Morgan Robertson and Allard Jongman

Transfer and proficiency effects on L2 French perception of nasal vowels: a gating study

Abstract: The process by which gradient acoustic cues are discretely categorized as mental phonological representations seems to be effortless for native speakers. For second language (L2) learners, perceiving the cues of the L2 often results in miscategorizations due to the influence of the first language (L1) system already in place. The present study examines L2 speech perception of nasal vowels by L1 English learners of French. A gating study was conducted to address L1 English learners' utilization of nasality in the acoustic signal as well as proficiency effects on the perception of nasal vowels. We adopt an approach which proposes that features extracted from the signal are mapped onto an underspecified lexicon, where lexical activation relies on a direct match or no-mismatch to the features extracted (Lahiri and Marslen-Wilson 1991). Crucially, in French, the feature [NASAL] is contrastive for vowels, but in English the feature is allophonic. Participants completed the LexTale-FR proficiency task (Brysbaert 2013) and a gating task in which they were presented with successively longer portions of CVN, CVC, and CṼ stimuli. Based on mixed-effects models, results showed that low-proficiency participants exhibited more L1 influence than high-proficiency participants; however, high-proficiency participants did not exhibit native-like behavior in the task. The results are interpreted within the framework of an underspecified lexicon as evidence for L1 influence at low-proficiency levels. Additionally, it is posited that advanced learners create a new category based on vowel quality rather than nasality.

Keywords: Second language acquisition, nasal vowels, perception, gating, underspecification

1 Introduction

The influence of the first language (L1) on the second language (L2) during the process of second language acquisition is a well-studied phenomenon. Some models of second language speech perception consider this influence rooted in phonological contrast (e.g., Best and Tyler 2007). Other models consider the impact to be more phonetic (e.g., Flege 1995; Flege and Bohn 2021). While these models often make similar predictions, understanding the level of mapping within the pho-

https://doi.org/10.1515/9783110783452-004

netic/phonological system is important. The present study explores the extent to which L2 acoustic-phonetic nasality activates L1 phonological representations. We investigate how the gradient cue of nasality in a French speech signal is categorized in the phonology of an L1 English speaker. The perception of this cue is complicated by the phonological distribution of the feature [NASAL] in both languages. In French, the feature [NASAL] is contrastive for vowels, but in English the feature is allophonic. Research in the field of second language acquisition predicts perceptual difficulties with this contrast for L1 English L2 French learners due to the phonetic similarity between the L1 and L2 sounds and the phonological differences in the distribution of the feature [NASAL].

Vowels in General American English do not exhibit a phonological nasal–oral contrast. Instead, there is an allophonic relationship between nasal and oral vowels in English: When a vowel occurs after a nasal consonant, the vowel becomes nasalized by way of anticipatory coarticulation. For example, though the word *bat* [bæt] has an oral low front lax vowel, the word *ban* [bæ̃n] includes a nasal vowel due to the anticipatory coarticulation of nasality from the final nasal consonant (Solé 1992: 35–36). This allophonic relationship is of interest to the present study because it is this allophony that is suspected to interfere with acquisition of the L2 contrast.

French vowels do exhibit a phonological contrast between non-high nasal and oral vowels. Specifically, the vowels [ɛ] as in *fait* [fɛ] 'does,' [a] as in *la* [la] 'the,' and [ɔ] as in *sobre* [sɔbʀ] 'sober' each have a contrastive nasal counterpart: [ɛ̃] as in *faim* [fɛ̃] 'hunger,' [ɑ̃] as in *lent* [lɑ̃] 'slow,' and [ɔ̃] as in *sombre* [sɔ̃bʀ] 'somber.' Some dialects of French also include the oral–nasal vowel contrast [œ] ~ [œ̃]; however, in many dialects (including the Northern Metropolitan French dialect that is taught in the classroom) the nasal vowel [œ̃] merges with [ɛ̃] (Walker 2001: 108) and is therefore not included in the present study. These oral–nasal counterparts play an important role in French morphology, serving functions such as distinguishing between masculine and feminine nouns, adjectives, and determiners, as in *Parisien* [paʀizjɛ̃] 'Parisian' (masc.) vs. *Parisienne* [paʀizjɛn] 'Parisian' (fem.), and distinguishing between singular and plural verb forms, as in *tiens* [tjɛ̃] 'I hold' vs. *tiennent* [tjɛn] 'they hold' (Walker 2001: 112–113). Notably, French exhibits a neutralization in which the nasal–oral contrast does not occur before a nasal consonant (cf. Sampson 1999 for a diachronic explanation). In French CVN sequences, like *somme* [sɔm] or *Seine* [sɛn], phonemically nasal vowels cannot occur. Even when *la liaison* occurs in French there are instances when an underlyingly nasal vowel will become oral when it appears before a nasal consonant, *bon* [bɔ̃] *ami* [a.mi] becomes *bon ami* [bɔ.na.mi] (cf. Steriade 1999 for an analysis). Additionally, in examining Northern Metropolitan French speakers, Dow (2020: 263) found that nasal coarticulation is systematically blocked when a non-high vowel occurs in this CVN, prenasal context. This results in an oral vowel with very little coarticulatory nasality occurring right before a nasal consonant, in the exact same sequence in which English speakers systematically nasalize the vowel.

Despite the morphophonological relationship between nasal and oral vowels in French, nasal vowels are not acoustically *nasalized oral vowels* (Carignan 2014: 25; Cohn 1990, 2007; Dow 2020: 241). While *phonologically*, /ɛ/ and /ɛ̃/ are often contrasted (i.e., *plein* /plɛ̃/ ~ *pleine* /plɛn/), *phonetically*, the nasal vowel manifests a formant structure closer to [æ̃] (i.e., *plein* [plæ̃] ~ *pleine* [plɛn]). All French nasal vowels have been shown to exhibit spectral changes compared to their oral counterparts, such that phonological /ɛ̃, ɑ̃, ɔ̃/ are phonetically realized as [æ̃, ɔ̃, õ], respectively (Walker 2001: 107; Dow 2020: 241). Native speakers of French could use spectral changes as a cue in addition to the presence of nasality. The question of whether non-native speakers also rely on these spectral changes is still unanswered. One hypothesis is that these spectral changes serve as a strong cue to identify *nasal* vowels (CṼ) as opposed to *prenasal* oral vowels (CVN). Crucially, when presented with a prenasal oral vowel, a non-native speaker may anticipate a nasal consonant because there is still likely to be sufficient nasality present in the signal.

