On the representation of /h/ by Quebec francophone learners of English

Paul John[1]* and Simon Rigoulot[2]

[1]CogNAC Research Group, Department of Modern Languages and Translation, Université du Québec à Trois-Rivières, Trois-Rivières, QC, Canada, [2]CogNAC Research Group, Department of Psychology, Université du Québec à Trois-Rivières, Trois-Rivières, QC, Canada

The current study investigates whether some of the variation in h-production observed among Quebec francophone (QF) learners of English could follow from their at times assimilating /h/ to /ʁ/. In earlier research, we attributed variation exclusively to QFs developing an approximate ("fuzzy" or "murky") representation of /h/ that is not fully reliable as a base for h-perception and production. Nonetheless, two previous studies observed via event-related potentials differences in QF perceptual ability, which may follow from the quality of the vowel used in the stimuli: /ɑ/ vs. /ʌ/ (detection vs. no detection of /h/). Before the vowel /ɑ/, /h/ exhibits phonetic properties that may allow it to be assimilated to and thus underlyingly represented as /ʁ/. If /h/ is at times subject to approximate representation (e.g., before /ʌ/) and at others captured as /ʁ/ (before /ɑ/), we would expect production of /h/ to reflect this representational distinction, with greater accuracy rates in items containing /ɑ/. Two-way ANOVAs and paired Bayesian *t*-tests on the reading-aloud data of 27 QFs, however, reveal no difference in h-production according to vowel type. We address the consequences of our findings, discussing notably why QFs have such enduring difficulty acquiring /h/ despite the feature [spread glottis] being available in their representational repertoire. We propose the presence of a Laryngeal Input Constraint that renders representations containing only a laryngeal feature highly marked. We also consider the possibility that, rather than having overcome this constraint, some highly advanced learners are "phonological zombies": these learners become so adept at employing approximate representations in perception and production that they are indistinguishable from speakers with bona fide phonemic representations.

KEYWORDS

L2 phonological acquisition, perceptual assimilation, approximate representations, variation, h-deletion, input constraints, phonological zombies

1 Introduction

Motivated by the findings of two earlier studies on the perception of /h/ (White et al., 2015; Mah et al., 2016), the study presented here investigates the production of /h/ before the vowels /ɑ/ (*hot*) and /ʌ/ (*hut*) by Quebec francophone (QF) learners of English. Francophones, whether from Quebec or elsewhere, struggle to produce /h/, the tendency being to delete this non-native phoneme (*hair → _air*) or even to epenthesize it before vowel-initial forms (*ankle →* [h]*ankle*) (John and Cardoso, 2009; John and Frasnelli, 2022). H-deletion constitutes the basic QF pronunciation error, instantiated notably in loanwords to Quebec French (*hotdog → _otdog*; Paradis and Lebel, 1994), with loanwords corresponding to a kind of "ground zero" for L2 acquisition. H-epenthesis is a form

of qualitative hypercorrection (Janda and Auger, 1992), along the lines of intrusive-r in English (Halle and Idsardi, 1997; Orgun, 2002). To narrow our scope, we focus here on the phenomenon of h-deletion.

Our position is that, at its source, h-deletion is not due to low-level articulatory difficulty nor to a phonological process that removes underlying /h/ from surface forms (e.g., resulting from constraints on output as in the Emergence of the Unmarked: Broselow et al., 1998). Instead, h-deletion follows from perceptual and representational problems. Under the perceptual reorganization that accompanies first language (L1) acquisition (Strange and Shafer, 2008), second language (L2) learners "redeploy L1 phonological knowledge" (Archibald, 2005), typically perceiving and representing novel phonemes according to L1 categories (a process referred to as "perceptual assimilation" by Best and Tyler, 2007, and auditory "equivalence classification" by Flege and Bohn, 2021). Perceptual assimilation accounts for the widespread phenomenon of substitution in L2 speech. For example, Russian learners of English realize /h/ as the L1 voiceless velar fricative /x/, and Spanish learners, depending on their variant of L1 Spanish (i.e., variants without /h/), realize /h/ as velar /x/ or uvular /χ/; similarly, QFs realize English /θ ð/ as /t d/ (*think that* → [t]*ink* [d]*at*) (Brannen, 2011). The process of h-deletion, however, suggests that QFs do not assimilate English /h/ to an L1 category. Apparently, no L1 phoneme is sufficiently similar to /h/ for assimilation to take place; instead, QFs fail to detect /h/ in speech output and consequently leave the segment out of underlying representations (URs). As a result, /h/-vowel minimal pairs such as *hair-air* are represented as homophonous /ɛr/, and the error of h-deletion in fact constitutes an accurate realization of the stored form.

The situation is, however, more complicated than the above scenario implies. First, QFs typically exhibit variable h-production rather than categorical deletion in English, which is hard to reconcile with the absence of /h/ in URs. If /h/ is missing from lexical entries, how do learners generate [h] in output at all? Indeed, learners manage to generate higher rates of [h] in items that should contain it (correct h-production) vs. items that should not contain it (hypercorrect h-epenthesis) (John and Cardoso, 2009). This suggests that learners must somehow lexically mark the distinction between h-ful and h-less items. One possibility is that they develop approximate (i.e., "fuzzy," as in Darcy et al., 2013; or "murky," as in John, 2006) representations for /h/ using non-linguistic diacritics rather than actual distinctive features (John and Frasnelli, 2022). That is, the minimal pair *hair-air* may be distinguished via an *ad hoc* marking that we capture graphically with a superscript question mark: /$^{?}$ɛr/ vs. /ɛr/. Learners are thought to develop approximate representations when they become aware (e.g., due to feedback) that their output diverges from that of native speakers (NSs). That is, when learners recognize their own tendency to delete /h/ and yet are at a loss to match this elusive speech sound to a phoneme category, they compensate by marking h-ful items as requiring special implementation (and h-epenthesis, incidentally, would simply be an instance of over-application of this special implementation to h-less items). Likewise, in perception, approximate representations correspond to an auditory level of processing: while QFs may fail to perceive /h/ phonetically/phonemically, they can detect acoustic

differences between h-ful and h-less items (see the discussion below of Mah et al., 2016). Acoustic / auditory perception usually decays rapidly (Werker and Tees, 1984; Werker and Logan, 1985); approximate representations are an attempt to preserve these low-level perceptual distinctions. Such markings are not part of the toolkit supplied by Universal Grammar; instead they are add-ons introduced from outside the Language Faculty when normal phonological acquisition fails. It is not surprising then that they are less reliable than bona fide feature-based representations both in enabling h-perception and in cuing h-production. This explains the considerable variability in learner performance. In essence, the marking merely reflects auditory processing and signals how a form should be phonetically implemented.

It is worth noting that this representational view of L2 variation runs counter to established sociolinguistic accounts of L1 variation. These usually situate variation in the derivational grammar, whether due to variable rules (Cedergren and Sankoff, 1974), to shifts between categorical grammars (Kroch, 1989), or to partially ordered (Anttila, 1997) or overlapping constraints (Boersma, 1997); representational explanations are restricted to subsets of the lexicon (e.g., lexical exceptions in Guy, 2007). Studies of L2 variation have tended to adopt a similar derivational approach (Dickerson, 1975; Preston, 1989; Cardoso, 2007). Attributions of L2 variation to a phonological process, however, rely on the problematic assumption of accurate representation, as well as failing to account for the parallel phenomenon of variable perception. A lexical account of variation, based on approximate representations, avoids these problems. Second, two electroencephalography (EEG) studies show contradictory results regarding QF h-perception. On the one hand, an *unattended oddball paradigm* of auditory *linguistic* stimuli corresponding to syllables with and without [h] ([hʌm]-[ʌm]) failed to generate Mismatch Negativity [MMN—an event-related potential (ERP) associated with a deviant stimulus after a series of standard stimuli] among QF participants; only NSs exhibited this ERP (Mah et al., 2016). This finding suggests QFs neither perceive nor lexically record the distinction between h- and vowel-initial forms. Furthermore, *non-linguistic* noise burst stimuli as in [f], [hf], and [θf] elicited comparable MMN responses among NSs and QFs, which suggests that QFs have no problem with low-level processing of the acoustic properties of [h]. On the other hand, an *attended oddball paradigm* targeting [h] ([hɑ]-[ɑ]) led to MMN for both NSs and QFs (White et al., 2015). This finding suggests QFs are able to perceive, and potentially record in lexical entries, differences between h- and vowel-initial items. The nature of this possible representation is open to debate: it could be an actual /h/, an approximate representation as in John and Frasnelli (2022), or something else entirely, as we consider next.

The perceptual tasks in the two studies differ in whether participants attended to the input and in the quality of the vowel used in the stimuli: [ʌ] vs. [ɑ]. Mah et al. (2016) attribute the difference in perceptual accuracy not to the absence vs. presence of attention but to the phonetic realization of /h/ before the two vowels. In phonetic implementation, /h/ undergoes considerable contextual conditioning such that, particularly when preceded by a pause, it takes on the oral articulatory properties of the following vowel (Keating, 1988; Ladefoged and Maddieson, 1996). That is, /h/ is realized as a voiceless version of the subsequent vowel sound.

