

Intrusive tokens of aspiration in French learners' L2 English

Christelle Exare

Laboratoire de Phonétique et Phonologie, CNRS (UMR 7018) / Paris3, Sorbonne-Nouvelle, France
christelle.exare@gmail.com

ABSTRACT

This paper addresses the production and the perception of ‘intrusive [h]’ in French learners’ L2 English (e.g. : *I ^hate pasta* instead of *I ate pasta*). Based on the phonetic description of L1 and L2 English word-initial vowels, this paper provides three explanations and three teaching applications to deal with such phonologically illicit occurrences. **First**, the onsets of English words are phonetically analysed in three types of data: i) a text read by 8 native English speakers and 10 French learners of English, ii) spontaneous speech elicited from 25 French learners and iii) a perception test taken by 30 French-speaking students. The results suggest that a pause, some glottalisation or some aspiration are three processes that increase the time span between two vowels in a hiatus context. **Second**, three processes accounting for “intrusive h” are put forward. **Third**, some phonetic training is proposed.

Keywords: [h], intrusion, aspiration, English as a Foreign Language, French learners

1. INTRODUCTION

Intrusive tokens of aspiration are commonly referred to as “epenthetic /h/” [21]. Although rare [19], such occurrences are perceptually salient and often strongly stigmatised, particularly in proficient learners [8, 26]. Since such errors generate semantic inconsistencies (*we ^heat*), they are supposed to deserve immediate and active correction. Some of the triggering parameters of intrusive aspirations have already been identified. Intrusive aspirations usually occur at stressed vocalic onsets. They are more common after a pause or a vowel, and when [h] can be found in the preceding context [21].

2. EXPERIMENTS

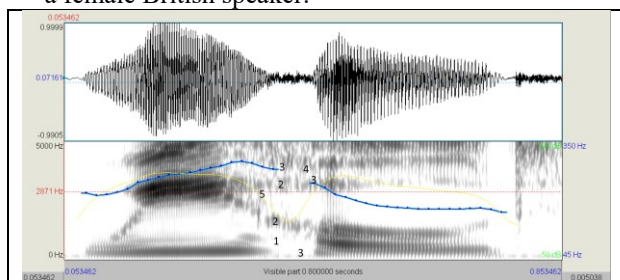
The onsets of some English words were analysed in three types of data: i) a text read by 8 native English speakers and 10 French learners of English, ii) spontaneous speech elicited from 25 French learners and iii) a perception test taken by 30 French-speaking students. The learners who took part in this study were all assessed at levels A2 or B1 of the Common European Framework of Reference for Languages. They all studied English in a classroom environment.

2.1 Production: Read speech

2.1.1 Protocol

The first experiment was a pilot study aiming at comparing hiatus resolution in L1 and L2 English. An *ad hoc* text including 20 noun phrases was read by 8 native English speakers and 10 fifteen-year-old female learners of English. The noun phrases consisted of the article THE followed by an adjective or a noun starting by a long vowel (/ɜ:/, /ɑ:/, /ɔ:/, and /i:/). For example, one sentence was: “*the autumn was so cold for the hawk and the eagle that they flew to the east.*” The native speakers were from diverse origins. Three were from the United Kingdom. Three were from the United States. One was Canadian and one was from New Zealand. The productions were coded according to the type of vowel phonation chosen for hiatus resolution. Using Praat (Boersma 2001), four types of vowel phonation were acoustically coded: i) modal, ii) glottalised, iii) aspirated iv) other (e.g.: elision or unpronounced item). The rules established by [13, 25, 27] were followed to label the segments. In this paper, vowel-initial glottalisation acoustically corresponds to what is perceptually described as “hard attack” or “hard onset.” In a gestural approach [6], and as they characterise vowels here, the terms “aspiration” and “glottalisation” refer to two articulatory gestures that interfere with voicing. Glottalisation and aspiration can either precede the voicing gesture, or overlap it. The acoustic correlates of true English [h], which is often described as the voiceless counterpart of the following vowel, are numbered in figure 1 [25]ⁱ. They are: a weakened first formant partly corresponding to the opening of the glottis (1), F2 and F3 clear and continuous transitions (2), a complete or incomplete interruption of voicing (3) and finally, aspiration noise in mid frequencies (4), and a drop in intensity (5).

