibu/ → [leibu]	[leatʁ]
C	/le/, /lez/	/{le,lez} + *ami/ → [lezami]	/{le,lez} + *ibu/ → [leibu]	[lezatʁ] or [leatʁ]

In (4), the asterisks in each row represent the type of W2 in the French lexicon which needs some additional treatment in order to surface correctly, given that row's W1 representational assumption. In accounts of type A, where *les* is stored with a liaison /z/, it is the h-aspiré W2s like *hibou* that behave exceptionally, and somehow must prevent the /z/ from surfacing. In accounts of type B, where *les* has no /z/ in its stored form, it is the W2s like *ami* which must do something 'exceptional', insofar as they somehow trigger a surface [z]. Finally, in accounts of type C, which store two W1 *les* allomorphs – one with /z/ and one without /z/ – both regular and h-aspiré W2s must somehow select the appropriate W1. There are many different proposals for what these additional treatments of the *W2 forms could be, but here we gloss over these details to focus on the predictions for nonce W2 behaviour. We assume simply that associating some special treatment with a W2 requires positive evidence, and that nonce words will therefore not be subject to this type of special treatment. This yields the set of predictions summarized in (5).

(5)   When combining a W1 trigger like *les* and a novel V-initial word like [atʁ]:
         (4A) accounts predict liaison
         (4B) accounts predict h-aspiré treatment (no liaison)
         (4C) accounts predict both possibilities should occur

3   Methods

Here we report on a production study that tests the predictions in (5) and follows up on the pilot reported in Tessier & Jesney (2020). Compared to the pilot, the method here is virtually unchanged, but a larger and better structured set of stimuli and participants is included.

3.1   *Participants*   A total of 49 participants completed the production experiment. Participation was restricted to adults age 18+ who identified one of their first languages as some variety of Canadian French, and who both grew up with the language and continue to use it frequently in adulthood. Participants were recruited via word of mouth and social media posting within the authors' networks.

Of the 49 respondents, 29 reported that they were current residents of the province of Quebec. These Quebec participants included 1 non-binary, 7 male, and 21 female people; they ranged in age from 23 to 65 years, with a mean age of 41. Most (22/29) had completed their secondary education in French, and most (24/29) had a post-secondary degree. The remaining 20 respondents were speakers of other varieties of Canadian French, but without a large enough dataset from any one such variety, we leave further investigation of this data to future work. Thus in what follows, we report data from the 28 Quebec-based participants who completed the production task with both real and nonce words; given their demographic data, we are confident that their speech variety can be overall classified as Laurentian French. 1 additional set of nonce word responses (from a speaker whose real words were lost due to technical difficulties) is also included.

3.2   *Stimuli*   The table in (6) lists the real W2 nouns included in the experiment: eight regular V-initial and h-muet words that are expected to trigger liaison, and eight h-aspiré words. We designed the h-aspiré list to include nouns of two and three syllables, with a variety of initial vowels. As the glosses in (6) show, the h-aspiré words were of neither extremely high nor extremely low frequency, and they did not include any arcane or technical terms.

(6)   Real word stimuli

V-initial		h-muet		h-aspiré	
amis	'friends'	*habitant*	'resident'	*hamac*	'hammock'
automne	'autumn'	*hiver*	'winter'	*haricot*	'bean'
éléphant	'elephant'	*hôpitaux*	'hospitals'	*hérisson*	'hedgehog'
ours	'bear'	*hôtel*	'hotel'	*héros*	'hero'
				hibou	'owl'
				Hollandais	'Dutch person'
				homard	'lobster'
				houblon	'hops'

The table in (7) lists the nonce words included in the experiment, classified according to syllable count and initial vowel. Nonce words were designed to be phonotactically licit and include common French segments. The third author and another L1 speaker-linguist of Laurentian French reviewed all items in order to rule out any that were overly similar to existing nouns. One item, a 3-syllable word with initial [u], was eliminated during pilot testing when one of the participants indicated that it reminded them of an existing proper noun.