This difference in distribution of nasal and oral vowels in English and French is an ideal testing ground for examining perception of phonological differences between an L1 and L2 since L1 English L2 French learners would have to perceive the nasality in the signal as a contrastive feature of the vowel rather than a cue for an upcoming nasal consonant. Though many studies have examined second language acquisition of a contrast that is allophonic in the L1, few studies have looked at English speakers' acquisition of this contrast with neutralization pattern in L2 French (Marquez Martinez 2016).

2 Nasal categories and abstraction

2.1 L2 speech perception

The Speech Learning Model (SLM, Flege 1995) and the later Speech Learning Model – revised (SLM-r, Flege and Bohn 2021) posit learning begins with a complete mapping of all L2 sounds onto preexisting L1 phonetic categories. The SLM proposes that the same mechanism that enables infants and young children to develop and hone their L1 categories drives the development of L2 categories. Namely, statistical input distributions shape phonetic categories and form a collection of sounds that match a narrow band of "prototypes." The L2 sounds heard are initially mapped onto the phonetically closest L1 category. Learning new categories is possible, according to the SLM, though gradual and time-consuming. However, the SLM predicts that the newly formed L2 categories will likely not match those of a monolingual speaker.

An important aspect of the SLM and the SLM-r (henceforth: SLM(-r)) is its conceptualization of phonetic categories as position-sensitive allophones rather than phonemes. SLM(-r) predicts that learners will gradually create a new phonetic category as they are further exposed to the richly detailed phonetic information that will define the formation of the new category through statistical distributions. Both models also stipulate that the more dissimilar the L1 and L2 sounds are, the more likely a new category will be formed. This is consistent with the position that L2 sounds are mapped on to phonetically similar L1 phonetic categories. Should the L2 sound be an extremely poor exemplar of the L1 category, it would be reasonable to create a new L2 category for that dissimilar sound.

An alternative view to the SLM(-r) is the perceptual assimilation model for second language learners (PAM-L2, Best and Tyler 2007). Where the SLM(-r) focused on the mapping of individual phones from L2 to L1, PAM-L2 is concerned with mapping contrasts from L2 to L1. According to PAM-L2, each phone in the L2 can either be categorized or uncategorized. Categorized phones can be perceived as good representatives of their native category, acceptable representatives, or poor representatives. Importantly, PAM-L2 posits that the L2 listener detects "higher-order phonetic properties that define category membership" (Tyler 2021: 2).

According to PAM-L2, when a non-native contrast is perceived as categorized, it can be assimilated to two categories, in which case discrimination is expected to be very good. Alternatively, the two sounds could be assimilated to a single category, and discrimination is likely to be poor. A category-goodness assimilation occurs when an L2 contrast is assimilated to a single category but one phone has a higher goodness of fit than the other phone. In these cases, PAM-L2 predicts a higher likelihood of a new category formation.

The present study examines the potential interference of L1 English allophony of pre-nasal consonant oral vowels (VN) with the perception of L2 French contrastive nasal vowels (Ṽ). Per the SLM(-r), L1 English speakers should have a position-specific phonetic category that encompasses oral vowels before nasal consonants and is cued by anticipatory coarticulatory nasality. This model would predict that learners will initially interpret vocalic nasality in the L2 signal (Ṽ) as mapping onto their preexisting phonetic category (VN). The model would also predict that learners are able to gradually create a new category, assuming that the sounds are phonetically dissimilar enough. Within PAM-L2, the French nasal–oral vowel contrast could be perceived as assimilating to a single category within the L1, indicating that there is no phonological distinction that can be used to discriminate the two phones. However, this single-category assimilation is likely to be a category-goodness assimilation in which the French oral vowels are a better fit to the native category than the nasal vowels. As mentioned above, French nasal vowels exhibit spectral differences compared to their phonological, oral counterparts such that /ɛ̃, ɑ̃, ɔ̃/ are realized phonetically as [æ̃, ɔ̃, õ]. This spectral difference could serve as a sufficient cue to create a new category apart from the [NASAL] feature. PAM-L2, importantly, does not overtly make claims about positionality of phones, so it is unclear how it would handle assimilation of French Ṽ to position-specific VN in English.

Several studies have shown that native language allophonic sounds are more challenging to discriminate than contrastive sounds (e.g., Boomershine et al. 2008). Since these phonetically distinct sounds map onto the same phonological category, they are effectively treated at some level as the same sound. This would suggest that the English nasal allophony might render [V] and [Ṽ] indistinguishable. Using a lexical decision with repetition priming task, Pallier, Colomé, and Sebastián-Gallés (2001) found that bilingual Spanish-Catalan speakers who were Spanish-dominant exhibited repetition priming for Catalan words that exhibited a non-Spanish contrast like /netə/ ~ /nɛtə/. The participants treated words with an /e/ ~ /ɛ/ contrast as homophonous. Catalan-dominant bilinguals exhibited no repetition priming for the same contrast since the two words are distinct in Catalan. Pallier et al. (2001) conclude that word recognition involves a language-specific phonological representation, such that the dominant language phonology influences lexical access in the nondominant language. These results show that even with a high degree of fluency in the second language, as was the case with the Spanish-dominant bilinguals, speakers still mapped the two sounds onto a single L1 category.