Before the low vowel [ɑ], [h] consequently resembles a voiceless uvular continuant. Since QFs have the devoiced uvular rhotic continuant [ʁ̥] in their repertoire (Tousignant, 1987; Walker, 2001; Sankoff and Blondeau, 2007), this might enable them to distinguish [hɑ] from [ɑ]: the tokens sound like [ʁ̥ɑ] vs. [ɑ]. This perceptual assimilation could explain the QF participants' MMN responses in White et al. (2015).

If Mah et al. (2016)'s explanation of the different findings holds, QFs assimilate instances of [h] before [ɑ] to the L1 category /ʁ/: it is for this reason that [h] in [hɑ] is detected rather than falling under the perceptual radar as in [hʌm]. By extension, instead of always omitting /h/, we might expect QFs to replace /h/ with /ʁ/ in URs containing the vowel /ɑ/. That is, just as QFs perceive [θ] as being an instance of /t/, thus storing the item *thank* as /tæŋk/, [h] before [ɑ] in *hot* would be heard and stored as /ʁɑt/. The subsequent realization of *hot* as [ʁ̥ɑt] could also sound sufficiently convincing to NS ears to pass as accurate. That is, perceptual assimilation can go both ways: NSs would interpret [ʁ̥] as /h/. Certainly, a listener would not process the output as an instance of h-deletion. To elaborate, by virtue of /t/ being a segment of English, NSs immediately recognize the substitution of [t] for /θ/; by virtue of the absence of /ʁ/ from the English inventory (typically, the English rhotic is alveolar /ɹ/ or retroflex /ɻ/; McMahon, 2002), however, NSs would not so readily detect the substitution of [ʁ̥] for /h/.

In sum, under Mah et al.'s (2016) proposal, we expect items with /ɑ/ (e.g., *hot, hop, hall, hard*) to be associated with higher rates of h-production than items with /ʌ/ (e.g., *hut, hug, hulk, honey*). That is, part of the variation observed in QF h-production may follow from how /h/ is at times included in URs (albeit as /ʁ/) and at others omitted or assigned an approximate representation. Just as Russian and Spanish learners may represent /h/ as /x/ or /χ/, francophone learners potentially represent a subset of items containing /h/ as /ʁ/, which should be reflected in production. The current study sought to establish whether QFs in fact show such variation in producing /h/ before /ɑ/ and /ʌ/. After the literature review in the next section, we outline the method used to test this prediction.

2 Background

In what follows, we consider the status and distribution of /h/ in both English and French and review previous work on QF perception and production of /h/.

2.1 On /h/ in English

In English, the voiceless glottal fricative /h/ is special in having only a single feature in its representation and a distribution limited to positions of prosodic prominence. Following arguments in Davis and Cho (2003), the representation for /h/ contains only [spread glottis], a laryngeal feature also associated with aspiration. The glottal fricative is thus unusual in having neither place of articulation nor manner features (Figure 1).

The restricted distribution of English /h/ runs parallel to that of aspirated stops. Barred from coda position, aspirated stops and /h/ appear only in word-initial onsets or, word-internally, in

Representation of /h/

o ROOT
LARYNGEAL
[spread glottis]

FIGURE 1
Representation of /h/.
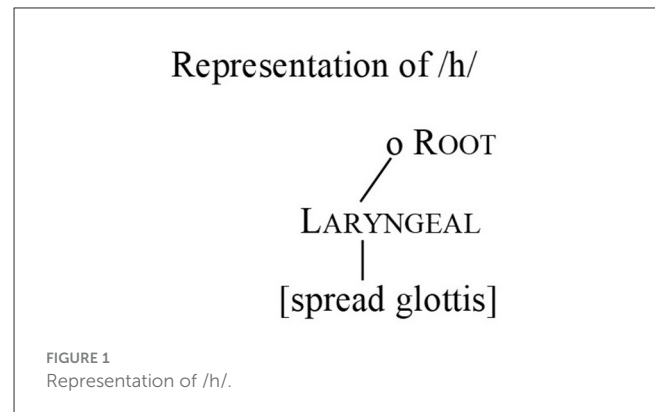
onsets of stressed syllables (Table 1). At the beginning of word-internal unstressed syllables, /pʰ tʰ kʰ/ are de-aspirated (compare *pre′*[pʰ]*are* and *pre*[p]*a′ration*) and /h/ is deleted (compare *pro′*[h]*ibit* and *pro*[_]*i′bition*). If we assume that, like /h/, aspiration is underlying in English (on this, see Harris, 1994), the phenomena of de-aspiration and h-deletion in prosodically weak contexts can be captured via a unified process delinking [spread glottis] (or even the entire laryngeal node: Lombardi, 1995). Function words also undergo optional delinking of [spread glottis] such that *give her* and *give to* surface variably with and without initial [h] in *her* and aspirated [tʰ] in *to* (variation is shown as √ ∼ X in Table 1).

In a further distributional limitation, English /h/ forms a branching onset only with the glide /j/ (e.g., *huge* =/hjudʒ/) and, in some varieties of English, with the glide /w/ (e.g., *where* =/hwɛr/). That is, unlike most obstruents, /h/ never combines with the liquids /r/ or /l/. One means of accounting for this restriction is to argue that, being a single-feature segment, /h/ does not have sufficient strength to license a dependent segment of greater complexity than a glide (Harris, 1997).

Finally, speakers of h-dropping varieties of English in Britain and Newfoundland routinely leave /h/ out, variably or categorically deleting it regardless of position in the word (Wells, 1982; Milroy, 1983). That is, /h/ alternates with absence in the output of speakers of these varieties, much as it does in the speech of QFs. Indeed, h-droppers who try to emulate h-ful speech sometimes produce epenthetic [h] (Häcker, 2002), just like francophone learners of English. We address the status of /h/ in French in the next section.

2.2 On /h/ in French

Absent from the French inventory, /h/ constitutes a new phoneme that QF learners of English need to acquire. In addition, no native phoneme is perceptually close to /h/: rather than being assimilated to an L1 phoneme, /h/ is generally not detected in input (LaCharité and Prévost, 1999; Melnik and Peperkamp, 2019). As a result, QFs seem initially to leave /h/ out of URs, although they may eventually construct an approximate representation employing *ad hoc* diacritics (John and Frasnelli, 2022). QFs also typically exhibit variable h-deletion, alternating between _*appy* and [h]*appy* as a realization of *happy*. All the same, it is not entirely clear why /h/ poses such a considerable challenge for francophones. As we discuss

TABLE 1 Distribution of [h] and aspiration in English.

	Word-initial		Word-internal		
	Onset (*content wd*)	Onset (*function wd*)	Onset (*stressed*)	Onset (*unstressed*)	Coda
[h]	√	√ ∼ X	√	X	X
[pʰ tʰ kʰ]	√	√ ∼ X	√	X	X

next, [h] can occur in French epenthetically, so it is not beyond learners' articulatory abilities. More significantly, [h] appears as an allophone of /ʃ/ in some varieties of Quebec French. This further confirms that physical production of the sound is unproblematic. Additionally, as discussed below, the occurrence of allophonic [h] suggests that the French phoneme inventory employs [spread glottis], the sole feature required to represent /h/. This view on the status of [spread glottis] in French is based on Harris's (1994) position that URs are fully specified, and phonological processes are limited to operations of spreading, delinking and default insertion. Since [spread glottis] is not an unmarked feature eligible for default insertion, we assume it is present in URs, emerging as allophonic [h] in a lenition process which delinks all other features. Arguably, the presence of [spread glottis] in L1 representations should make it easier to access for representing L2 phonemes such as /h/. This follows from Brown's (1998) position that L2 phonemes *per se* are not problematic for acquisition, only L2 distinctive features (i.e., features not employed in L1 representations). Admittedly, Brown's proposal is made within the model of Minimally Contrastive Underspecification (Avery and Rice, 1989), which postulates that only those features required to establish contrasts are specified in URs. In this model, since [spread glottis] is not employed contrastively in French, the feature may be left out of URs. However, if we omit [spread glottis] from French representations, the lenition process which generates [h] in Quebec French is hard to account for. This brings us back to the question of why QFs cannot readily access [spread glottis] to represent English /h/. To resolve the conundrum, we propose the presence in the phonological component of a constraint on phonemic representations that are exclusively composed of laryngeal features. This Laryngeal Input Constraint makes /h/ a marked phoneme and accounts for why it is so problematic for francophones.

The sound [h] is at times realized in interjections in French— Walker (2001) mentions *hop!* and *hein?*—and even before vowel-initial content words instead of a glottal stop. While, as in many languages (Lombardi, 2002), glottal stops are the preferred epenthetic consonant in French (notably for marking *h-aspiré* words: Gabriel and Meisenburg, 2009), an epenthetic [h] can also occur.[1] Although epenthetic [h] may be more common in song (see footnote 1), its occurrence even there suggests that the francophone

problem with English /h/ is not superficially phonetic, related to articulatory difficulty; francophones are physically capable of producing the sound. Instead, /h/ appears to constitute a phonological problem for L2 learners of English, but the precise nature of the problem is hard to pinpoint.