(7)   Nonce word stimuli

	[a]	**[e / ɛ]**	**[i]**	**[o / ɔ]**	**[u]**
1 syllable	[af]	[ɛb]	[ibl]	[ol]	[uʃ]
	[atʁ]	[ɛdʁ]	[im]	[opl]	[udʁ]
3 syllables	[aʃeli]	[ekidu]	[inaʃu]	[olakiz]	[ubʁazi]
	[aspusin]	[epʁaluʃ]	[itiʁuʃ]	[olinuʃ]	

The words in (6) and (7) were recorded in the frame sentences described in section 3.3, with a preceding word that is not a liaison trigger. All stimuli were recorded by a native speaker of Laurentian French using a high-quality microphone. Multiple recordings of each sentence were made, and the clearest repetitions were selected for the final stimulus set.

3.3   *Procedure*   Participants completed the study online via web browser. The study's introductory materials consisted of a consent/information screen, a mandatory 'hearing test' that made it very difficult to continue without using headphones, and a practice screen that allowed the respondent to record themselves and play back their recording to ensure good quality.

The main body of the experiment consisted of a series of trials, each with two parts. These are exemplified in (8) and (9) below. First, as in (8), a sentence appeared on the screen with one word missing, indicated with a blank line. At the same time, participants heard the sentence read aloud, with the blank filled by the trial's target word. All of these sentences presented the target W2 in a non-liaison context, following the W1 *joli*, *jeune, chaque*, or *six*. These W1s invariantly end in [i], [n], [k] and [s], respectively, and so give no evidence regarding the following word's liaison status.

(8)   Study trial part 1 (example)
         Seen on screen: Ici il y a six _____ .
         Heard over headphones: «Ici il y a six [**ami**]»
                                      Here there are six friends.

This was followed immediately with a new sentence appearing on the screen, again with one missing word indicated by a blank line, as in (9). When ready, participants clicked a button on the screen to begin recording. They then read the sentence aloud, inserting the target word that had filled the blank on the previous screen. When finished, they clicked the button again to end recording and continue to the next trial. In this second set of sentences, the target word was always preceded by one of three W1 liaison triggers: the singular indefinite determiner *un*, the plural definite determiner *les* or the adjective *petit*. As noted in section 2.1, determiners are generally obligatory triggers of liaison, while liaison is optional with adjectives.

(9)   Study trial part 2 (example)
         Seen on screen: Ce sont les _____ du prof.
         Recorded by participant: «Ce sont [**lezami**] du prof» OR «Ce sont [**leami**] du prof»
                                      They are the prof's friends.

The trials were presented in two blocks: first the 16 real word items and then the 19 nonce word items. In both blocks, all participants saw the trials in the same order, but the trials were pseudo-randomized to mix item types within a block – i.e., V-initial vs. h-muet vs. h-aspiré were mixed among the real words, and syllable count and vowel quality were mixed among the nonce words.

After finishing the nonce word block, participants completed a related acceptability study (see section 5.2) and then a short demographic survey. Finally, participants could provide an e-mail address to receive a 15 CAD gift certificate to an online French bookstore as a token of appreciation.