Alternatively, Beddor and Strange (1982) examined native speakers of English and of Hindi, languages in which the nasal–oral relationship for vowels is allophonic and contrastive, respectively. The researchers created synthetic productions that systematically ranged from /ba/–/bɑ̃/ and had their participants discriminate and identify the stimuli. Hindi speakers – whose native language has a nasal–oral contrast in both consonants and vowels – were able to discriminate categorically between each of the pairs. English speakers *were* able to discriminate between the nasal and oral vowels; however, the discrimination pattern was more continuous than categorical. Discussing this study, Martinez, Goad, and Dow (2021: 4) suggest that this was due to the perception of a nasal consonant after the nasal vowel, making the contrast more English-like V ~ VN rather than the less English-like V ~ Ṽ. The scenario in which [Ṽ] is mapped onto [VN] is consistent with the SLM(-r) because of the claim that it is position-specific allophone categories that are mapped onto. Since the vowel in an English VN sequence exhibits significant coarticulatory nasality, learners map the L2 Ṽ onto the phonetic category of VN.

Weber and Cutler (2004) and Cutler, Weber, and Otake (2006) present interesting results from experiments on native Dutch listeners of English and native Japanese listeners of English, respectively. While the SLM(-r) predicts that similar

L2 sounds map onto preexisting L1 phonetic categories, Weber and Cutler (2004) and Cutler et al. (2006) show that this mapping could be asymmetric in nature. When a pair of L2 sounds is mapped onto the same L1 category, these studies show that one of the sounds proves to be more dominant than the other, similar in nature to PAM-L2's category goodness assimilation. It is often claimed that L1-Dutch L2-English learners map the English vowels [ɛ] and [æ] onto a single L1 phonetic category. However, in an eye-tracking study, Weber and Cutler (2004) revealed asymmetric competitor activation such that target *panda* caused [ɛ]-competitor activation while the target *pencil* did not cause [æ]-competitor activation. This asymmetry was replicated in Cutler et al. (2006) with perception of the English [r]/[l] contrast by native Japanese speakers. Results from the eye-tracking study showed that /r/-initial targets induced competitor looks to /l/-initial words, while /l/-initial words did not cause competitor looks to /r/-initial words. The lack of symmetry, which would have been expected if the learners treated /r/-initial and /l/-initial words as homophonous, suggests a more complicated mapping from L2 to L1 phonetic categories. Unlike PAM-L2's category goodness assimilation pattern, learners in Cutler et al. (2006) distinguished between two phones at the phonetic level based on acoustic-phonetic proximity to the L1 category. Additionally, this distinction, though extant at an acoustic-phonetic level, did not rise to the level of lexical access and word recognition.

With respect to change over time in perception of L2 sounds, the SLM(-r) and PAM-L2 posit that new sounds *can* be learned, and new categories *can* be created. For the SLM(-r), it is possible if the two sounds in question are phonetically sufficiently distinct; for PAM-L2, it is possible if the category goodness rating of one phone is higher than that of the other. Evidence of learning a new contrast is shown in Herd, Jongman, and Sereno (2013), which investigated the [d]/[ɾ] contrast in L1 English learners of Spanish and found that perception of segments that are allophonic in the L1 but contrastive in the L2 can in fact improve with training. The learners scored the lowest in the perception pretest for the [d]/[ɾ] contrast, but after perception training, their scores improved in the posttest. The perception training consisted of participants completing a forced choice identification task in which they selected which of a minimal pair (e.g., *loro* 'parrot' vs. *lodo* 'mud') they heard. After each selection, participants were provided with feedback and then heard the token repeated regardless of accuracy. After six training sessions held over the course of 2–3 weeks, learners *did* improve in their perception of these two segments on the posttest even though they are allophonic in their native language. This suggests that though initially very difficult to perceive, native allophonic segments can be perceived as contrastive with more targeted exposure to this contrast.

2.2 Abstract underlying phonological representations

A complete understanding of the process of speech perception and word recognition requires determining how gradient acoustic cues are mapped onto discrete categories within the mind of the speaker. Some approaches (e.g., Lahiri and Jongman 1990) have posited that listeners extract *features* from the acoustic signal that then map directly onto lexical items. The competition occurs when more than one lexical representation shares the extracted feature. When [nasal] occurs within the signal, its presence activates many competitors that all share the feature [nasal].

Lahiri and Marslen-Wilson (1991) is one of the most influential studies on the perception of vowel nasality. To determine how listeners interpret featural information in the acoustic signal, they investigated perception of oral and nasal vowels in a variety of sequences for both native English and native Bengali speakers. Their hypothesis was rooted in the concept of underspecification: listeners use an abstract, underspecified representation to interpret the incoming signal as opposed to using the fully specified acoustic surface representation. Lahiri and Marslen-Wilson (1991) envisaged a process of extracting information from the acoustic signal in a nonlinear, continuous fashion. Specifically, the surface vowel [ã] might cause the listener to extract the feature [NASAL] from the signal; their hypothesis was that this feature would activate the lexical representations that directly match this feature as well as those that simply do not mismatch. To test this hypothesis, Lahiri and Marslen-Wilson examined the distribution of the feature [NASAL] in native English and native Bengali because in these two languages the same cue, nasality, is present in the surface representation for both underlyingly nasal and oral vowels in Bengali and for just underlyingly oral vowels in English. This provided an interesting testcase for determining what speakers do with the signal as it comes in.

For each language, Lahiri and Marslen-Wilson (1991) only looked at sequences found in the native language. For English speakers, they looked at CVN and CVC sequences; for Bengali speakers, they examined CVN, CVC, and CṼC sequences. They used a gating methodology, which the present study largely adopts. To be systematic about responses relative to vowel onset and offset, Lahiri and Marslen-Wilson used 40 ms gates with gating always beginning at the fourth glottal pulse after the onset of the vowel and a final gate being added at vowel offset. Therefore, the first and final gates were variable in length, dependent on the stimulus. When the vowel offset occurred within 10 ms (within one glottal pulse) of the previous gate, the last gate was simply lengthened to end at the vowel offset. When the vowel offset occurred more than 10 ms after the end of the previous gate, an additional gate was added which ended at the offset of the vowel. After the vowel portion of the stimulus, the gating continued at 40 ms increments until the end of the word. Lahiri and Marslen-Wilson presented the gates with an interstimulus interval (ISI) of 6,000 ms between each gate, with a single tone playing as a warning for the next gate and a double tone serving as a warning for a new word.