Figure 1. Spectrogram of *we hold*, pronounced by a female British speaker.



In this French corpus of L2 English most tokens of intrusive aspiration could be segmented as [h]. In some erratic cases, however, /h/ was categorically perceived, although no real segment could be isolated. Instead, breathy vowels could be found. From an articulatory point of view, a breathy vowel has a double source: vocal fold vibration and aspiration noise. Hence the phrase “voiceless vowel” does not seem to be appropriate to describe such occurrences, as they are indeed produced with voicing, despite incomplete vocal folds adduction that results in aspiration noise.

2.1.2 Results

First, glottalisation is preferred to modal voice by both groups of speakers (natives and non-natives) at THE#V boundary, where # stands for a word boundary. However, the F0 curve tends to drop at the onset of the second vowel in native productions, while it goes up when an intrusive token of aspiration is produced in non-native productions. **Second**, three categories of learners emerge according to their choice of vowel phonation at V#V boundary. Most students produce both modal vowels and glottalised vowels in this context, but 5 of them significantly prefer glottalisation to modal voicing. 1 student produces 17 intrusive tokens of aspiration (in a total of 20 noun phrases) and no glottalisation at all. On the contrary, 2 other students never insert [h], while i) they utter glottalised vowels more than the others and ii) they always elide initial /h/ in contexts like “*the harm*,” “*the hawk*,” “*the heart*” and “*the heater*.” **Third**, glottalisation seems to be preferred by the native speakers who do *not* speak Southern Standard British English (see [11] for more detail on American glottal stops in V#V contexts). In this pilot study glottalised and modal voicing are two options equally chosen by British speakers. This is in line with [18]: vowel-initial glottalisation is not a forbidden option for native speakers of British English.

Table 1. Type of phonation uttered by i) 3 native speakers of Southern Standard British English and ii) 3 native speakers of American English, at V#V boundary, in a THE#V context. For each speaker the same 20 noun phrases were analysed.

	UK	US
modal	27	14
glottalised	31	45
other	2	1

A chi-square test for independence compares the two distributions and categorical variables (in grey in the table above). The result indicates a significant association between the speaker’s linguistic origin and the type of phonation, as $\chi^2(1, n = 117) = 6.69$ with $p = 0.009$. So a relationship can be found between the type of phonation (“modal” vs glottalised”) and the linguistic origin of the speaker.

2.2 Production: Spontaneous speech

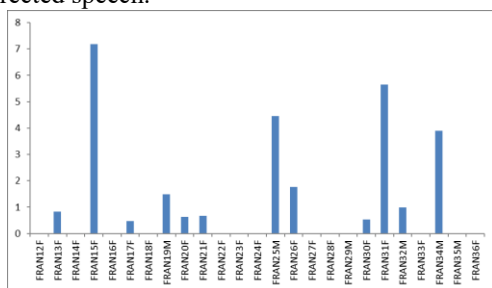
2.2.1 Protocol

25 French students (19 females and 6 males), aged 15 years old, were recorded during a semi-structured interview. The experimenter asked them questions about their studies, their hobbies and the trip they had taken two weeks before the recording took place. The goal was to elicit utterances like “*I asked*” or “*she answered*.” The acoustic data were collected through a headset microphone. On the whole, 160 minutes of L2 English casual speech uttered by 25 learners were recorded, orthographically transcribed, and acoustically analysed.

2.2.2 Results

The main results concern i) intra- and inter-speaker variability, ii) the phonetic distribution of intrusive h, and iii) the role of the task. **First**, only 48 intrusive tokens of aspiration were identified and isolated at L2 English word onsets. They are produced by 12 (in 25) learners. Their frequency exhibits high inter- and intra-speaker variability. 2 students insert more than 10 intrusive aspirations in their recordings that approximate 10 minutes. 10 students do not utter intrusive aspirations at all during their interviews.

Figure 2. Ratio of intrusive aspirations to the total number of #V contexts in 25 French learners' semi-directed speech.