4 Results

4.1 *Real and nonce words* The boxplot in Figure 1 presents the overall results; dots show each participant's rate of liaison production for the four word types. As expected, for both regular V-initial and h-muet words, there was a high rate of liaison across participants – V-initial words *Median* = 100.00%, *Mean* = 97.32% (*SD* = 10.41), h-muet words *Median* = 100.00%, *Mean* = 92.86% (*SD* = 15.00). Traditional h-aspiré words showed substantially lower rates of liaison (*Median* = 12.5%, *Mean* = 19.64%, *SD* = 22.93). Nonce words patterned with neither the V-initial nor h-aspiré words, instead showing intermediate rates of liaison (*Median* = 68.42%, *Mean* = 65.27%, *SD* = 26.25).
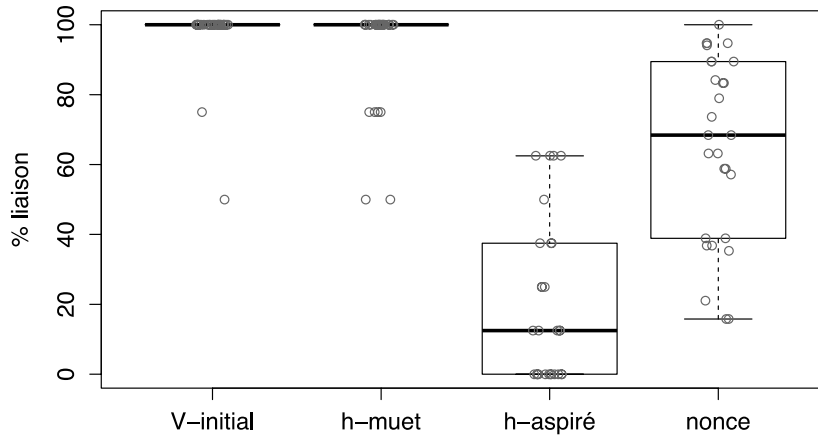

Figure 1: Participant rates of liaison production across W2 types

While there was a high degree of variation in precise rates of liaison production across participants, particularly with respect to h-aspiré and nonce words, the overall pattern in Figure 1 was generally replicated in individual participants' data. Figure 2A plots each participant's rate of liaison with nonce words (*y*-axis) against their rate of liaison with traditional h-aspiré words (*x*-axis). All but four dots fall above the diagonal line, indicating that almost all participants did more liaison with nonce words than with h-aspiré words, as suggested by the overall data shown in Figure 1.
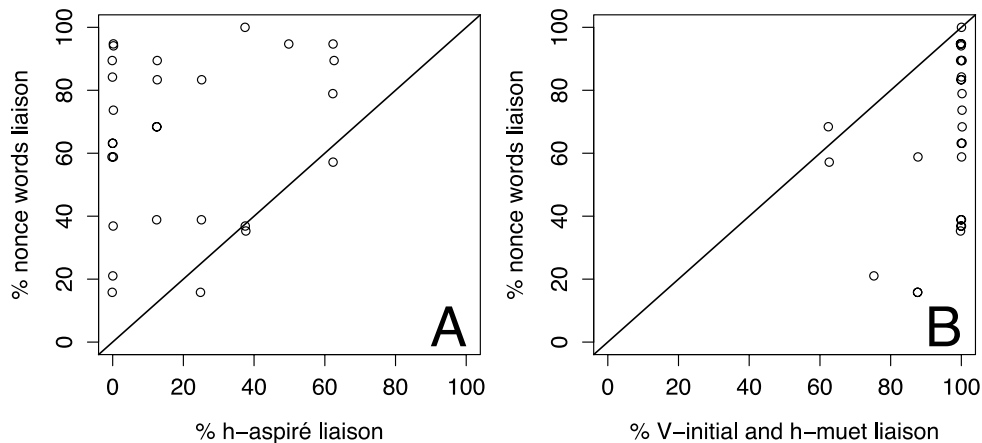

Figure 2: Individual participants' rates of liaison for h-aspiré words vs. nonce words (A) and vowel-initial and h-muet words vs. nonce words (B)

Figure 2B plots each participant's rate of liaison in nonce words (again on the *y*-axis) against their rate of liaison for vowel-initial and h-muet words (*x*-axis). Here, only one dot falls above the diagonal line, demonstrating that almost all participants did more liaison with vowel-initial and h-muet words than with nonce words.