Arguably, instances of epenthetic [h] are merely added during phonetic implementation. That is, they may be like excrescent stops in nasal-fricative sequences such as *prin*[t]*ce* or *Chom*[p]*sky* in English, which seem to be articulatory effects generated phonetically rather than phonologically (Ohala and Ohala, 1993; Feldscher and Durvasula, 2017). It is of course difficult to determine definitively whether a speech phenomenon has its source in the phonological system or in phonetic implementation. Categorical phenomena tend to be phonological, whereas variable / optional phenomena could be either. For example, epenthetic glottal stops (a variable phenomenon in English and many other languages) may result from a phonological preference for syllables with onsets (e.g., the Onset constraint in Prince and Smolensky, 1993/2004), or they could simply reflect ease of articulation, emerging like excrescent stops during phonetic implementation. The same applies to epenthetic [h] in French. Under the analysis of epenthetic [h] as a product of the phonetic system, its occurrence in French would not necessarily aid in the perception and eventual phonological acquisition of /h/.

When [h] is generated as an allophone of an underlying segment, however, this is clearly a phonological process, which can conceivably facilitate acquisition of underlying /h/. Some varieties of Quebec French instantiate a process of debuccalization, whereby /ʃ/ is reduced to [h] in onset position: *chocolat chaud,* for example, is realized as [h]*ocolat* [h]*aud* (Bittner, 1995; Paradis and LaCharité, 2001; Morin, 2002). Similar processes are observed in other languages, [h] being in Spanish a well-attested product of lenition of coda /s/ and a pre-deletion segment: /s/ → [h] → Ø (File-Muriel and Brown, 2010; Núñez-Méndez, 2022). Brazilian Portuguese instantiates a similar pattern for coda rhotics: /ʁ/ → [h] → Ø (Rennicke, 2015). Interestingly, QFs seem to readily process [h] as an allophone of /ʃ/ even if this pronunciation does not occur in their own variety. There is no evidence that QFs without this allophonic process struggle to understand speakers who do realize /ʃ/ as [h]; indeed, although to our knowledge the matter has not been formally investigated, informal observations suggest that non-debuccalizing QFs are capable of imitating speakers who realize /ʃ/ as [h]. Again, articulation of /h/ is resolutely not the problem.

Under the view that lenition involves feature loss (Harris, 1990, 1994, 1997; Honeybone, 2008), debuccalization consists of suppression of all other features in the representation except [spread glottis]. The process can be captured in feature geometry (e.g., Clements, 1985) via delinking of the supralaryngeal node, leaving only the laryngeal node with [spread glottis] as dependent

---

1   An example of [h] in an interjection occurs in commercials for *Familiprix* where *aha* is unmistakably realized as [aha]: https://www.youtube.com/watch?v=aURNJoNgUIQ. An epenthetic [h] before a vowel-initial form occurs in the introductory song for the animated series *Wakfu*, where *héros malgré lui* is realized as [h]*éros*: https://www.youtube.com/watch?v=_7TvSNdgKik. Likewise, Edith Piaf clearly sings [h]*Allez, venez, Milord!* on two occasions in the following version of her well-known song: https://www.youtube.com/watch?v=nZcdI1u_9o8.

feature. The cross-linguistic occurrence of [h] as a product of lenition, as well as its tendency to alternate with zero, is thus consistent with the view of /h/ as a single-feature phoneme. It also indicates that [spread glottis] must be part of the underlying representation for Spanish /s/, Brazilian Portuguese /ʁ/ and Quebec French /ʃ/, even though none of these languages have /h/ in their phoneme inventory. This is hard to reconcile with Brown's (1998) claim that L2 phonemes using distinctive features employed in L1 representations should be relatively easy to acquire. If QFs already require [spread glottis] in their L1 (albeit as a non-contrastive feature), acquisition of /h/ should be relatively straightforward. However, production data from previous studies suggest that acquiring /h/ is highly challenging: despite extensive English-language studies (mean: 12.06 years), only 12 of the 50 QF participants in John and Frasnelli (2022) showed no h-deletion even in a limited reading-aloud task. QFs who perform on a par with NSs of English in the perception and production of /h/ are thus the exception.

Despite [spread glottis] potentially being available to QFs from their L1, learners have difficulty acquiring a phonemic representation containing only this distinctive feature. Possibly something other than the absence of the feature from the L1 is at work to block the acquisition of phonemic /h/. We propose the existence of an input constraint. Phonological constraints in recent decades have been construed in Optimality Theory (OT; Prince and Smolensky, 1993/2004; see also McCarthy, 2002) as applying purely to output. Although, under Lexicon Optimization, input representations generally reflect surface forms, the principle of Richness of the Base (Smolensky, 1996; Davidson et al., 2004) considers that input forms are entirely unconstrained (for challenges to Richness of the Base, cf. Vaysman, 2002; Gouskova, 2023). In theory, this means that L2 learners should have no problem developing URs that contain novel segments, including highly marked phonemes; their only challenge should be to re-rank markedness constraints such that the segments can emerge in output.

Conceivably, however, the phonological component incorporates input constraints that make certain URs dispreferred (i.e., a Restriction on the Base). For example, a constraint on underlying segments comprised exclusively of laryngeal features would favor phoneme inventories that lack /h/; it would likewise exclude /ʔ/, glottal stops being comprised solely of the feature [constricted glottis]. A [spread glottis] Laryngeal Input Constraint would account for the absence of /h/ from the phoneme inventories of Spanish, Brazilian Portuguese and Quebec French even though [h] appears in the output of these languages. The same case can be made for languages such as English that lack underlying /ʔ/ but employ glottal stops as a surface allophone: they could contain a [constricted glottis] Laryngeal Input Constraint. Interestingly, the proposed constraint contradicts the claim that glottal is an unmarked feature. For example, according to Lombardi (2002) and de Lacy (2006), glottal constitutes the least marked place of articulation. Our position is that, while this unmarked status may hold at the surface, it seems not to apply underlyingly.

While diverging from OT-based output constraints, the notion of restrictions on input recalls Morpheme Structure Constraints. For example, the Obligatory Contour Principle (OCP) bars the presence in a UR of adjacent identical features (e.g., in the initial conception, tones: Leben, 1973). Though the OCP was later expanded to apply to output, triggering and blocking phonological derivations (McCarthy, 1986; Yip, 1988), it was originally conceived of as a constraint on URs much like the Laryngeal Input Constraint. The OCP has also been depicted as a soft constraint that is not always respected (Odden, 1986); underlying structures that violate the OCP are avoided since more marked, but not strictly ruled out. The same seems to be the case with the Laryngeal Input Constraint: while /h/ and /ʔ/ are dispreferred, phoneme inventories can nonetheless contain these laryngeal segments.

The distinction between phonological constraints on output (as in OT) and input (as proposed here) is intriguing insofar as the content of the two constraint types may be contradictory. Output constraints target either *faithfulness* ("output is identical to input") or *markedness* ("output is less marked than input"). Consequently, output that diverges from input is necessarily less marked. Following this argument, [h] should be an unmarked segment, as it is an allophone of underlying /ʃ/, /s/, and /ʁ/ in Quebec French, Spanish, and Brazilian Portuguese. According to the Laryngeal Input Constraint, however, /h/ and /ʔ/ are marked structures in URs. Apparently, the markedness status of a segment can differ between underlying and surface levels of representation.

Concentrating on surface realizations of laryngeal segments, an interesting parallel can be made between the glottal fricative [h] in QF and the glottal stop [ʔ] in English, both of which have epenthetic as well as allophonic status. As mentioned above, QFs arguably process epenthetic [h] in French as a purely phonetic effect, filtering it out as a linguistically irrelevant segment unrelated to any phoneme. In this sense, instances of epenthetic [h] (in *hop!* and *hein?* and elsewhere) are processed in similar fashion to inserted glottal stops in English (Garellek, 2012). In our experience (e.g., the first author's, as a NS of English), anglophones have difficulty detecting epenthetic [ʔ] in the speech signal. To notice the presence of glottal stops, listeners have to attune their ears to a phonetic level of processing, which degrades rapidly under normal speech perception (Werker and Tees, 1984; Werker and Logan, 1985). Our impression is that neither epenthetic [h] in French nor epenthetic [ʔ] in English is particularly salient since neither corresponds to an underlying segment. Consequently, the presence of epenthetic [h] and [ʔ] in the L1 would not necessarily aid in the acquisition of phonemic /h/ and /ʔ/ in an L2.

However, like [h] in Quebec French, [ʔ] in some varieties of English can also be an allophone of an underlying segment. That is, one and the same form in output can have either epenthetic or allophonic status. Just as [h] can be an allophone of /ʃ/ in Quebec French, intervocalic /t/ can be realized as [ʔ] in British English as in *butter* → *bu*[ʔ]*er* (Harris and Kaye, 1990). Interestingly, while English speakers are largely unaware of epenthetic [ʔ] (i.e., it passes under the perceptual radar), allophonic [ʔ] is readily perceived. While processing allophonic [ʔ] is effortless and automatic, detecting epenthetic [ʔ] requires concentration on the signal. Despite epenthetic vs. allophonic [h] and [ʔ] having comparable phonetic properties, the segments are perceived quite differently according to their status. When [h] and [ʔ] are epenthetic, they tend not to be noticed; but when [h] and [ʔ] are allophonic, they are readily perceived.