Bengali speakers showed a clear preference for selecting CṼC responses over CVN responses when they encountered phonemic nasality in the signal (56.8% CṼC responses, 5.2% CVN responses to CṼC stimuli). The authors consider this finding to be in line with the underspecification mapping to the recognition lexicon because if the fully specified surface representation was used, Bengali listeners would have shown ambiguity between CVN and CṼC responses. Instead, they associated the nasality in the signal with mapping to an underlyingly nasal vowel and not an oral vowel that is nasalized by way of assimilation. In the trials in which the stimulus was CVN, it was not until the offset of the vowel, where more of the coarticulatory information from the following nasal consonant became evident, that Bengali listeners switched responses from CṼC to CVN.

Comparatively, English speakers showed much more ambiguity in their responses to CVN stimuli, providing CVC responses 59% of the time, while providing CVN responses 41% of the time. When responses are analyzed per gate, it can be clearly seen that CVN responses increase as more of the vowel is presented. According to an abstract underspecification account, this ambiguity is expected as the vowel nasalization is interpreted as evidence of the upcoming nasal consonant. However, as Lahiri and Marslen-Wilson point out, these findings can also be accounted for by the fully specified surface account in that there is nasalization in the signal, which can be used to discriminate between CVN and CVC words.

Since the present study focuses on French as the target L2, it is necessary to know how native French speakers might perform in a similar gating task. Ingram, Park, and Mylne (1997) conducted an experiment similar to Lahiri and Marslen-Wilson (1991) with native French speakers to determine if the difference in organization from Bengali to French, namely the suppression of regressive nasal assimilation, would influence the way speakers responded to gated nasal–vowel and oral–vowel stimuli. In Ingram et al.'s (1997) study, participants heard gated stimuli from seven CVC, CVN, and CṼC triplets. The methodology of this study followed that of Lahiri and Marslen-Wilson with gates beginning at the fourth glottal pulse and continuing every 40 ms thereafter. The entire word was gated, and participants heard all gates.

When presented with a CVC stimulus, French speakers responded overwhelmingly with CVC responses (93%). When presented with CṼC stimulus, the responses were split evenly between CVC (49%) and CṼC (48%). Last, when presented with CVN, speakers selected CVN only 6% of the time but responded CVC 86% of the time. Ingram et al. (1997) conclude that even though French has little coarticulatory nasalization, the native French speakers perceive nasality "virtually identically" to the Bengali speakers of Lahiri and Marslen-Wilson (1991). The predominant response when there was no nasality in the signal by native speakers of French was CVC.

When nasality was present in the signal, French speakers clearly favored CṼC over CVN responses. These findings are consistent with Lahiri and Marslen-Wilson's (1991) underspecification lexicon. If the feature [NASAL] is specified for CṼC but underspecified for CVC and CVN, we would expect the overwhelming CVC responses when nasality is not present and equal CVC and CṼC responses when nasality is present. When the feature [NASAL] is present, it presents a no-mismatch case with CVC words which are unspecified, which explains the relatively high number of CVC responses to CṼC stimuli. While this was not addressed in Ingram et al. (1997), the lack of CVN responses to all stimuli might be reflective of the relative low token-frequency of CVN words in the language. Considering there is little coarticulatory nasality and no phonological allophony in French CVN sequences, it is possible that listeners treat CVN as a subset of CVC, which explains the fewer responses.

The present study is closely related to the findings of Lahiri and Marslen-Wilson (1991) and Ingram et al. (1997), as it aims to perform a similar task but in a second language, to determine whether the responses are consistent with L1 phonological distribution of nasal vowels or L2 phonological distribution.

2.3 Underspecification and nasal vowels

Lahiri and Marslen-Wilson's (1991) findings support an underspecified lexicon. In the literature, nasal vowels are either treated as equipollent with nasal segments being treated as [+nasal] and oral segments being treated as [–nasal] or privative with only nasal segments being specified for [NASAL]. More recent work favors privative features, specifically supporting the idea of underspecification. In a privative feature system, nasal vowels and consonants are specified as [NASAL], while oral vowels and consonants remain unspecified. Being unspecified, oral segments are not treated as a set or group that shares common features or undergoes similar processes (Kotzor et al. 2022). In an underspecification lexicon, the absence of a feature in the signal would not result in a mismatch to a segment that contains that feature; rather it would be categorized as a "no-mismatch." In an equipollent system where both nasality and orality are specified as [+nasal] and [–nasal], respectively, one would predict different results when nasality is not present in the signal because [–nasal] is a direct "mismatch" to [+nasal] rather than a "no-mismatch."

Kotzor et al. (2022) conducted a study with Bengali native speakers to determine whether nasality is featurally represented as equipollent or privative and how native listeners utilize phonetic and phonological nasality in processing. They used a cross-modal (audio-visual) priming task with a forced-choice response. The researchers laid out clear predictions for both a privative approach which assumes underspecification and for a surface phonetic approach that assumes equipollent features based on the phonetic information in the signal. If the experiment supported the surface phonetic account, the researchers expected faster latencies only when the prime was *identical* to the target in onset (e.g., CṼ prime would match [CṼC] and [CṼN] target) regardless of underlying representation. Kotzor et al. (2022) posit that for the privative system, there are only two scenarios that would be a direct match from prime to target and thus result in faster latencies and lower error rates: /CṼC/ and /CVN/ (surface: [CṼN]) primes match to CṼC targets. Every other combination of prime-target pairing should result in a no-mismatch condition.

To test these predictions, participants were asked to indicate if the prime they heard belonged to the left or right word that appeared on a screen in Bengali script. Results showed significant support for the privative, underspecified system. The prime-target match pair that resulted in the fastest latencies was the predicted CṼC identity pair. Additionally, faster latencies were found with [CṼN] primes to CṼC targets, as predicted. Due to the results of both Lahiri and Marslen-Wilson (1991) and Kotzor et al. (2022), the present study assumes the underspecification model and the predictions made in Section 3.4 are based in an underspecified lexicon.