Figure 2 also illustrates the extent of variability across participants. As the clustering of dots on the right side of Figure 2B shows, participants were generally traditional in their treatment of vowel-initial and h-muet words. Twenty-two of the twenty-eight participants produced liaison in every case, and only two participants produced liaison with fewer than 7/8 items. Participants were considerably more variable in their treatment of h-aspiré words. Just under half of the participants (12 of 28) replicated the traditional behaviour described in the literature, producing no liaison with any of the eight h-aspiré words, while the remaining participants produced liaison with between one and five h-aspiré words. There was no clear correlation between an individual's rate of liaison with h-aspiré and nonce words (Pearson's *r* = 0.327, *p* = 0.09); participants who produced no liaison with h-aspiré words varied between 15.79% and 93.74% in their production of liaison with nonce words. However, the general absence of dots below the diagonal line in Figure 2A suggests that a given participant's rate of liaison with h-aspiré words acts as a lower bound on their rate of liaison for nonce words – a point that we return to in section 5.2.

4.2 *Predictors of liaison rates in nonce words* Several effects were noted among the nonce words. First, as anticipated, overall liaison rates were higher when W1 was the determiner *un* (70.41%) or *les* (71.35%) than when it was the adjective *petit* (54.17%). Second, the length effect noted in the literature was replicated here; liaison was produced 60.15% of the time with monosyllabic W2s vs. 71.54% of the time with trisyllabic W2s. Finally, the W2 initial vowel quality effect was also found here; the lowest rate of liaison was seen in nonce words with initial [u] (52.38%). A mixed logistic regression model with participant as a random factor, shown in (9), confirmed the statistical significance of these results (R Core Team 2019).
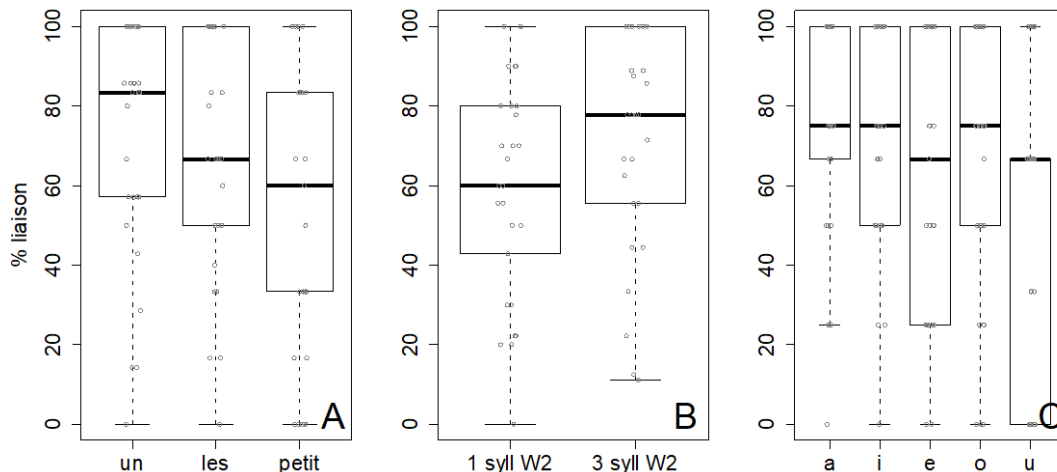

Figure 3: Nonce word liaison rates for participants based on W1 *un* vs. *les* vs. *petit* (A), syllable count of the nonce W2 (B), and initial vowel of the nonce W2 (C)

(10)   Predictors of nonce word liaison (fixed effects)

	Estimate	S.E.	*z* value	Pr > |z|	
(Intercept)	1.1114	0.4397	2.528	0.011	*
W1 = *un* (vs. *les*)	0.3597	0.2921	1.231	0.218	
W1 = *petit* (vs. *les*)	−0.8587	0.2806	−3.060	0.002	**
W2 length = 3 sylls (vs. 1 syll)	0.8157	0.2362	3.452	< 0.001	***
W2 vowel = e (vs. a)	−0.4292	0.3529	−1.216	0.224	
W2 vowel = i (vs. a)	−0.0696	0.3597	−0.193	0.857	
W2 vowel = o (vs. a)	−0.6360	0.3465	−1.835	0.066	
W2 vowel = u (vs. a)	−1.0775	0.3681	−2.927	0.003	**