This distinction, albeit speculative, strikes us as insightful for how QFs perceive and represent /h/ in English. As the studies

reviewed in the next section show, it seems that QFs often fail to detect English [h] in speech, leaving it out of URs or at some point developing an approximate representation that allows for variable detection of this sound. In this sense, English [h] is processed like epenthetic [h] in French or epenthetic [ʔ] in English. Nonetheless, not all instances of [h] in English are necessarily equal: depending on the type of adjacent vowel, which affects how [h] is articulated, [h] may have phonetic properties that allow QFs to perceive it as a surface form of French /ʁ/. In this case, input [h] is associated to an underlying segment, much as [ʔ] in English at times corresponds to underlying /t/. It is just that this underlying segment is /ʁ/ rather than /h/. While the latter representation, according to the Laryngeal Input Constraint, is hard to acquire, QFs will easily construct the former representation, given that the phoneme is already available in the L1 inventory.

2.3 QF perception and production of English /h/

The QF tendency to delete /h/ in English conceivably derives from a difficulty in distinguishing h-initial and vowel-initial forms. From a bottom-up perspective, if QFs cannot hear the difference between *heat* and *eat* any more than an anglophone distinguishes a realization of *eat* with or without a glottal stop, they will not record any distinction in URs. The minimal pair will consequently be stored as homophonous /it/. From a top-down perspective, the presence of a particular phoneme category in a listener's inventory leads to automatic detection of instances of the category—this is what makes speech comprehension so effortless. QF learners of English are disadvantaged in not having /h/ in their phoneme arsenal, thus impeding their ability to detect the speech sound and to record it in URs. As we see below, earlier studies on the perception and production of /h/ by francophones have findings either consistent with this "defective UR" account or pointing to the need for a more nuanced view.

The ERP findings in Mah et al. (2016) largely support the notion that QFs fail to record /h/ in URs. In an unattended oddball paradigm using both linguistic and non-linguistic auditory stimuli, participants listened to a series of repeated stimuli (standards) interspersed with different stimuli (deviants). Detection of a deviant in a stream of standards is associated with the ERP MMN. Both QFs and NSs of English exhibited MMN with *non-linguistic* noise burst stimuli ([f], [hf], [θf]), which suggests that QFs have no trouble with auditory detection of [h]. Only NSs exhibited MMN, however, with *linguistic* stimuli containing [h] ([ʌm], [hʌm]). This suggests that QFs fail to perceive [h] when processing the signal as speech, consistent with their leaving /h/ out of URs.

Nonetheless, there is a mismatch between what the perception data indicate about URs (apparently inaccurate) and the production data (surprisingly accurate): in a reading-aloud task, more than half the participants realized all 8 tokens of /h/. Mah et al. attribute this production accuracy to the influence of orthographic cues that guide the realization of h-initial forms. While previous studies have found some evidence that exposure to orthographic forms can scaffold acquisition of confusable phoneme contrasts (e.g., in a non-word learning task with Dutch participants targeting the /ɛ/-/æ/ contrast in English: Escudero et al., 2008), the role of orthography

in guiding speech production is far from clear. Indeed, QFs were found to produce more instances of h-epenthesis in reading aloud than in spontaneous speech (John and Cardoso, 2009); that is, despite the evidence for h-initial vs. vowel-initial forms being in full view, orthography failed to promote more accurate output. We are thus skeptical of claims that learners can reliably use written forms to guide production.

It is not that orthography plays no role in L2 phonological acquisition (see Bassetti et al., 2015; Hayes-Harb and Barrios, 2021, for overviews). For example, when learners transfer grapheme-phoneme correspondences from the L1 to the L2, this may interfere with production (e.g., Rafat, 2016). Indeed, the grapheme < h > does not correspond to any speech sound in French, as shown in homophones such as *aine-haine* /ɛn/ ("groin"-"hate"), *eau-haut* /o/ ("water"-"high"), and *ache-hache* /aʃ/ ("wild celery"- "axe"). Certainly, French contains orthographically h-initial items referred to as *h-aspiré* that act as though they are consonant-initial (Charette, 1991; Tranel, 1995). These forms block linking processes such as *liaison* that supply an onset to otherwise onsetless forms. Nonetheless, the beginning of an *h-aspiré* item does not itself correspond to an actual speech sound. Consequently, QFs are used to processing < h > as a silent letter. To complicate matters, < h > is not even a reliable indicator of /h/ in English since it remains unpronounced in a few high-frequency words (e.g., *hour, honor, honest*). A recent pilot project indicates that the inconsistency of the grapheme-phoneme correspondences for English /h/ could have a confounding effect on the development of accurate lexical representations (Jackson and Cardoso, 2023). In brief, the fact that learners do not associate < h > with any speech sound in French (i.e., unlike letters such as < j > or < s >) and that < h > is not always pronounced in English makes the grapheme more likely to generate h-deletion than to promote h-production.

It is also not clear that QFs always fail to perceive the distinction between h-initial and vowel-initial forms. While the results of the *unattended oddball paradigm* in Mah et al. (2016) point in this direction, the findings in White et al. (2015) are not consistent with this view. In an *attended oddball paradigm*, the latter researchers found similar MMN responses among QFs and NSs with auditory stimuli using the syllables [hɑ] and [ɑ]. This suggests that QFs are able to perceive [h] under certain conditions, whether they can record /h/ in URs or not.

Possibly, the difference in the findings of the two EEG studies is due to the presence or absence of attention; that is, QFs can perceive the [h]-Ø contrast as long as they are attending to the phonetic input. In our view, approximate representations certainly require special effort for learners to draw on them in perception and production. If such *ad hoc* markings are associated with /h/, it comes as no surprise for attention to assist perception. Mah et al. (2016), however, intriguingly attribute the difference to the vowel type used in the two studies: [ɑ] vs. [ʌ]. From our perspective, their position deserves further consideration and investigation. The absence of phonological place features means /h/ undergoes considerable contextual conditioning during phonetic implementation. The low vowel [ɑ] influences the realization of /h/, creating an acoustic effect that resembles a voiceless uvular continuant (Keating, 1988; Ladefoged and Maddieson, 1996). Since QFs often devoice the uvular rhotic continuant /ʁ/ in their inventory, such that /ʁ/→ [χ] (Tousignant, 1987; Walker, 2001; Sankoff and Blondeau, 2007), this could facilitate their ability to

distinguish [hɑ] from [ɑ]. For QFs, this pair conceivably sounds like [ʁɑ]-[ɑ], a contrast that is easy for them to process, unlike the [hʌm]-[ʌm] contrast.

By extension, if QFs assimilate [h] to /ʁ/ before [ɑ], they should also substitute /ʁ/ for /h/ in URs of items containing /ɑ/. Consequently, while an item such as *hut* might be stored as /ʌt/ (/h/ omitted), an item such as *hot* would be stored as /ʁɑt/ (/ʁ/ substituted for /h/). If /ʁɑt/ is then realized as [ʁɑt], it could strike NSs' ears as sufficiently close to /h/; certainly listeners would not have the impression /h/ has been deleted. Indeed, when Spanish or Russian speakers substitute [x] for /h/, NSs of English automatically classify the input as a realization of /h/, even if it sounds phonetically unconventional. Essentially, NSs themselves assimilate [x], a sound missing from their inventory, to the closest L1 phoneme, namely /h/. Similarly, NSs could process the QF realization [ʁ] as /h/. This is not the only possibility, however, for how QFs might represent /h/; they may instead develop an approximate representation using *ad hoc* diacritics, as we consider next.

While the two EEG studies show either presence or absence of MMN, consistent with ability/inability to perceive and possibly represent /h/, John and Frasnelli (2022) found highly variable QF perception and production of /h/; they also observed considerable inter-participant variation. Indeed, one of the advantages of behavioral over ERP data is that precisely this kind of variation is easier to discern. Because ERP responses are quite subtle, extensive data are required from numerous participants for patterns to emerge. Consequently, while differences between larger groups (e.g., QFs vs. NSs) are detectable, differences between individuals are typically lost. The intra- and inter-participant variation observed in John and Frasnelli (2022) is consistent with varying degrees of gradient perception across learners rather than simple ability/inability to perceive /h/. Perception was tested via an attended oddball paradigm task with trials where the fourth item was either the same or different from the preceding three. Stimuli were mono- or disyllabic real and non-words involving a variety of vowel sounds (e.g., *heat-heat-heat-eat, old-old-old-hold, hice-hice-hice-ice, enk-enk-enk-henk*), although the analysis did not include vowel quality as a condition. At the end of each trial, participants indicated via keyboard press whether the final item was the *same* or *different*. For the condition targeting /h/, QFs showed lower mean accuracy rates than NSs (64 vs. 96%) as well as considerably wider ranges (0.8–100% vs. 91–100%), with individual QF participant rates distributed evenly across the broad range. That is, QFs did not only show either poor or nativelike perception but everything in-between. QF h-production rates also covered a wide range (23.81–100%), and accuracy rates in perception and production were highly correlated.