3 The present study

Considering the literature reviewed above, the present study aims to add to the current knowledge by addressing the following research questions:

1. How do L1 English, L2 learners of French (L2ers) utilize nasality in the incoming signal for nasal vowels? Do L2ers interpret the nasality in light of their L1 distribution of nasal vowels, attributing the nasality to an upcoming consonant? Or are L2ers able to assign the nasality to the underlying nasal vowel?
2. Does more exposure to the L2 attenuate the effects of transfer? Will proficiency in the L2 predict accurate identifications in the gating task?

To address these questions, a gating study was conducted to investigate specifically how L2ers perceive the French sequences CVC, CVN, and CṼ.

3.1 Predictions

Following the underspecification account provided in Lahiri and Marslen-Wilson (1991) and Kotzor et al. (2022), predictions are outlined below using a match/mismatch/no-mismatch framework. It is expected that when there is a “mismatch,” activation of the competitor will not occur and few responses of that competitor will be given. When the input signal does not directly mismatch the target/competitor (i.e., a “match” or “no-mismatch”), it is expected that both of these responses will be given. It also stands to reason that a “match” competitor would receive more responses than a “no-mismatch” competitor.

Under the SLM(-r), if L1 English L2 French learners perceive nasality through their L1 phonological system, surface nasality would activate the specified feature of [NASAL] for the following nasal consonant. The prediction from SLM(-r) would be that L2 nasal vowels map onto the L1 *position-specific allophone* of a pre-nasal consonant oral vowel (VN). PAM-L2 makes no specific claims on positionality of contrasts, which makes it challenging to know what it would predict with respect to the Ṽ ~ VN contrast. Table 1 shows the match/mismatch/no-mismatch distribution expected from L1 English influence. Mapping the L2 nasal vowel onto the L1 VN allophone would result in a match with CVC when no nasality is present in the signal. Both CVC and CVN would have underlyingly underspecified oral vowels, which would match with orality in the signal. If the listeners interpret the nasality in CṼ as belonging to an upcoming nasal consonant, they would be likely to behave as the English speakers in Lahiri and Marslen-Wilson (1991), who were predicted to experience a no-mismatch to CVC and CVN when nasality was present and a match to underspecified CVC and CVN when no nasality was present (Lahiri and Marslen-Wilson 1991: 262). CVC should also be the predominant response regardless of stimulus, similar to English speakers in Lahiri and Marslen-Wilson (1991: 275) who responded to CVC stimuli with 83.4% CVCs and to CVN stimuli with 59.3% CVCs, presumably because of the no-mismatch mapping of the underspecified oral vowel.

Table 1: Predicted perceptual match/mismatch if L2 learners perceived the contrast as monolingual native English speakers (similar to those in Lahiri and Marslen-Wilson 1991).

			Predicted matches		
Stimulus	**Underlying**	**Surface nasality**	**/CVC/**	**/CṼ/**	**/CVN/**
CVC	/CVC/	–	Match	Match	Match
CṼ	/CVN/	+	No-mismatch	No-mismatch	No-mismatch
CVN	/CVN/	–	No-mismatch	No-mismatch	No-mismatch

If the L2 sounds (Ṽ and VN) are mapped onto the single L1 phonetic category, that means that when nasality is present in the signal, CṼ and CVN should essentially be treated as homophonous, much like Spanish-Catalan bilinguals did with /e/ ~ /ɛ/ in Pallier et al. (2001). PAM-L2’s category goodness assimilation pattern might reflect more of an asymmetry of the variety seen in Weber and Cutler (2004) and Cutler et al. (2006), where the phone that was phonetically closest to (or in PAM’s terms the best fit to) the native category is preferred over the less similar. In this case, perception would depend on which phone is considered dominant, that is, more phonetically similar to the L1 category. If learners map L2 sounds V(N) and Ṽ onto the same native category, there is reason to believe that the Ṽ would be the more similar phone, as it has more nasality in the signal than V(N) which is expected of an English vowel in a VN sequence. However, due to the formant shifts discussed previously, it could be that the vowel in the French VN sequence is phonetically closer to the L1 category. That is, the formant structure of French /ɛ/ in /ɛn/ might be closer to English /ɛ/ in /ɛn/ because the formants shift in /ɛ̃/ to become more like [æ̃].

Proficiency is also expected to play a role in the perception of French nasal vowels. The SLM(-r) posits that new categories can be created assuming that the L2 and L1 sounds are sufficiently distinct. As mentioned above, there is reason to suggest that the Ṽ is sufficiently distinct from the vowel in a VN sequence to warrant creation of a new category. It is predicted that more proficient learners will have created a new category of nasal vowels, thus performing more like the native speakers of French with a specified [NASAL] category for Ṽ. This would manifest in the data as proficiency predicting the number of CVN responses to CṼ, such that as proficiency increases, the number of CVN responses to CṼ stimuli decreases.

3.2 Methodology

3.2.1 Participants

Fifteen L1 English learners of French (3 males, 12 females, aged 18–46) were tested. They went through a familiarization task and gating task. Participants were compensated with a $10 Amazon gift card for their participation. One participant was omitted from analysis due to being bilingual in English and Portuguese. Portuguese has a contrastive nasal–oral vowel system similar to French, which likely influenced this participant's results. We examined our participants' proficiency in French by using a proficiency task called LexTale-FR (Brysbaert 2013), in which participants were asked to determine whether a word was a real French word or not. Correct responses added one point each while incorrect responses subtracted one point each. The scores from the LexTale-FR test are used in analysis to determine if proficiency influenced the results of the gating study.

3.2.2 Stimuli

The stimuli for this experiment were recorded by a female native speaker of French from Chateau-Thierry, France, which is located in the North of France just outside the Paris Region. They consist of CVN, CVC, and CṼ (sometimes CṼC) French triplet sequences. All sequences are monosyllables, and some of the onsets or codas have more than one consonant. Crucially, however, each target triplet has the same onset and nucleus (apart from nasality of the vowel). The CVN sequences consist of an oral vowel followed by a nasal coda consonant (e.g., *veine* [vɛn] 'veine'). The CVC sequences consist of an oral vowel followed by a non-nasal coda (e.g., *veste* [vɛst] 'jacket'). The CṼ sequences consist of a nasal vowel (e.g., *vin* [vɛ̃], 'wine'); in some cases, there is a non-nasal consonant following the nasal vowel (e.g., *banque* [bɑ̃k] 'bank'). Each word is associated with a particular image used in the familiarization task and again in the gating task. Images were used to avoid too much focus on the orthography.