As noted in section 4.1, twelve of the twenty-eight participants were traditional in their treatment of real h-aspiré words, never producing liaison with these items, while the remaining sixteen participants were innovative and produced liaison with at least one real h-aspiré word. As we would expect, given the data in Figure 2A, the more traditional group had a slightly lower rate of liaison in nonce words (62.61%) than did the innovative group (67.35%). Even considered separately, however, both groups showed a significant effect of W2 syllable count on the probability of nonce-word liaison: for the innovative group $\beta = 0.925$, $SE = 0.339$, $p = .006$; for the traditional group $\beta = 0.703$, $SE = 0.349$, $p = 0.04$. Only the innovative group also showed significant effects of W1 being *petit* ($\beta = -1.329$, $SE = 0.405$, $p = 0.001$) and of the initial vowel of W2 being [u] ($\beta = -1.486$, $SE = 0.531$, $p = 0.005$). Figure 4 shows that both groups displayed the same overall patterns, however, suggesting that the failure to find a significant effect of W1 and initial vowel in the case of the traditional group was due to a lack of statistical power, rather than a qualitative difference in the two groups' treatment of nonce words.
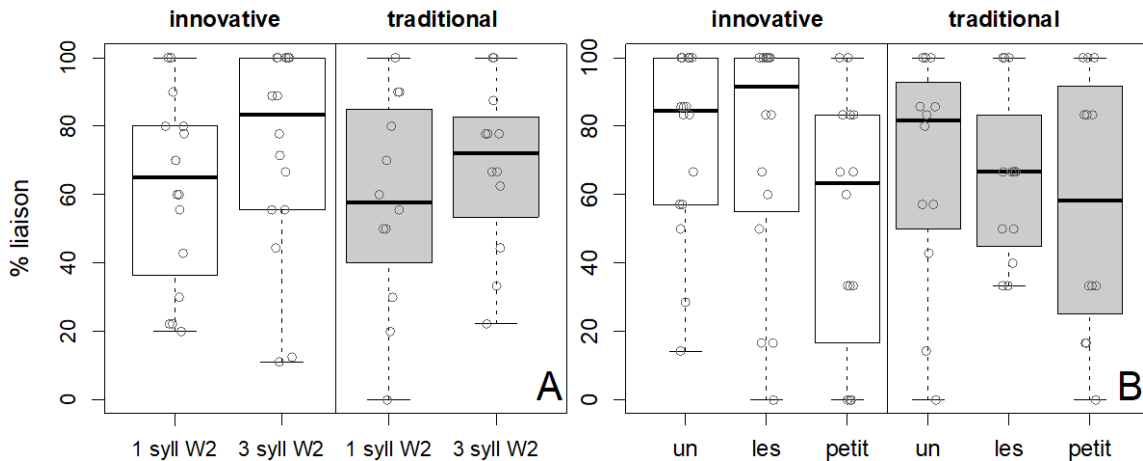

Figure 4: Nonce word liaison rates for innovative group (white bars) and traditional group (grey bars) based on syllable count of the nonce W2 (A), and W1 *un* vs. *les* vs. *petit* (B)

5   Discussion

5.1   *Assessing predictions of liaison accounts*   In the aggregate, nonce words in our study were intermediate in their liaison behaviour: they had lower levels of liaison than regular V-initial and h-muet words, and higher levels of liaison than h-aspiré words. To return to the predictions laid out in section 2.3, this finding is most in keeping with accounts of type (4C). Under these accounts, there are two co-existing representations of each W1: one that includes the liaison consonant /œ̃n, lez, pǝtit/, and one that omits the liaison consonant and ends with a vowel /œ̃, le, pǝti/. In the case of real words, each W2 selects for the appropriate allomorph, so that a real vowel-initial word will select /œ̃n, lez, pǝtit/, and a real h-aspiré word will select /œ̃, le, pǝti/. With a nonce word, where category membership has not been established through previous evidence, the speaker is somewhat free to choose either allomorph and produce the item with or without liaison.