QFs thus generally struggle to perceive and produce /h/ but nonetheless show gradient differences between individuals. Some QFs exhibit very poor perception and production, consistent with failure to record /h/ in URs; others perform on a par with NSs, consistent with having acquired /h/. Most, however, perform somewhere between these two poles, a distribution consistent with their having developed some distinction between h- and vowel-initial forms but falling short of full acquisition of phonemic /h/. Instead, John and Frasnelli argue for approximate ("fuzzy" or "murky") representations. These are indicated in URs with a

diacritic such as a superscript question mark that reflects the murky status of the representation. Items such as *hut* and *hot* are thus presumed to be stored as /$^{?}$ʌt/ and /$^{?}$ɑt/. It may, however, be more appropriate to think of these markings as separate from actual phonological representations. Similar to orthographic information associated with a lexical entry, approximate representations may constitute extra-phonological add-ons. These reflect a special phonetic quality that is both detectable when the input is attended to and reproducible when output is formulated with sufficient control and effort. Unlike representations employing features supplied by Universal Grammar, approximate representations would allow for only variable perception and production of /h/. As a function of experience, learners should show improvement in performance, thus accounting for the wide distribution in their perceptual and productive abilities.

Nonetheless, not all of the variation in h-production and perception observed among QFs is necessarily due to approximate representation of /h/; some variation may be due to assimilation of /h/ before /ɑ/ to /ʁ/. That is, as well as representing an item such as *hut* as /$^{?}$ʌt/, QFs may represent an item such as *hot* as /ʁɑt/, substituting /ʁ/ for /h/. Although inaccurate in terms of the target, the phonemic representation /ʁ/ should permit QFs to distinguish consistently between h- and vowel-initial items in perception and production, but only h-initial items containing /ɑ/. Additionally, while development of an approximate representation constitutes a strategy for getting round the Laryngeal Input Constraint, assimilation of /h/ to /ʁ/ means the constraint does not even apply to the input.

2.4 Research question and hypothesis

In the current study, we test the prediction that QFs assimilate /h/ before /ɑ/ to /ʁ/ and thus produce lower rates of h-deletion in items such as *hot* than *hut.* Our aim is to answer the following research question and verify the hypothesis given below.

Research question: Do QFs differ in their production of /h/ before /ɑ/ and /ʌ/? Hypothesis: QFs will show high accuracy in the production of /h/ in items where the following vowel is /ɑ/, but only variable production of /h/ in items where the following vowel is /ʌ/.

More graphically, h-production should show:

i) Accuracy: /hɑt/ > /hʌt/ (where the symbol ">" means "greater accuracy than").
ii) Variation: /hʌt/ > /hɑt/ (where the symbol ">" means "greater variation than").

The method used to test our hypotheses is outlined next.

3 Materials and methods

3.1 Participants

A total of 34 QFs (9_{males}, M age = 28.00 yrs, range = 18–57; $25_{females}$, M age = 30.79 yrs, range = 21–53) were recruited mainly among the student bodies of francophone universities in Quebec, a majority French-speaking region of Canada. For context, we should explain that French is the sole official language in Quebec, although

anglophones constitute a significant minority in certain regions, notably in Montreal, the largest city in the province. School boards in Quebec are divided along linguistic lines, with francophones attending schools run by the French school board, where English language instruction is usually introduced in the later years of primary school. Consistent with this situation, language background questionnaires administered before the production task established that the participants generally started learning English in a classroom setting from an early age (M age = 9.19 yrs; range = 5–18 yrs) and for an extended duration (M = 12.93 yrs). It should be noted that the questionnaires revealed considerable variation in both degree and type of exposure, as well as in age of initial exposure, such that it would be difficult to investigate any correlation between age or degree of exposure and production accuracy. Hypothetically, we might anticipate a divide between participants in Montreal and outlying regions, since the potential for contact with NSs is greater in Montreal. In practice, however, contact and English language use in Montreal can vary wildly, from virtually none to occasional or sustained contact, whether with neighbors, friends or colleagues/customers, and this can change considerably from one period in a person's life to another. Conversely, residing in a region with few anglophones does not preclude contact in diverse settings such as the workplace, foreign travel, immersion exchanges, and online environments (e.g., input from Netfix shows, virtual exchanges in gaming). Likewise, while we might presume participants who started learning English in the classroom from age 5 or 6 would have an advantage, such early exposure in the Quebec context is typically limited to 1 h per week, in which case the amount of exposure is too limited to provide an edge in acquisition. In brief, in the face of such diversity of experience, we did not attempt to establish correlations with h-production.

3.2 Materials

In an online environment using Zoom, participants were recorded reading aloud a series of 36 expressions (e.g., *some **hot** apple pie*; *a big **hug***) and 25 sentences (e.g., *True love is **hard** to find; She lives in a mud **hut***) presented one-by-one on PowerPoint slides. In all, the task contained 80 target items containing /h/ followed by either /ɑ/ (*hot, hard*) or /ʌ/ (*hug, hut*)—see Appendix A for the full list of expressions and sentences. Equal numbers of target items with /ɑ/ and /ʌ/ appeared in both the phrases and sentences.

3.3 Data analysis

The recordings were coded impressionistically by a NS of English, who indicated whether /h/ was deleted or preserved in the target item, as well as noting instances of h-epenthesis, although these were not included in the actual analysis. If the rater judged that /h/ was preserved in a given item, this was coded as an instance of accurate h-production. Using R software for statistical computing (R Core Team, 2021), we carried out two-way ANOVAs with accuracy of h-production as dependent variable and vowel

TABLE 2 QF accuracy rates for h-production (%).

		Descriptive statistics		
Independent variables		Mean	Std deviation	Range
Vowel type	/ɑ/	76.82	29.07	0-100
	/ʌ/	77.12	26.89	0-100
Stimulus type	Phrases	77.73	26.57	5-100
	Sentences	76.21	31.46	0-100

type (/ɑ/ vs. /ʌ/) and stimulus type (phrase vs. sentence) as intra-participant independent variables.

Effects for vowel type were anticipated, with items containing /ɑ/ (*hot*) expected to show higher rates of accuracy than items containing /ʌ/ (*hut*). Although the "stimulus type" variable was included more for exploratory purposes, we might also expect higher rates of h-production to be associated with phrase stimuli, given that short phrases are less cognitively challenging to process and articulate than full sentences. This effect is particularly anticipated if /h/ is captured via approximate representations in the QF lexicon, since such representations are thought to entail greater effort than true phonemic representations.

Since classical ANOVAs cannot be used to support the null hypothesis (absence of difference in h-production), we subsequently ran Bayesian t-tests to directly compare the amount of evidence in favor of the null hypothesis for both vowel type and stimulus type. These analyses yielded a Bayes Factor (BF_{10}) that corresponds to the ratio of evidence in favor of the alternate model (i.e., where there is a difference as a function of the parameter considered) vs. the evidence in favor of null model (where there is no difference). Specifically, BF values >1 indicate the strength of evidence in favor of the alternate model or, should BF be <1, the lower the value the stronger the evidence in favor of the null hypothesis (Rouder et al., 2009, 2012). Benchmark scores are: BF_{10} between 1 and 1/3 are considered weak (barely worth mentioning); between 1/3 and 1/10, they are considered substantial; and <1/10, they are considered strong evidence in favor of the null hypothesis (Jeffreys, 1961). The higher this value, the greater the evidence in favor of the alternative hypothesis, with benchmarks BF_{10} between 1 and 3 considered weak evidence and between 3 and 10 substantial (Jeffreys, 1961).

4 Results

Table 2 shows the mean accuracy rates in h-production by the 34 QF participants (2,720 tokens in all) according to the key independent variable of vowel type (/ɑ/ vs. /ʌ/) and the further variable of stimulus type (phrase vs. sentence).

It should be mentioned that considerable inter-participant variation was observed, with the accuracy rates for h-production of individual participants distributed across a range from near-categorically inaccurate to categorically accurate (2.5–100%). We can also report that 13 participants produced between 1 and 4 instances of h-epenthesis (e.g., *a broken* [h]*arm, the front* [h]*office*). For the purposes of the two-way ANOVAs and *t*-tests, the data

from the 7 categorically accurate h-producers were removed, since the purpose is to analyze variation. Note that including these 7 participants did not change the results of the ANOVA and Bayesian *t*-tests indicated next.

Classical ANOVAs run on participant accuracy rates could not reveal any effect for vowel type ($F_{[1,26]} = 1.00$, $p = 0.327$, $\varepsilon^2 = 0.001$) and stimulus type ($F_{[1,26]} = 0.17$, $p = 0.685$, $\varepsilon^2 < 0.001$), nor any interaction between these factors ($F_{[1,26]} = 1.35$, $p = 0.256$, $\varepsilon^2 = 0.003$). The results suggest that we cannot rule out the null hypothesis (i.e., an absence of difference in accuracy rates related to these factors). Paired Bayesian *t*-tests were run to compare the accuracy rates of participants as a function of vowel type and of sentence type. Following previously indicated benchmarks, the Bayesian paired *t*-tests provide moderate evidence for the null hypotheses, that is, that vowel type and sentence type had no influence on accuracy ($BF_{10} = 0.220$ and $BF_{10} = 0.236$, respectively).