We used nine triplets, resulting in a total of 27 target stimuli. In addition to the target stimuli, filler stimuli were included. These filler stimuli consisted of five triplets with a non-nasal consonant and a high vowel, such as C[i] (e.g., *lit* [li] 'bed'), C[y] (e.g., *lu* [ly] 'read' past participle), and C[u] (e.g., *loup* [lu] 'wolf'), resulting in a total of 15 words. A complete list of stimuli as well as the images associated with them can be found in the Appendix in Tables 5–6. The native speaker was asked to record each stimulus (both filler and target) three times to obtain at least one quality recording for each word. The stimuli were presented in random order.

3.2.3 Familiarization task

Testing L2 learners from a variety of learning backgrounds as well as needing very specific triplets in the stimuli, it is difficult to know whether participants are familiar with the stimuli presented to them. Therefore, before completing the main gating task, a familiarization task was run to ensure that the participants knew the stimuli of the study. The stimuli were presented orthographically on the screen with the associated picture. The recording of the word was not played during this task to prevent practice effects or teaching to test effects. The familiarization task consisted of a practice phase and a testing phase. In the practice phase, participants were presented with a picture and visual word combination in standard orthography on the screen. Each word-picture combination was presented in randomized order.

In the testing phase of the familiarization task, participants were tested on their associations with the images. In this phase, participants saw the orthographic word they learned and then were instructed to select the picture associated with the word. For example, a participant saw “flamme” on the screen and then was presented with three possible pictures to choose from. The three options consisted of the correct picture and the other two stimuli that were a part of that triplet. The order of the three presented options was randomized in terms of position (left, center, right).

3.2.4 Gating task

We modeled the gating task after Lahiri and Marslen-Wilson (1991). As such, the recorded stimuli were broken up into gates of 40 ms starting at the fourth glottal pulse of the vowel. All 27 target stimuli had at least 5 gates, while some had 6 (n = 11) and some had 7 gates total (n = 2). The gates were numbered $-n$ to 0 depending on the length of the stimuli, with 0 being the offset of the vowel. Based on Lahiri and Marslen-Wilson’s methodology, a boundary was always set at the offset of the vowel, so in cases where the previous boundary ended less than 10 ms before the offset of the vowel, the length of gate 0 was increased to account for that. If the difference between the previous boundary and the vowel offset was longer than 10 ms, an additional gate was added. For example, Figure 1 shows a CVN stimulus recorded by the native speaker broken up into 40 ms gates. The interval between gates −1 and 0 is less than 40 ms but greater than 10 ms. At each gate, the participant heard more and more of the stimulus up until the offset of the vowel. Unlike in Lahiri and Marslen-Wilson’s experiment, the rest of the word was not played, which prevented the listeners from verifying answers and having practice effects. On average, it took participants 7 min to complete the background questionnaire, 5 min to complete the familiarization task in which participants were familiarized with the images and associated them with the correct words, 2.5 min to complete the testing of familiarization, and 16.5 min to complete the gating task, resulting in an average total time of 33 min.

Participants were asked to select from three options which picture they associated with what they were hearing at each gate. They were instructed to give their best guess even if they were unsure. A practice round of 10 trials was then given to participants to familiarize them with the format. After the practice trials, another instruction screen appeared, on which they had to click “Next” when they were ready to begin the experiment. The task itself consisted of a fixation cross that stayed on screen for 250 ms and then a screen on which three images were presented while the audio was played. The fixation cross served as a cue for the beginning of a new trial. The picture selection screen remained until participants made their selection.

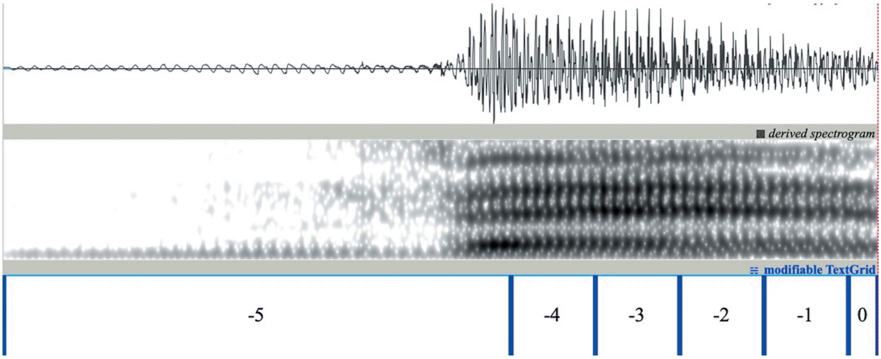


Figure 1: Example waveform and spectrogram of the CV portion of the French word "veine" with gates beginning at the fourth glottal pulse and continuing every 40 ms (final gate is 14 ms).

4 Data analysis

4.1 Acoustic analysis

All measurements and stimulus preparation took place using Praat speech analysis software (Broersma and Weenink 2022). We measured duration, F1, F2, and F3 and nasality (A1-P0, Styler 2017). The average duration for the CVC stimuli was 662 ms, with the average vowel duration being 178 ms (average number of gates: 5.22). For CVN stimuli, the average total duration was 598 ms, with the average vowel duration being 187 ms (average number of gates: 5.33). For CṼ stimuli, the average total duration was 485 ms, with the vowel duration averaging at 209 ms (average number of gates: 6.11). The nasal vowels in the CṼ sequences were expected to be slightly longer because they were in open syllables.