The other two types of liaison accounts in (4) do not clearly predict the kind of variable nonce word behaviour seen in our data. In type (4A) accounts, where there is a single W1 allomorph that includes the liaison consonant /œ̃n, lez, pǝtit/, liaison is expected to apply to all vowel-initial nonce words. On these accounts, speakers must learn that a certain set of W2 lexical items – h-aspiré words – are exceptional in blocking the liaison consonant from surfacing. Without evidence demonstrating that a nonce word is exceptional, nonce words should consistently undergo liaison. On the other hand, the type (4B) accounts' single W1 allomorph without a liaison consonant /œ̃, le, pǝti/ predicts that liaison will <u>not</u> apply to any vowel-initial nonce words. On these accounts, speakers must learn a form of exceptionality associated with the "opposite" set of W2 lexical items – i.e. regular h-muet and V-initial words. To the extent that a given nonce word does not present evidence of such exceptionality, liaison is expected to be blocked.

5.2 *Implications*   The results of this study indicate that speakers are able to select either liaison or non-liaison W1 forms for nonce words. The rate at which participants blocked liaison with nonce words – i.e., treated them as though they were traditional h-aspiré words – was highly variable across speakers, ranging from 0% to 84.21%. This said, speakers generally chose non-liaison at a rate that was higher than the proportion of h-aspiré words in the relevant subset of the French lexicon, which we estimated above to be about 5%. In other words, the overall rate of (non-)liaison with nonce words cannot be viewed simply as frequency matching.

It is possible that liaison blocking was artificially boosted in our nonce word study because h-aspiré words accounted for half of the real words that participants were exposed to in the preceding block (8 V-initial / h-muet items vs. 8 h-aspiré items). While this might indeed have primed speakers for liaison blocking to some extent, it still seems to us that the nonce words showed a substantially higher rate of non-liaison than would be expected if speakers were drawing solely on their knowledge of the real word W2 lexicon to estimate the probability that a new word should resist liaison. Further evidence for this conclusion comes from the fact that we did not see a significant correlation between individual participants' treatment of nonce words and the extent of their innovative behaviour with real h-aspiré words. As discussed in section 4.1, participants' individual rates of liaison with real h-aspiré words seemed to merely provide a lower bound on their rates of liaison with nonce words, rather than predicting those rates precisely.

The fact that participants differed to such an extent in how frequently they used liaison with real h-aspiré words is also informative. As mentioned in section 3.2, our real words were common h-aspiré masculine nouns chosen with the goal of including some different word lengths and initial vowel qualities. We speculate that looking at more of the h-aspiré lexicon, including more infrequent words, would only reveal even more variation across speakers. Despite this variability, participants' treatment of nonce words still showed a sensitivity to the statistical properties of the dictionary lexicon's set of h-aspiré words. We interpret this connection, which holds both for the group and for individual speakers, as strong evidence that participants were indeed using their knowledge of real h-aspiré words, of whatever shape, to decide the liaison fate of nonce words.

The further fact that these effects were seen just as strongly (if not more so) in the group that regularized h-aspiré words to a greater extent (the innovative group in Figure 4), makes this especially interesting. One potential account of participants' variability would be that each speaker has a different set of real words that do vs. do not undergo liaison in their lexicons. While this appears to be true in their productions within the context of the experiment, their nonce word behavior at the group level still suggests a set of shared intuitions regarding the phonological properties of the h-aspiré lexicon. In other words, even those innovative speakers who regularly use liaison with a relatively large number of real h-aspiré words must still be relying on traditional h-aspiré representational knowledge when determining the realization of nonce words.