5 Discussion

While QFs typically struggle to detect [h] in input, Mah et al. (2016) proposed that the phonetic quality of [h] before the vowel [ɑ] leads QFs to hear [h] as [ʁ], a common realization of the L1 segment /ʁ/. This would explain why, in two EEG studies using an oddball paradigm to target the ERP MMN, White et al. (2015) detected MMN responses among QF participants, whereas Mah et al. did not: the former study employed stimuli containing [ɑ], and the latter stimuli containing [ʌ]. We extrapolated that the presumed perceptual assimilation that facilitates h-perception before [ɑ] should lead QFs to represent /h/ as /ʁ/, but only for items where the following vowel is /ɑ/ (*hot*) and not /ʌ/ (*hut*). If such is the case, the distinction should be clearly reflected in h-production: QFs should show decidedly higher rates of h-production for items containing /ɑ/ than /ʌ/. Indeed, we expected QFs to show only variable h-production for items such as *hut,* represented in our view as /ʔʌt/ (John and Frasnelli, 2022). The approximate representation of /h/ via a superscript question mark (more properly an add-on external to the actual phonological representation) is what permits QFs, despite the absence of /h/ from their segmental repertoire, to distinguish h-initial from vowel-initial items such as *heat-eat* in their lexicon. The lexical distinction leads QFs to realize higher rates of h-production in actual h-initial items than hypercorrect h-epenthesis in vowel-initial items (John, 2006; John and Cardoso, 2009). In brief, according to our hypothesis, items such as *hot,* by virtue of /h/ being replaced with /ʁ/, should exhibit essentially categorical h-production, whereas items such as *hut,* with murkily specified /h/, should only exhibit variable h-production.

With data from a reading-aloud task involving 27 QFs, however, two-way ANOVAS and paired Bayesian *t*-tests revealed no difference in the realization of /h/ before /ɑ/ vs. /ʌ/ (vowel type) nor in phrases vs. sentences (stimulus type). QFs showed comparably variable h-production regardless of vowel, and the hypothesized greater accuracy in h-production for *hot* vs. *hut* failed to materialize. Instead, the null hypothesis (i.e., vowel type has no effect on h-production) was confirmed. This suggests that /h/ has the same representation for QFs regardless of the adjacent vowel, and the variation found in QF h-production in no way derives from

learners' at times assimilating /h/ to /ʁ/. In this case, the fact that MMN was observed among QFs in the oddball paradigm task in White et al. but not Mah et al. remains to be explained. We suggest that the design difference involving presence/absence of attention is responsible: in the former study, participants attended to the stimuli, whereas in the latter, they did not. Interestingly, attention is not an absolute requirement for h-perception, since MMN was observed among NS participants in Mah et al. It seems only QF learners of English need to attend to the signal in order for /h/ to be detected in the linguistic input.

In our view, the need for attention in order for QFs but not NSs to detect /h/ is consistent with a difference in the representation of this segment: unlike NSs, QFs do not record in URs an actual phonemic representation for /h/, involving the laryngeal feature [spread glottis]; instead, they use an approximate representation that bypasses distinctive features. Such *ad hoc* markings are less reliable and require greater effort than representations using UG-based features in supporting speech perception and production. Although we might have expected the association between approximate representations and effort/attention to result in greater accuracy in short phrases than full sentences, this variable failed to emerge as an influence on h-production. Nonetheless, considerable differences are consistently encountered between QFs and NSs in h-perception and production. In addition, QFs exhibit wide ranges in performance (e.g., John and Frasnelli, 2022): while learners initially have difficulty using approximate representations to support perception and production of /h/, over time they get better at the task, such that performances may rival those of NSs, with their actual feature-based representation of /h/.

Conceivably, the reading-aloud task, compared with spontaneous speech, may be particularly conducive to the kind of focused attention required for learners to draw on approximate representations as a cue for h-production. QFs are acutely aware of their difficulties with English /h/, so part of their success in h-production may be due to the task facilitating efforts to control articulatory behavior. Attention involves concentrated awareness directed toward a particular input, as with White et al.'s (2015) attended oddball paradigm, or output, as with our reading-aloud task (for a review of the concept of attention, see Lindsay, 2020). It may be that the 7 participants who showed categorical h-production were particularly skilled at the kind of heightened vigilance required to produce a speech sound that lacks a phonemic representation in their lexicon.

That QFs generally fail to develop an accurate representation for /h/ is unusual, given that [spread glottis], as we argued in the Background section, is present in L1 representations. Although [spread glottis] is not a contrastive feature in French, we considered, contra Brown (1998), that it should be available for developing L2 representations, making /h/ relatively easy to acquire. Since QFs have considerable difficulty acquiring /h/, we suggested that the phonological component contains a Laryngeal Input Constraint that renders highly marked any representations based exclusively on laryngeal features. Such a proposal runs counter to the OT view that constraints apply only to output, URs being entirely unconstrained. The principle of Richness of the Base (Smolensky, 1996; Davidson et al., 2004) seems to be disproved by L2 phenomena such as QF acquisition of /h/; the base itself is apparently subject to markedness constraints. When it

comes to /h/, QFs circumvent these by constructing an approximate representation that employs a non-feature-based diacritic. The proposal is important because it provides us with a means of characterizing certain challenges in L2 phonological acquisition. These are not problems of phonological or phonetic output nor even perceptual problems *per se*, but problems of underlying representation. Approximate representations also permit us to understand the considerable variation that characterizes L2s: variation follows directly from the nature of the *ad hoc* status of the representation.

It remains an open question as to whether some QF learners of English overcome the Laryngeal Input Constraint and eventually develop an accurate representation for /h/. The presence of 7 participants with categorically accurate h-production points to the possibility that these learners have in fact acquired /h/. Nonetheless, it is also possible that, given a more extended task or a task less conducive to attention, these highly proficient h-producers would have eventually slipped up and exhibited occasional instances of h-deletion. More intriguingly, another possibility is that some learners become so adept at drawing on approximate representations to ensure h-production that their performances are indistinguishable from those of NSs, despite the differences in representation of /h/. As such, some learners may constitute what John and Frasnelli (2022) refer to as "phonological zombies." The term "zombie" is borrowed from debates on the philosophy of mind (Chalmers, 1996) and is in no way intended to disparage L2 learners. The point of the zombie concept is that, while we each have privileged access to our own internal worlds, including subjective thoughts and feelings that confirm that we are personally in possession of consciousness, we can never be entirely sure about those around us. Despite exhibiting behavior consistent with a similar inner life, others may not experience consciousness exactly as we do. Indeed, we cannot be sure others are endowed with consciousness at all: we may be surrounded by zombies who only show the outward signs of consciousness. By extension, some L2 learners may exhibit an ability to perceive and produce /h/ that is consistent with their having acquired an accurate representation, but we cannot be sure that this is the case: they may be phonological zombies who perform on a level with NSs, despite /h/ being captured in their lexicon by approximate rather than phonemic representations. The challenge for future research is to design an experiment able to distinguish between phonological zombies and L2 learners who have in fact acquired an accurate representation for /h/ or other L2 sounds.

Finally, the current study contains certain limitations. Notably, while our analysis of the reading-aloud data investigates the influence of the following vowel on h-production, we did not include an examination of preceding sounds. As can be seen in Appendix A, items beginning with /h/ were preceded by different consonant or vowel sounds and at times by no sound at all. Since the preceding phonetic environment has been shown to influence rates of h-epenthesis (John and Cardoso, 2009), it remains possible that this variable also affects h-production. Furthermore, the information participants provided in the Language Background Questionnaire did not permit us to develop a clear portrait of age of acquisition and degree of exposure to L2 English. For example, although some participants reported starting English instruction as early as 5 or 6 years old, the exposure was

not necessarily sufficient to provide an advantage over learners who started at an older age. Likewise, exposure is hard to quantify and compare. To demonstrate, one participant reported "occasional interactions with anglophone friends" while another indicated having spent "8 months working with anglophones." Determining which situation constitutes greater exposure proved impossible, thus preventing us from exploring the influence of this variable in our data. Finally, we did not perform a fine-grained phonetic analysis of QF tokens of /h/, which could further elucidate whether /ʁ/ is at all substituted for /h/ underlyingly. In future research, it would be worthwhile to use a tool such as PRAAT for an acoustic analysis of QF and NS productions of [h] in English, comparing these with QF productions of [ʁ] in French.