Formant values were calculated using a Praat script (Reetz 2020), which computes the mean formant values at the midpoint of the specified vowel. Figure 2 shows the F1/F2 vowel space for [a, ɛ, ɔ] in CVC (pink) and CVN (green) sequences as well as [ɑ̃, ɛ̃, ɔ̃] in CṼ (blue) sequences. As expected from the literature, the F1 values of the nasal vowel [ɛ̃] (blue triangle) are higher and F2 values are lower indicating a more retracted and lower production than the oral counterpart (Dow 2020: 241). Additionally, the nasal vowel [ɑ̃] (blue circle) shows lower F1 and F2 values as well, indicating that a more accurate phonetic transcription would resemble the oral vowel [ɔ] (pink and green squares). This difference in formant values from oral to nasal vowels is consistent with the findings from previous literature, in which French nasal vowels are produced with a more retracted tongue for the mid front vowel /ɛ̃/ and a lower tongue with greater lip rounding for the mid and low back vowels /ɑ̃, ɔ̃/ (Carignan 2014: 25; Dow 2020: 241). Our speaker, however, does not seem to have a large spectral difference, at least with respect to F1/F2, between /ɔ̃/ (blue square) and /ɑ̃/ (blue circle).

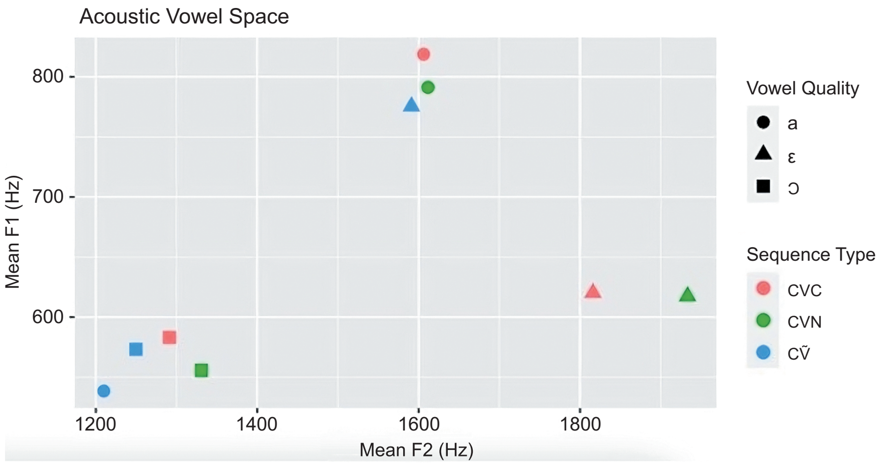


Figure 2: F1/F2 vowel space for the native speaker's nasal vowels and oral counterparts. Shapes represent the vowel quality. Colors represent the sequence in which the vowel appeared.

Nasality was measured using A1-P0 (Styler 2017). A1-P0 is an often-used measure that takes the amplitude of the highest harmonic in the first formant frequency (F1) and subtracts from it the amplitude of the nasal pole that typically occurs in the low frequency range (250–450 Hz; Styler 2017: 2470). The opening of the velopharyngeal port introduces poles and zeros associated with the nasal cavity causing the amplitude of the nasal pole (P0) to increase while lowering the amplitude of F1. Thus, the nasality measure of A1-P0 will be smaller when the opening to the velopharyngeal port is larger and the vowel is more nasal. These measures are derived via spectral analysis at a selected timepoint within the vowel. For the purposes of the present study, A1-P0 was calculated at the 75% mark of the vowel in each gate. This was done to get an understanding of how the nasality changed over the course of the gates and how much nasality each gate introduced to the listener. A Praat script (Styler 2018) was used to derive the measure with any errors being flagged and manually calculated. The data was then aggregated by Gate and Sequence Type. Figure 3 shows the amount of nasality in the vowels of each Sequence Type plotted across each gate.

The CṼ sequences (e.g., *vin* [vɛ̃] 'wine') exhibit lower A1-P0 values (more nasality) from the earliest gates, while the CVN sequences (e.g., *veine* [vɛn] 'vein') show

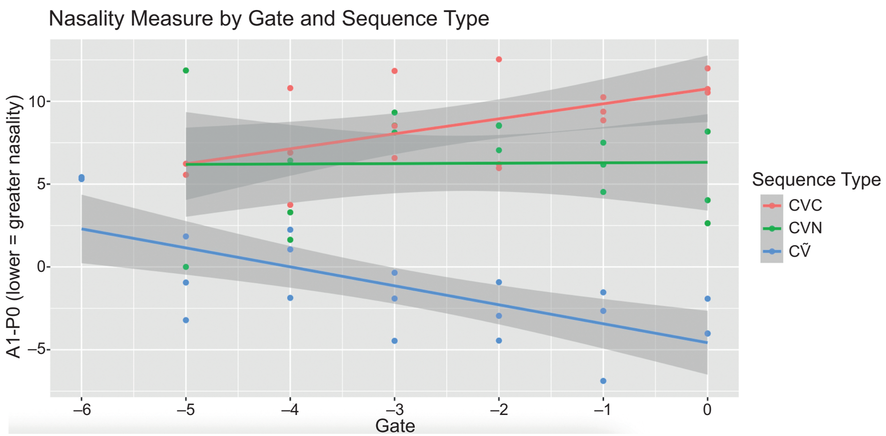


Figure 3: A1-P0 plotted as a function of Gate and Sequence Type.

less nasality than CṼ sequences, but more nasality than CVC sequences (e.g., *veste* [vɛst] 'jacket'). Consistent with Dow's (2020) findings, the nasality of the CVN does not dramatically increase toward the end of the vowel, which suggests little coarticulatory nasality. Expectedly, there is some overlap at the early gates between the oral vowels in CVC and CVN sequences, which supports the conclusion that these are both phonemically oral.

4.2 Statistical analysis

In addition to qualitative analyses of percent Responses by Sequence Played and by Gate, we performed two statistical analyses using R programming language (R Core Team 2022). The first of which was a generalized logistic mixed-effects model to understand the overall accuracy of the participants (termed "Accuracy Model"). For this analysis, the glmer() function was used from the lme4 package in R (Bates et al. 2015). Using a forward stepwise approach, we started with an empty model that included the dependent variable of Accuracy and the random effect of Participant. Gradually, each fixed effect was added, and a likelihood ratio test was run to determine if the additional fixed effect bettered the model. The final model included the dependent variable of Accuracy, the fixed effects of Sequence Played, Gate, and Proficiency.