Intra-speaker variability can also be observed. 13 lexically similar utterances were carefully compared, like “*we asked.*” For each pair uttered by the same learner, one item had an intrusive glottal fricative at the onset of the vowel-initial word, while the second item was realised with some glottalisation. In other words, learners tend to choose either some glottalisation or some aspiration in order to pull away two vowels in a hiatus. A pause is a third option, that may, in addition, increase the distance between the two vowels. **Second**, in the present corpus, intrusive tokens of aspiration only surface i) in strict initial position or ii) after a vocalic sound, which may be an intrusive word-final schwa. When the word-final phoneme preceding the aspiration is a plosive (e.g.: *eat, trip, about*) or a fricative (e.g.: *is, practise*), a phonetic intrusive schwa emerges before illicit [h]. For instance, “*I love animals*” is pronounced [aɪləvə#haniməlz]. **Third**, the type of speaking task taken by the two different groups of participants seems to play a role in the emergence of phonetic tokens of aspiration. An independent-samples t-test indicates there is a significant difference ($t = 2.13$; $p = 0.04$) in the proportion of intrusions between read speech ($M = 0.12$, $SD = 0.2$) and casual speech ($M = 0.01$, $SD = 0.02$).

2.3 Perception

2.3.1 Literature

The phoneme /h/ is hardly perceptible in weak positions in Turkish and other languages [24]. Syllable-initial English /h/ is visually well-perceived [20]. Although it may acoustically be confused with initial /p/ by native and foreign speakers [9], English /h/ is equally recognized by native speakers and French learners [22]. The learners’ phonetic representations of /h/ seem to be at the origin of their particular treatment of English /h/ [23].

2.3.2 Protocol

A perception experiment was conducted in order to determine how French learners perceive: i) glottalised

vowels and ii) word-initial English /h/ when it is licit (produced by native speakers) or illicit (inserted by French learners). The hypotheses are i) learners perceive /h/ adequately in both conditions (licit English /h/ vs French intrusive aspiration), ii) initial glottalised vowels are perceived as /h/. 30 older and more advanced French students took the AXB test. 4 series of 20 stimuli were taken from i) the Aix-MARSEC Database [2], ii) the corpus of read speech described in section 2.1 and iii) the corpus of casual speech described in section 2.2. Each stimulus is a word followed by the first syllable of a second word starting either by a (glottalised vs modal) vowel or by (licit vs illicit) initial /h/. The syllabic boundary is that given by [29]. After a mock test the listeners had to determine whether they heard a vowel or [h] at the onset of the second word, in the 80 stimuli. 2400 responses were analysed.

2.3.3 Results

The 30 listeners performed well. Initial vowels are perceived as vowels (82 % of correct answers) while [h] is generally identified as /h/ (80 %). However, [h] is better identified when it is pronounced by English speakers than when it is pronounced by French learners: $\chi^2(1, 1200) = 9.18$ with $p = 0.002$. Although a tendency to perceive glottalised vowels as /h/ is noted, no significant effect of the type of phonation on the perception of the onset can be found: $\chi^2(1, 1200) = 1.06$ with $p = 0.3$.

Table 3. AXB perception test taken by 30 French listeners. The targets are L1 and L2 English vocalic and aspirated word onsets. Number and ratio of responses for each category. 4 categories of 20 stimuli are presented: i) modal vowels : (M)V ; glottalised vowels: (G)V; L2 English intrusive [h]: *H; L1 English licit [h]: H.

SOUND		Played			
		(M) V	(G) V	*H	H
Perceived	V	512	499	141	99
	H	88	101	459	501
	%V	85,33%	83,17%	23,50%	16,50%
	%H	14,67%	16,83%	76,50%	83,50%

3. DISCUSSION

3.1 Proficiency level

No consistent relationship between the learners’ tendency to insert and drop [h] could be found. However, it is hypothesized here that a learner has a good command of the oral features if they:

- i) preserve /h/ where the realisation of the phoneme is expected (the elision rate is below 10% of all possible cases).
- ii) never insert intrusive aspirations (the insertion rate is below 4% of all possible cases).

Some future research may confirm whether the calculation of a H-ratio based on the rates of elision and insertion is a reliable level indicator.