6 Conclusion

Phenomena that involve L2 segments, such as the deletion of /h/ by QFs investigated here, are frequently variable (John and Frasnelli, 2022). This raises the question of how L2 segments are represented. If L2 representations are simply accurate, it remains to be seen why learners struggle to produce the target sounds and, more crucially, why learners have parallel problems in perception. QFs not only delete /h/, but they also fail to detect /h/ in the speech signal, the implication being that the state of the intervening representation is responsible for both. If, conversely, L2 representations are inaccurate, that is, if /h/ is simply left out of URs, it remains to be seen how QFs are able at times to produce and perceive /h/. Previous research has suggested that learners resort to an approximate representation of L2 phonemes (John and Frasnelli, 2022). Such diacritic representations are not as reliable as actual feature-based representations, but they allow learners to perceive and produce /h/ to varying degrees. Two earlier EEG studies, however, suggested another possible source for variation in h-production and perception: while [hʌm]-[ʌm] stimuli in an oddball paradigm failed to generate MMN among QFs (Mah et al., 2016), consistent with /h/ being left out of URs, [hɑ]-[ɑ] stimuli were accompanied by MMN (White et al., 2015), possibly because the phonetic properties of [h] before [ɑ] allow it to be assimilated to, and consequently represented as, the L1 phoneme /ʁ/. If /h/ is represented as /ʁ/ in items containing /ɑ/ (*hot*) but not /ʌ/ (*hut*), we hypothesized that QFs should show lower rates of h-deletion in such items. While QFs showed considerable variation in h-production in a reading-aloud task, none of this variation, however, could be attributed to the type of vowel occurring in an item. Our conclusion is that /h/ must be represented identically in the QF lexicon regardless of vowel type. While some QFs may simply leave /h/ out of URs such that *hot* and *hut* are stored as /ɑt/ and /ʌt/ and others may overcome the Laryngeal Input Constraint to develop the accurate representations /hɑt/ and /hʌt/, most of the participants in our study seem to have developed approximate representations based on non-linguistic markings: /ˀɑt/ and /ˀʌt/.

By extension, instead of attributing the different findings in the two EEG studies to the quality of the vowel, we conclude that the presence vs. absence of attention during the task was responsible for whether /h/ was detected: only when QFs pay attention to the signal does MMN emerge. Interestingly,

this observation is consistent with QFs having an approximate representation for /h/. While actual phonemic representations are associated with automatic and effortless processing of speech, whether in perception or production, approximate representations require effort and attention. It is unusual that QFs should resort to an approximate representation for /h/, given that [spread glottis], the sole feature required to represent this phoneme, is available from their L1 inventory. According to Brown (1998), this should make /h/ relatively easy for QFs to acquire. To resolve this conundrum, we suggest the presence of a Laryngeal Input Constraint that makes representations composed exclusively of laryngeal features particularly marked and hence dispreferred. This and other constraints on underlying representation are potentially what make certain L2 segments so difficult to acquire and what make L2 learners turn to alternate forms of non-feature-based representation to solve the puzzle of L2 phonological acquisition.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Comité d'éthique de la recherche de l'Université du Québec à Trois-Rivières. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

PJ: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Writing—original draft, Writing—review & editing. SR: Formal analysis, Methodology, Writing—original draft, Writing—review & editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/flang.2023.1286084/full#supplementary-material

References

Anttila, A. (1997). "Deriving variation from grammar: a study of Finnish genitives," in *Variation, change and phonological theory*, eds. F. Hinskens, R. Van Hout, and L. Wetzels (Amsterdam: John Benjamins), 935–68.

Archibald, J. (2005). Second language phonology as redeployment of LI phonological knowledge. *Canad. J. Linguist.* 50, 285–314. doi: 10.1017/S0008413100003741

Avery, P., and Rice, K. (1989). Segment structure and coronal underspecification. *Phonology* 6, 179–200. doi: 10.1017/S0952675700001007

Bassetti, B., Escudero, P., and Hayes-Harb, R. (2015). Second language phonology at the interface between acoustic and orthographic input. *Appl. Psycholing.* 36, 1–6. doi: 10.1017/S0142716414000393

Best, C. T., and Tyler, M. D. (2007). "Nonnative and second-language speech perception: commonalities and complementarities," in *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*, eds. M. J. Munro and O.-S. Bohn (Amsterdam: John Benjamins), 13–34. doi: 10.1075/lllt.17.07bes

Bittner, M. (1995). *Réalisation des constrictives/ ʃ /et/ʒ/en parler saguenéen: étude acoustique.* Unpublished Master's thesis, Université du Québec à Chicoutimi, Chicoutimi. doi: 10.1522/15 14294

Boersma, P. (1997). "How we learn variation, optionality, and probability," in *Proceedings of the Institute of Phonetic Sciences 21* (Amsterdam: University of Amsterdam), 43–58.

Brannen, K. (2011). *The perception and production of interdental fricatives in second language acquisition.* Unpublished doctoral dissertation, McGill University.

Broselow, E., Chen, S.-I., and Wang, C. (1998). The emergence of the unmarked in second language phonology. *Stud. Second Lang. Acquisit.* 20, 261–280. doi: 10.1017/S0272263198002071

Brown, C. (1998). The role of the L1 grammar in the acquisition of L2 segmental structure. *Second Lang. Res.* 14, 136–193. doi: 10.1191/0267658986695 08401

Cardoso, W. (2007). The variable development of English word-final stops by Brazilian Portuguese speakers: a stochastic optimality theoretic account. *Lang. Variat. Change* 19, 219–248. doi: 10.1017/S0954394507000142

Cedergren, H. J., and Sankoff, D. (1974). Variable rules: performance as a statistical reflection of competence. *Language* 50, 333–355. doi: 10.2307/412441

Chalmers, D. J. (1996). *The Conscious Mind: In Search of a Fundamental Theory.* Oxford: Oxford University Press.

Charette, M. (1991). *Conditions on Phonological Government.* Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511554339

Clements, G. N. (1985). The geometry of phonological features. *Phonol. Yearb.* 2, 225–252. doi: 10.1017/S0952675700000440

Darcy, I., Daidone, D., and Kojima, C. (2013). Asymmetric lexical access and fuzzy lexical representations in second language learners. *Mental Lexicon* 8, 372–420. doi: 10.1075/ml.8.3.06dar

Davidson, L., Smolensky, P., and Jusczyk, P. (2004). "The initial and final states: theoretical implications and explorations of Richness of the Base," in *Constraints in phonological acquisition,* eds. R. Kager, J. Pater, and W. Zonneveld (Cambridge: Cambridge University Press), 321–368. doi: 10.1017/CBO9780511486418.011

Davis, S., and Cho, M.-H. (2003). The distribution of aspirated stops and /h/ in American English and Korean: an alignment approach with typological implications. *Linguistics* 41, 607–652. doi: 10.1515/ling.2003.020

de Lacy, P. (2006). *Markedness – Reduction and Preservation in Phonology.* Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511486388

Dickerson, L. J. (1975). The learner's interlanguage as a system of variable rules. *TESOL Quart.* 9, 401–408. doi: 10.2307/3585624

Escudero, P., Hayes-Harb, R., and Mitterer, H. (2008). Novel second-language words and asymmetric lexical access. *J. Phonet.* 36, 345–360. doi: 10.1016/j.wocn.2007.11.002

Feldscher, C., and Durvasula, K. (2017). Excrescent stops in American English. *Proc. Linguistic Soc. Am.* 2, 1–15. doi: 10.3765/plsa.v2i0.4064

File-Muriel, R. J., and Brown, E. K. (2010). The gradient nature of s-lenition in Caleño Spanish. *Lang. Variat. Change* 23, 223–243. doi: 10.1017/S0954394511000056

Flege, J. E., and Bohn, O.-S. (2021). "The revised speech learning model (SLM-r)," in *Second Language Speech Learning: Theoretical and Empirical Progress,* ed. R. Wayland (Cambridge: Cambridge University Press), 3–83. doi: 10.1017/9781108886901.002

Gabriel, C., and Meisenburg, T. (2009). "Silent onsets? An optimality-theoretic approach to French *h aspiré* words," in *Variation and Gradience in Phonetics and Phonology, eds.* F. Kügler, C. Féry, and R. van de Vijver (Berlin: Mouton de Gruyter), 163–184. doi: 10.1515/9783110219326.163

Garellek, M. (2012). Glottal stops before word-initial vowels in American English: distribution and acoustic characteristics. *UCLA Work. Papers Phonet.* 110, 1–23. Available online at: https://escholarship.org/uc/item/7m55b8bb

Gouskova, M. (2023). MSCs in positional neutralization: the problem of gapped inventories. *Phonology.*

Guy, G. R. (2007). Lexical exceptions in variable phonology. *Univ. Pennsylvania Work. Papers Linguist.* 13, 109–119. Available online at: https://repository.upenn.edu/handle/20.500.14332/44657

Häcker, M. (2002). "Intrusive [h] in present-day English accents and <h>-insertion in medieval manuscripts," in *New perspectives on English historical linguistics, volume II: Lexis and transmission,* eds. C. Kay, C. Hough, and I. Wotherspoon (Amsterdam: John Benjamins), 109–123. doi: 10.1075/cilt.252.09hac

Halle, M., and Idsardi, W. J. (1997). "*r,* hypercorrection, and the Elsewhere Condition," in *Derivations and constraints in phonology,* ed. I. Roca (Oxford: Clarendon), 331–348. doi: 10.1093/oso/9780198236894.003.0010

Harris, J. (1990). Segmental complexity and phonological government. *Phonology* 7, 255–300. doi: 10.1017/S0952675700001202

Harris, J. (1994). *English Sound Structure.* Oxford: Blackwell.