This idea is reinforced by the results of the acceptability study completed by the same group of participants. After completing the production task reported here, the participants provided judgements of the relative acceptability of liaison vs. non-liaison pronunciations for a variety of real h-aspiré and h-muet words using a six-point Likert scale. This task revealed a significant difference in the relative acceptability of innovative realizations of h-aspiré vs. h-muet words in obligatory liaison contexts following the W1 determiners *un* and *les*. While all participants clearly preferred traditional realizations of both word types, they were significantly more accepting of innovative realizations of h-aspiré words (i.e., those with liaison) than they were of innovative realization of h-muet words (i.e., those without liaison). H-aspiré forms with liaison were rated as better formed ($M = 1.822$, $SD = 1.64$) than *h*-muet forms without liaison ($M = 1.15$, $SD = 1.10$); $t(148) = 3.003$, $p = 0.003$. This difference in the means, along with the relatively greater standard deviation for h-aspiré words, indicates that, as a group, participants viewed liaison as at least somewhat acceptable for h-aspiré words. Thus, we conclude that there remains a clear, shared lexicon of h-aspiré words that can be used as a basis for lexical generalization, despite the fact that speakers are less absolute in their blocking of liaison with h-aspiré forms than traditional descriptions suggest.

5.3 *Future directions*   Taken together, our results present many directions for future work. We will highlight just two of these here.

First, the preliminary acceptability study results described in the previous section suggest that speakers' representations of h-aspiré words are perhaps in a state of flux. Participants vary in the extent to which they

are willing to realize real h-aspiré words with liaison, but as a group they show awareness of the properties of this class of words in their acceptability responses and in their production of nonce words. It remains to be determined what this might mean for models of phonological and lexical representation.

Second, while it seems clear that an account of liaison and h-aspiré behaviour must allow both W1 and W2 to play a role, the path by which this shared lexical responsibility can be learned is not evident. Existing data regarding children's acquisition of liaison provides considerable evidence that they can initially attribute relatively more responsibility to W2 representations, leading to errors where a liaison consonant is attached to a vowel-initial word even in the absence of a W1 – e.g., realizing *ami* as **n**ami or **z**ami in non-liaison contexts (see especially Chevrot et al. 2009). Studying how children move from this early understanding of liaison through to the adult patterns reported in the present work should, we think, provide insight into both the grammar of liaison and the representation of the relevant Laurentian French lexical classes.

References

Booij, Geert & Daan de Jong. 1987. The domain of liaison: theories and data. *Linguistics* 25: 1005-1025.

Buerkin-Pontrelli, Andrea, Jennifer Culbertson, Géraldine Legendre & Thierry Nazzi. 2017. Competing models of liaison acquisition: evidence from corpus and experimental data. *Language* 93(1): 189-219.

Bybee, Joan. 2001. Frequency effects on French liaison. In Joan Bybee & Paul Hopper (eds.), *Frequency and the emergence of linguistic structure*, 337-359. Amsterdam: John Benjamins.

Cajolet-Laganière, Hélène, Pierre Martel, Chantal-Édith Masson & Louis Mercier (eds.). 2022. Le dictionnaire Usito. Sherbrooke, QC: Université de Sherbrooke. (usito.usherbrooke.ca) (Accessed 2023-01-15.)

Chevrot, Jean-Pierre, Damien Chabanal & Céline Dugua. 2007. Pour un modèle de l'acquisition des liaisons basé sur l'usage: Trois études de cas. *Journal of French Language Studies* 17: 103-128.

Chevrot, Jean-Pierre, Céline Dugua & Michel Fayal. 2009. Liaison acquisition, word segmentation and construction in French: a usage-based account. *Journal of Child Language* 36: 557-596

Côté, Marie-Hélène. 2005. Le statut lexical des consonnes de liaison. *Langages* 158: 66-78.

Côté, Marie-Hélène. 2011. French liaison. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, 2685-2710. Malden, MA: Wiley-Blackwell.

de Jong, Daan. 1994. La sociophonologie de la liaison orléanaise. In Chantal Lyche (ed.), *French generative phonology: retrospective and perspectives*, 95-130. Salford : ESRI

Durand, Jacques & Lyche, Chantal. 2008. French liaison in the light of corpus data. *French language studies* 18: 33-66.

Encrevé, Pierre. 1988. *La liaison avec et sans enchaînement: Phonologie tridimensionnelle et usages du français*. Paris: Editions du Seuil.

