Learning a model of gradient French liaison

French liaison, h-aspiré and ə-deletion (Côté (2008)) are phenomena that interact in complex ways. Liaison consonants $\mathcal{L}$ are weak elements that only appear in inter-word contexts (W$_1$ __ W_2) when W$_2$ is vowel-initial – unless W$_2$ is an h-aspiré word like *héros* 'hero' which appears vowel-initial but (a) resists liaison and ə-deletion after *le/la* 'the' and (b) optionally preserves ə in preceding *une* 'a' (fem.) We present an account of these interacting phenomena that improves on a recent analysis by Zuraw and Hayes (2017) (Z&H), where the phenomena we cover go well beyond their purely orthographic data and our analysis avoids their proliferation of lexically indexed constraints like USE-X for an allomorph X.

Our account extends an analysis by Smolensky and Goldrick (2016) (S&G) that goes beyond their hand-calculated values by (a) providing a tested, error-driven learning algorithm for constraint weights and feature activations and (b) proposing that the partially-activated input values in their account result in split morpheme boundaries ((1) below), a move that can account for gradient patterns of what Z&H refer to as 'alignancy': the tendency for morphemes to align with syllable edges. In S&G's Gradient Symbolic Computation framework (GSC), a type of Harmonic Grammar with weighted constraints and partially-activated features, liaison is derived from the coalescence of partially-activated input consonants $\mathcal{L}_1$ and $\mathcal{L}_2$ that occur finally in W_1 ($\mathcal{L}_1$) and initially in W_2 ($\mathcal{L}_2$). A liaison consonant $\mathcal{L}$ surfaces iff its aggregate activation surpasses an epiphenomenal threshold determined by weighted MAX and DEP constraints. A MAX constraint rewards a surfacing input by MAX's weight times the input activation. A DEP constraint entails a harmonic penalty for a surfacing input by DEP's weight times the deficit between full and input activation.

We adopt their proposal that h-aspiré words have no initial $\mathcal{L}_2$, resulting in lack of required activation for a $\mathcal{L}$ to surface. We relativize their proposed Alignment (actually anchoring) constraints to vowels vs. consonants. Our learning algorithm found a strong weight to ANCHOR-V-LEFT-MORPHEME-LEFT-σ.

(1) ANCHOR-V-LEFT-MORPHEME-LEFT-σ: "If the left edge of a vowel V$_i$ is aligned with the left edge of a morpheme M$_j$ in the input, then the left edtges of V_i and M_j align with a syllable in the output." (Rewards positive Harmony to the weight of the constraint.)

This constraint accounts for the tendencies of h-aspiré words (a) to preserve a preceding ə after *le/la* and optionally after *une* and (b) to syllabify a preceding C as a coda *quel hibou* (kɛl.i.bu)) 'what an owl'. These cases all earn a harmonic reward for satisfying the constraint, which is only weakly satisfied by words like *ami* 'friend' with a proposed $\mathcal{L}_2$ at their left edge. Gradient morpheme boundaries, shown by superscripted values beside square brackets in *petit ami* 'boyfriend' below are commensurate with subscripted activations of edge segments. (See also (6) on p. 2.)

(2) [pəti]$^{.52}$t$_{.48}$]$^{.48}$ $[^{.09}\{t, z, n\}_{.09}[^{.91}$ami]

Z&H observe interacting gradient tendencies towards alignant behaviour (a) among h-aspiré words as W_2 and among alignant allomorphs of lexemes like *beau∼bel*, *le/la∼l'* as W_1, accounting for them through morpheme- or group-specific constraints such as USELE for W_1 and a set of lexically indexed ALIGNMORPHLSYLLL constraints for W_2. Our model avoids the proliferation of lexically-indexed constraints by encoding these gradient behaviours through the activations of $\mathcal{L}_1$ and $\mathcal{L}_2$ in the input. We account for weak-h-aspiré words like *Hollandais* that optionally permit liaison, through a weak $\mathcal{L}_2$ activation and the alignant tendency of W_1s like *le* through the amount of input activation on ə, where persistence of ə results in alignancy. Because the anchoring constraint considers a partially-activated morpheme-edge input segment to split the morpheme boundary before and after it, it rewards syllable-edge alignment according to the amount of input activation, whose magnitude will determine the tendency towards alignancy. (See (6) on p. 2.)