Harris, J. (1997). Licensing inheritance: an integrated theory of neutralisation. *Phonology* 14, 315–370. doi: 10.1017/S09526757980 03479

Harris, J., and Kaye, J. (1990). A tale of two cities: London glottalling and New York City tapping. *Linguist. Rev.* 7, 251–274. doi: 10.1515/tlir.1990.7.3.251

Hayes-Harb, R., and Barrios, S. (2021). The influence of orthography in second language phonological acquisition. *Lang. Teach.* 54, 297–326. doi: 10.1017/S0261444820000658

Honeybone, P. (2008). "Lenition, weaking and consonantal strength: tracing concepts through the history of phonology," in *Lenition and Fortition,* eds. J. Brandão de Carvalho, T. Scheer, and P. Ségéral (Berlin: Mouton de Gruyter), 9–92. doi: 10.1515/9783110211443.1.9

Jackson, S., and Cardoso, W. (2023). "Orthographic interference in the acquisition of English /h/ by Francophones," in *Second language pronunciation,* eds. U. K. Alves and J. I. A. de Albuquerque (Berlin/Boston: De Gruyter Mouton), 229–248. doi: 10.1515/9783110736120-009

Janda, R. D., and Auger, J. (1992). "Quantitative evidence, qualitative hypercorrection, sociolinguistic variables – and French speakers' 'eadhaches with English h/Ø. *Lang. Commun.* 12, 195–236. doi: 10.1016/0271-5309(92)90015-2

Jeffreys, H. (1961). *Theory of Probability (3rd edition).* Oxford: Oxford University Press.

John, P. (2006). *Variable h-epenthesis in the interlanguage of francophone ESL learners.* Unpublished Master's thesis, Concordia University.

John, P., and Cardoso, W. (2009). "Francophone ESL learners' difficulties with English /h/," in *Recent research in second language phonetics/phonology: Perception and production,* eds. M. A. Watkins, A. S. Rauber, and B. O. Baptista (Newcastle upon Tyne: Cambridge Scholars Publishing), 118–140.

John, P., and Frasnelli, J. (2022). On the lexical source of variable L2 phoneme production. *The Mental Lexicon* 17, 239–276. doi: 10.1075/ml.22002.joh

Keating, P. (1988). Underspecification in phonetics. *Phonology* 5, 275–292. doi: 10.1017/S095267570000230X

Kroch, A. (1989). Reflexes of grammar in patterns of language change. *Lang. Variat. Change* 1, 199–244. doi: 10.1017/S0954394500000168

LaCharité, D., and Prévost, P. (1999). "The role of L1 and of teaching in the acquisition of English sounds by francophones," in *Proceedings of the 23rd Annual Boston University Conference on Language Development,* eds. A. Greenhill, H. Littlefield, and C. Taro (Somerville, MA: Cascadilla Press), 373–385.

Ladefoged, P., and Maddieson, I. (1996). *The Sounds of the World's Languages.* Oxford: Blackwell.

Leben, W. (1973). *Suprasegmental phonology.* PhD dissertation, MIT. Distributed by Indiana University Linguistics Club.

Lindsay, G. W. (2020). Attention in psychology, neuroscience, and machine learning. *Font. Comput. Neurosci.* 14, 29. doi: 10.3389/fncom.2020.00029

Lombardi, L. (1995). Laryngeal neutralization and syllable wellformedness. *Nat. Lang. Lingu. Theory* 13, 39–74. doi: 10.1007/BF00992778

Lombardi, L. (2002). Coronal epenthesis and markedness. *Phonology* 19, 219–251. doi: 10.1017/S0952675702004323

Mah, J., Goad, H., and Steinhauer, K. (2016). Using event-related brain potentials to assess perceptibility: the case of French speakers and English. *Front. Psychol.* 7, 1469. doi: 10.3389/fpsyg.2016.01469

McCarthy, J. J. (1986). OCP effects: gemination and antigemination. *Ling. Inquiry* 17, 207–263.

McCarthy, J. J. (2002). *A Thematic Guide to Optimality Theory.* Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511613333

McMahon, A. (2002). *An Introduction to English Phonology.* Edinburgh: Edinburgh University Press.

Melnik, G. A., and Peperkamp, S. (2019). Perceptual deletion and asymmetric lexical access in second language learners. *J. Acoust. Soc. Am.* 145, EL13. doi: 10.1121/1.5085648

Milroy, J. (1983). "On the sociolinguistic history of /h/ -dropping in English," in *Current topics in English historical linguistics,* eds. M. Davenport, E. Hansen, and H. Nielsen (Odense: Odense University Press), 37–53.

Morin, Y.-C. (2002). Les premiers immigrants et la prononciation du français au Québec. *Rev. Québécoise Linguist.* 31, 39–78. doi: 10.7202/006844ar

Núñez-Méndez, E. (2022). Variation in Spanish /s/: overview and new perspectives. *Languages* 7, 1–50. doi: 10.3390/languages7020077

Odden, D. (1986). On the role of the obligatory contour principle in phonological theory. *Language* 62, 353–383. doi: 10.2307/414677

Ohala, J., and Ohala, M. (1993). The phonetics of nasal phonology: theorems and data. *Phonet. Phonol* 5, 225–249. doi: 10.1016/B978-0-12-360380-7.50013-2

Orgun, C. O. (2002). English *r*-insertion in optimality theory. *Nat. Lang. Ling. Theory* 19, 737–749. doi: 10.1023/A:1013313827798

Paradis, C., and LaCharité, D. (2001). Guttural deletion in loanwords. *Phonology* 18, 255–300. doi: 10.1017/S0952675701004079

Paradis, C., and Lebel, C. (1994). Contrasts from segmental parameter settings in loanwords: core and periphery in Quebec French. *Toronto Work. Papers Ling.* 13, 75–94.

Preston, D. R. (1989). *Sociolinguistics and Second Language Acquisition.* Oxford: Blackwell.

Prince, A., and Smolensky, P. (1993/2004). *Optimality Theory: Constraint Interaction in Generative Grammar.* Hoboken: Blackwell Publishers. doi: 10.1002/9780470759400

R Core Team (2021). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. Available online at: https://www.R-project.org/ (accessed August 23, 2023).

Rafat, Y. (2016). Orthography-induced transfer in the production of English-speaking learners of Spanish. *Lang. Learn. J.* 44, 197–213. doi: 10.1080/09571736.2013.784346

Rennicke, I. E. (2015). *Variation and change in the rhotics of Brazilian Portuguese*. Unpublished doctoral dissertation, Universidade Federal de Minas Gerais.

Rouder, J. N., Morey, R. D., Speckman, P. L., and Province, J. M. (2012). Default Bayes factors for ANOVA designs. *J. Mathem. Psychol.* 56, 356–374. doi: 10.1016/j.jmp.2012.08.001

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., and Iverson, G. (2009). Bayesian t-tests for accepting and rejecting the null hypothesis. *Psychon. Bull. Rev.* 16, 225–237. doi: 10.3758/PBR.16.2.225

Sankoff, G., and Blondeau, H. (2007). Language change across the lifespan: /r/ in montreal French. *Language* 83, 560–588. doi: 10.1353/lan.2007.0106

Smolensky, P. (1996). *The initial state and 'Richness of the Base' in optimality theory*. Technical Report JHU-CogSci-96–4, John Hopkins University, ROA-154.

Strange, W., and Shafer, V. L. (2008). "Speech perception in second language learners," in *Phonology and Second Language Acquisition,* eds. J. G. Hansen-Edwards and M. L. Zampini (Amsterdam: John Benjamins), 153–191. doi: 10.1075/sibil.36.09str

Tousignant, C. (1987). Les variantes du/R/montréalais: Contextes phonologiques favorisant leur apparition. *Rev. québécoise Linguist. Théor. Appl.* 6, 73–113.

Tranel, B. (1995). "Current issues in French phonology: liaison and position theories," in *The Handbook of Phonological Theory,* ed. J.A. Goldsmith (Oxford: Blackwell), 798–816.

Vaysman, O. (2002). Against richness of the base: evidence from Nganasan. *Ann. Meet. Berkeley Lingu. Soc.* 28, 327–338. doi: 10.3765/bls.v28i1.3848

Walker, D. C. (2001). *French Sound Structure*. Calgary, University of Calgary Press.

Wells, J. C. (1982). *Accents of English 1: An Introduction*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511611759

Werker, J. F., and Logan, J. S. (1985). Cross-language evidence for three factors in speech perception. *Percept. Psychophys.* 37, 35–44. doi: 10.3758/BF03207136

Werker, J. F., and Tees, R. C. (1984). Phonemic and phonetic factors in adult cross-language speech perception. *J. Acoust. Soc. Am.* 75, 1866–1878. doi: 10.1121/1.390988

White, E. J., Titone, D., Genesee, F., and Steinhauer, K. (2015). Phonological processing in late second language learners: the effects of proficiency and task. *Bilingualism* 20, 162–183. doi: 10.1017/S1366728915000620

Yip, M. (1988). The Obligatory Contour Principle and phonological rules: a loss of identity. *Lingu. Inquiry* 19, 65–100.