3.2 Three explanatory processes

Illicit tokens of aspiration at word onsets in L2 English can be considered as occurrences of hypercorrection [14], which may result from three distinct, yet possibly complementary processes. **First**, some students may produce [ʔ] instead of [h] and /or [h] instead of [ʔ] because of an incomplete assimilation of the English [ʔ] ~ [h] contrast. **Second**, initial h-insertion (combined with final ə-insertion) may signal some optional phonological repair, restoring French rhythm through typically CV#CV syllabic structures. **Third**, an intrusive aspiration may be an intrusive gesture of glottal opening [6], sometimes overlapping the voicing of the vowel (i.e. when [h] can hardly be segmented) or a glottal constriction gesture that fails to reach its target (i.e. an inchoative glottalisation). In fact, glottalisation and aspiration both correspond to glottal tension [28], although glottalisation correlates with a closed glottis and aspiration correlates with an open glottis. Confusion between the two modes of articulation may come from the learner's lack of control over glottal gestures. From a phonetic point of view, word-initial intrusive tokens of aspiration in L2 English should be studied in relation to word-initial glottalised vowels.

4. TEACHING APPLICATIONS

Relying on multimodal input, the author's teaching proposal aims at raising a learner's awareness of i) glottal control for aspiration, glottalisation, and continuous modal voicing across word boundaries and ii) some syllabic specificities of French and English that make word boundaries potential stumbling blocks in French learners' L2 English, iii) variability in the realisations of /h/. The framework for this set of activities is that of the weak-interface model [15, 16], which posits that instruction should combine explicit and implicit knowledge.

4.1 Multimodal input

Acquisition of languages is favoured by multimodal input. Relying on visual and auditory input can increase the linguistic and cognitive abilities of children with communication disorders [3]. Subtitling

videos has been shown to improve the results of second-language learners in aural comprehension tasks [5, 10]. However, to train students in speaking, casual speech or repetition should be preferred to reading tasks, in order to minimise the risks of producing intrusive aspirations.

4.2 Training in glottal control

Students ought to be encouraged to distinguish between two processes to perform vowel-to-vowel linking: i) continuous voicing (corresponding to the uninterrupted vibration of vocal folds) as opposed to ii) glottalisation (corresponding to an abruptly closed glottis). Kinesthaetic activities, like the placement of one's hand near the glottis can be useful here. To make students perceive what glottalisation is, an analogy with French can be drawn. Glottal stops are usually produced at the beginning of exclamations like "Hein?" Also, the study of a popular song like "Hello" [1] can allow the teacher to introduce the contrast between continuous voicing, glottalisation and aspiration.

4.3 Syllabic specificities

French learners should be taught about syllabic structures. French prefers CV syllables while English favours CVC syllables [12]. French words ending with a phonological consonant tend to be uttered with a final --phonetic-- intrusive schwa in French [7] and in L2 English [17]. Hence students should be explicitly told that they do not need to add extra schwas to final consonants in L2 English.

4.4 Variation in the realisations of English /h/

French learners should also be taught about the prosody-driven variation in /h/ realisations. It is important to stress that English /h/ can be weakened in unstressed positions. In particular the opposition between HAVE (lexical verb) and HAVE (auxiliary) is a major difficulty for French learners. This gradient characteristic of /h/ could be represented as variable-sized symbols (e.g.: fences). Visualizing real-time spectrograms may also be an interesting tool to help students improve their pronunciation.ⁱⁱ

6. REFERENCES

- [1] Adele. 2015. *Adele 25*. Audio CD. XI Recordings.
- [2] Auran, C., C. Bouzon & D. Hirst. 2004. The Aix-MARSEC Project: An Evolutive Database of Spoken English. *Proceedings of the Second International Conference on Speech Prosody*, 561–564. Nara, Japan: Bel, B. & Marlen, I.
- [3] Beukelman, D. & P. Mirinda. 1998. *Augmentative and Alternative Communication: Management of*