To test the learnability of this model, we ran an error-driven algorithm, similar to Boersma and Pater's (2016) Gradual Learning Algorithm, on inter-related data, comprising 55 examples. For each example, the candidate with the greatest Harmony is calculated from input activations and constraint weights. If the wrong winner is chosen, it increases the weights of constraints and activations on segments that favour the desired winner and decreases those favouring the false winner. It found constraint weights and input activations that derived all the examples correctly. The tableaux below illustrate (3) surfacing of $\mathcal{L}$ through coalescence, (4) lack of $\mathcal{L}$ before h-aspiré, (5) persistent consonants in feminine forms resulting from a pure activation input ϕ for the feminine morpheme which boosts $\mathcal{L}_2$ to make it always surface. (5) captures the generalization that weak segments appear in specific morphological environments: e.g., a final weak $\mathcal{L}$ in masculine prenominal adjectives (*petit*) vs. strong C in feminine (*petite*). Z&H's model would need to capture this alternation through separate listings or a morpheme-specific constraint for each lexeme of the form USE-X where X is a citation form, missing the generalization that weak segments appear in specific morphological environments. This also leads to awkward formulations for morphemes like z_{pl}, which surfaces as a null segment in the alignant cases that Z&H derive with a "Use-X" constraint. We would need a constraint USE-NULL for the plural, with its citation form being a null morpheme.

(3) petit ami	⚓L-V	ONSET	MAX	DEP	H
'boyfriend'	1.07	-0.72	0.6	-0.58	
$\mathcal{L}_1$=$(0.48 \cdot t)_1$ $\mathcal{L}_2$=$\{t, z, n\}$ =.09					
/pəti$\mathcal{L}_1$/ + /$\mathcal{L}_2$ami/					
(a) pə.ti.a.mi	0.78	-0.72			0.06
(b) **pə.ti.t$_{12}$a.mi**			0.34	-0.25	0.11

(4) petit héros 'little hero'	⚓L-V 1.07	ONSET -0.72	MAX 0.6	DEP -0.58	H
/pəti$\mathcal{L}_1$/ + /eʁo/					
(a) **pə.ti.e.ʁo**	1.07	-0.72			0.35
(b) pə.ti.t$_1$eʁo			0.29	-0.30	-0.01

(5)	MAX	DEP	NOCODA	H
	0.6	-0.58	-0.34	
petit 'small (m.)' /pəti$\mathcal{L}_1$/ $\mathcal{L}_1$=$(0.48 \cdot t)_1$				
(a) **pə.ti**				0
(b) pə.tit	0.29	$-.30$	-0.34	-0.35
petite 'small (f.)' /pəti$\mathcal{L}_1$/ + ϕ_{fem} ϕ_{fem} = 0.35				
(a) pə.ti				0
(b) **pə.tit**	0.50	-0.10	-0.34	0.06

In the GSC framework, the probability of a candidate is proportional to the exponential of its Harmony, where learning constraint weights solves the same problem that MaxEnt models solve except for the additional problem in GSC of learning activity levels. If two candidates differ only slightly in Harmony, their probabilities will be similar and we predict optionality. This property of GSC accounts for optional liaison before weak-h-aspiré words like *Hollandais* 'Dutch', where, in our account, *Hollandais* ends up with a very weak $\mathcal{L}_2$ activation. Z&H explain such phenomena through lexically indexing varying weights of an Alignment constraint to different classes of alignant words.

In (6) weak $\mathcal{L}_2$ activation on *Hollandais* of $0.02 \times 3 = 0.06$ means that its left morpheme boundary is split in the input 94%∼6%, making the reward from ⚓L-V $1.07 \times 0.94 = 1.01$ for the alignant candidate, which results in near-tied Harmony in the tableau. This can account for the optionality of alignancy for this word as described by Tranel (1996).

(6) le/l' Hollandais	⚓L-V	ONSET	MAX	DEP	H
lϵ + $\mathcal{L}_2$/olandɛ ϵ=ə$\cdot 0.25$, $\mathcal{L}_2$/=$\{t, z, n\}'_2$=0.02, 0.02, 0.02					
'the Dutch'	1.07	-0.72	0.6	-0.58	
(a) **lə.o.lan.dɛ**	1.01	-0.72	0.15	-0.45	-0.01
(b) **lo.lan.dɛ**					0

References Marie-Hélène Côté. 2008. Empty elements in schwa, liaison and h-aspiré: The French Holy Trinity revisited. pp. 61-103 in Hartmann J.M., Hegedüs V., van Riemsdijk H. (eds.) Sounds of silence: Empty elements in syntax and phonology, Elsevier. ⊕ Paul Boersma & Joe Pater. 2016. Convergence properties of a gradual learning algorithm for Harmonic Grammar. In John McCarthy & Joe Pater (eds.): Harmonic Serialism and Harmonic Grammar, 389-434. Sheffield: Equinox. ⊕ Paul Smolensky and Matt Goldrick. 2016. Gradient Symbolic Representations in Grammar: The case of French Liaison. ROA 1552. ⊕ Kie Zuraw and Bruce Hayes. 2017. Intersecting constraint families: An argument for harmonic grammar. Language 93: 497-548. ⊗ Bernard Tranel. 1996. Current Issues in French Phonology: Liaison and Position Theories, in The Handbook of Phonological Theory, J. Goldsmith, ed.