Faust, Noam. 2016. *Exceptional liaison words in French.* Ms., Univ. Paris 8. (https://www.academia.edu/27762471/Exceptional_liaison_words_in_French) (Accessed 2023-01-15.)

Jesney, Karen & Anne-Michelle Tessier. in prep. Production, perception and acquisition of liaison in Laurentian French: the development of representations. Ms., Carleton University and UBC.

Klausenburger, Jürgen. 1974. Rule inversion, opacity, conspiracy: French liaison and elision. *Lingua* 34: 167-179

Laks, Bernard. 2009. Dynamiques de la liaison en français. In Luc Baronian & France Martineau (eds.), *Le français d'un continent à l'autre*, 237-267. Québec, QC: Presses de l'Université Laval.

Milner, Jean-Claude. 1967. French truncation rule. *MIT Quarterly Progress Report* 86: 273-283.

Moore-Cantwell, Claire. 2016. *The representation of probabilistic phonological patterns: neurological, behavioral, and computational evidence from the English stress system.* PhD dissertation. University of Massachusetts Amherst.

Morin, Yves-Charles & Jonathan Kaye. 1982. The syntactic bases for French liaison. *Journal of Linguistics* 18: 291-330.

Pater, Joe, Robert Staubs, Karen Jesney & Brian Smith. 2012. Learning probabilities over underlying representations. In *Proceedings of the Twelfth Meeting of the Special Interest Group on Phonetics, Phonology, and Morphology (SIGMORPHON)*, 62-71. Montréal, PQ: Association for Computational Linguistics.

Plénat, Marc. 2008. La liaison "obligatoire" avec et sans enchaînement. In Jacques Durand, Benoît Habert & Bernard Laks (eds.), *Congrès mondial de linguistique française - CMLF'08,* 1657-1667. Paris: Institut de Linguistique Française.

Pope, Mildred. 1934. *From Latin to Modern French.* Manchester University Press.

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. (https://www.R-project.org)

Schane, Sanford. 1968. *French phonology and morphology.* Cambridge, MA: MIT Press.

Selkirk, Elisabeth & Jean-Roger Vergnaud. 1973. How abstract is French phonology? *Foundations of Language* 10: 249-254.

Smith, Brian W. 2015. *Phonologically conditioned allomorphy and UR constraints*. PhD dissertation. University of Massachusetts Amherst.

Smolensky, Paul & Matthew Goldrick. 2016. Gradient Symbolic Representation in grammar: The case of French liaison. Ms., Johns Hopkins University and Northwestern University. (ROA-1552).

Smolensky, Paul, Eric Rosen & Matthew Goldrick. 2020. Learning a gradient grammar of French liaison. In Hyunah Baek, Chikako Takahashi, Alex Hong-Lun Yeung (eds.), *Proceedings of the 2019 Annual Meeting of Phonology*. Washington, DC: Linguistic Society of America.

Ternes, Elmar. 1977. Konsonantische Anlautveränderungen in den keltischen und romanischen sprachen. *Romanistisches Jahrbuch* 28: 19-53.

Tessier, Anne-Michelle & Karen Jesney. 2021.. Learning French liaison with Gradient Symbolic Representations: errors, predictions, consequences. In Ryan Bennett, Richard Bibbs, Mykel Loren Brinkerhoff, Max J. Kaplan, Stephanie Rich, Nicholas Van Handel & Maya Wax Cavallaro (eds.), *Proceedings of the 2020 Annual Meeting on Phonology*. Washington, DC: Linguistic Society of America.

Tranel, Bernard. 1995. Current issues in French phonology: liaison and position theories. In John A. Goldsmith (ed.), *The Handbook of Phonological Theory*, 798-816. Cambridge, MA: Blackwell.

Zuraw, Kie. 2000. *Patterned Exceptions in Phonology*. PhD dissertation. University of California, Los Angeles.

Zuraw, Kie & Bruce Hayes. 2017. Intersecting constraint families: An argument for Harmonic Grammar. *Language* 93(3): 497-548.