- Severe Communication Disorders in Children and Adults*. Brookes.
- [4] Boersma, P. 2001. Praat, a system for doing phonetics by computer. *Glott International* 5(9/10). 341–345.
- [5] Borrás, I. & R. C. Lafayette. 1994. Effects of Multimedia Courseware Subtitling on the Speaking Performance of College Students of French. *The Modern Language Journal* 78(1). 61–75.
- [6] Browman, C. P. & L. Goldstein. 1992. Articulatory Phonology: An Overview. *Phonetica* 49(3–4). 155–180.
- [7] Carton, F. 1999. L'épithèse vocalique en français contemporain: étude phonétique. *Faits de langues* 7(13). 35–45.
- [8] Charlot, C. 2015. *Rapport de jury. Concours de l'agrégation externe d'anglais. Session 2015*. CNDP.
- [9] Cutler, A., A. Weber, R. Smits & N. Cooper. 2004. Patterns of English phoneme confusions by native and non-native listeners. *The Journal of the Acoustical Society of America* 116(6). 3668–3678.
- [10] Danan, M. 2004. Captioning and Subtitling: Undervalued Language Learning Strategies. *Meta: Journal des traducteurs. Translators' Journal* 49(1). 67–77.
- [11] Davidson, L. & D. Erker. 2014. Hiatus resolution in American English: The case against glide insertion. *Language* 90(2). 482–514.
- [12] Delattre, P. 1965. *Comparing the phonetic features of English, French, German and Spanish. An interim report*. Heidelberg: J. Groos Verlag.
- [13] Dilley, L., S. Shattuck-Hufnagel & M. Ostendorf. 1996. Glottalisation of word-initial vowels as a function of prosodic structure. *Journal of Phonetics* 24. 423–444.
- [14] Eckman, F. R., Gregory K. Iverson & J. Song. 2013. The role of hypercorrection in the acquisition of L2 phonemic contrasts. *Second Language Research* 29(3). 257–283.
- [15] Ellis, Nick C. 1994. *Implicit and Explicit Learning of Languages*. London ; San Diego: Academic Press Inc.
- [16] Ellis, R. J. 2008. *The Study of Second Language Acquisition*. Oxford University Press. (30 June, 2014).
- [17] Faure, G. 1975. *Manuel pratique d'anglais parlé*. Paris: Hachette.
- [18] Fuchs, R. 2015. Word-Initial Glottal Stop Insertion, Hiatus Resolution and Linking in British English. *Proceedings of Interspeech 2015*, 1675–1679. Dresden, Germany. (25 February, 2017).
- [19] Janda, R., D. & J. Auger. 1992. Quantitative Evidence, Qualitative Hypercorrection, Sociolinguistic Variables --and French Speakers' eadaches with English h/Ø. *Language and Communication* 12(3/4). 195–236.
- [20] Jiang, J., A. Alwan, E.T. Auer & L. Bernstein. 2001. Predicting Visual Consonant Perception from Physical Measures. *Eurospeech*.
- [21] John, P. & W. Cardoso. 2008. Francophone ESL Learners and h-epenthesis. *Concordia Working Papers in Applied Linguistics* 1. 76–97.
- [22] LaCharité, D. & P. Prévost. 1999. The Role of L1 and teaching in the acquisition of English sounds by francophones. *Proceedings of BUCLD 23*, 373–385. Cascadilla.
- [23] Mah, J., K. Steinhauer & H. Goad. 2006. The Trouble with /h/: Evidence from ERPs. *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006)*, 80–87. Somerville, MA: Mary Grantham O'Brien, Christine Shea, and John Archibald.
- [24] Mielke, J. 2003. The interplay of speech perception and phonology: Experimental evidence from Turkish. *Phonetica* 60(3). 208–229.
- [25] Pierrehumbert, J. & D. Talkin. 1991. Lenition of /h/ and glottal stop. *Papers in Laboratory Phonology II*, 90–117. Cambridge UK: Cambridge University Press. (14 June, 2010).
- [26] Rotgé, W. 2008. *Rapport de jury. Concours de l'agrégation externe d'anglais. Session 2008*. Ministère de l'Éducation Nationale. Ministère de l'Enseignement Supérieur et de la Recherche.
- [27] Umeda, N. 1978. Occurrence of Glottal Stop in Fluent Speech. *The Journal of the Acoustical Society of America* 64(1). 88–94.
- [28] Vaissière, J. 1986. Variance and Invariance at the Word Level. *Invariance and Variability in Speech Process*, 534–539. Lawrence Erlbaum Associates. (27 August, 2015).
- [29] Wells, J. 2008. *Longman Pronunciation Dictionary*. Pearson Longman.

ⁱ For more detail on the acoustic cues for [h], the reader may want to refer to Kenneth Stevens (1998). *Acoustic Phonetics*. The MIT Press. Cambridge, Massachusetts. (Chapter 8.3.1.5)

ⁱⁱ The author wants to thank Timothy De Cotis, Lucas Drouhot, and two anonymous reviewers for their kind and helpful comments.