We were interested in understanding what choices participants made when they were incorrect. When presented with a CṼ stimulus, were they more likely to incorrectly select CVC or CVN and how was this affected by Gate? As a result, we ran a multinomial logistic mixed-effects model that included the dependent variable of Response and the random effect of Participant (termed "Response Model"). This model required the use of the mblogit() function from the mclogit package (Elff 2022). Once again, the fixed effects of Sequence Played and Gate were added in a forward step-wise fashion. The interaction of Sequence Played and Gate significantly improved the model compared to the noninteraction mode.

5 Results

5.1 Accuracy model

The results of the generalized linear mixed-effects model, which included Accuracy as the dependent variable and Sequence Played and Proficiency as the fixed effects and Participant as a random effect, are presented in Table 2. Because the fixed effect of Sequence Played is a three-level factor (CVC, CVN, CṼ), the model was run with CVN as the reference level.

Table 2: Results from the generalized linear mixed-effects model.

Predictors	***β***	***SE***	***p***	***Sig. Level***
(Intercept)	−1.28	0.209	<0.001	***
Sequence Played [CVC]	0.15	0.113	0.175	
Sequence Played [CṼ]	1.50	0.121	<0.001	***
Gate	0.18	0.0297	<0.001	***
Proficiency	0.01	0.005	0.0153	*

Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Positive β-coefficient is associated with an increase in accuracy.

The simple effect of Sequence Played indicates that CṼ was more accurately selected compared to CVN. A releveling of the model to CVC showed that CṼ was also more accurately selected compared to CVC ($\beta = 1.34$, $p < 0.001$). Plotting out percent responses by Sequence Played verifies this finding (Figure 4), showing that participants were much more accurate at identifying CṼ compared to CVC and CVN.

A significant simple effect of Gate indicates that as the gate increased (more of the word was played), the accuracy increased. This finding is expected because as the gates increase, participants have more information on which to base their responses.

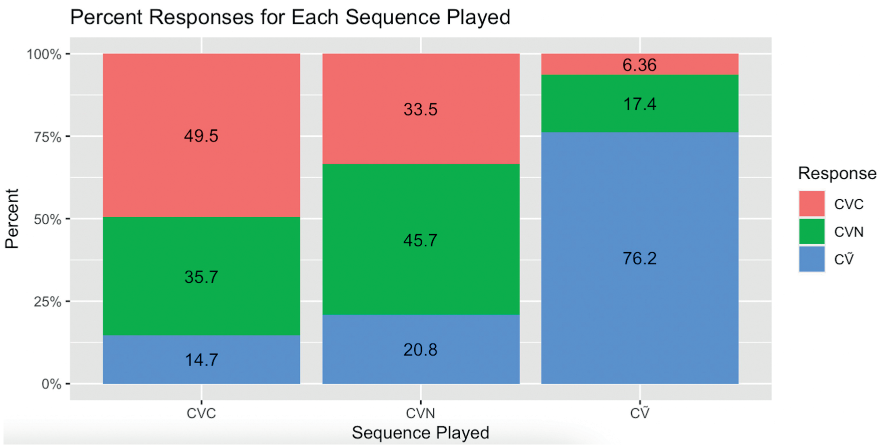


Figure 4: Percentage of responses for each Sequence Played.

Last, the Accuracy model showed a significant effect of Proficiency: as Proficiency increased, Accuracy in identification increased as well. Participants took the LexTale-FR proficiency task and received a score by tallying the total number of correct responses and subtracting the total number of incorrect responses (Brysbaert 2013). The range of participants' LexTale-FR scores was –2 to 74 with a mean of 17.5 and a median of 9. Some scores were negative as participants were penalized for incorrect answers to avoid yes or no biases. Figure 5 shows percent responses to

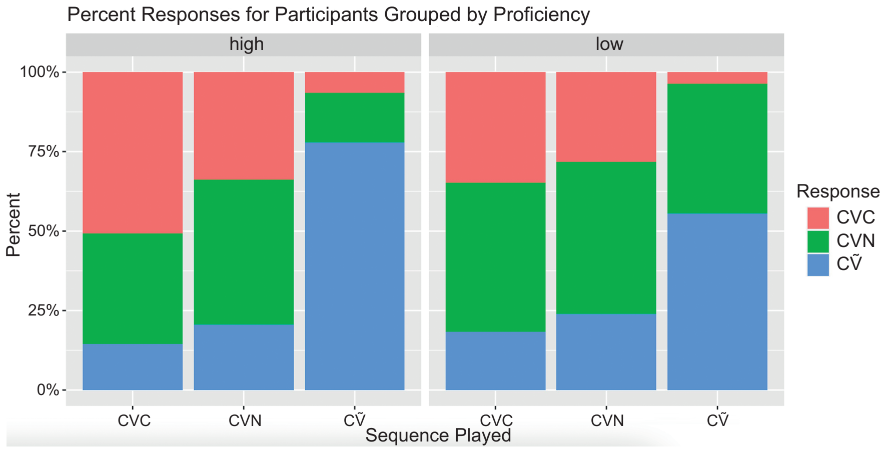


Figure 5: Percent responses for each vowel sequence played for participants grouped by proficiency (high-proficiency on left, low-proficiency on right).

each Sequence Played grouped by Proficiency (high and low). This figure clearly shows that high-proficiency participants were more accurate at identifying CṼ and CVC sequences.

5.2 Response model

The results of the multinomial logit mixed-effect model, which included Response as a dependent variable and Sequence Played and Gate as fixed effects and Participant as a random effect, are presented in Table 3. Again, since Sequence Played is a three-level categorical variable, the model was run with reference set to CVC.

The model showed an effect of Sequence Played at varying levels of significance depending on the reference of both Response and Sequence Played. The findings from the Accuracy model are replicated here, such that the learners were more likely to select accurate Response CṼ for a CṼ Sequence as opposed to inaccurately selecting CVN or CVC (ref = CVN: row 3, Table 3B, and ref = CVC: row 3,

Table 3: Results from the multinomial logit mixed-effects model.

	A. Response: CVN versus CVC				
	Predictors	***β***	**SE**	***p***	**Sig. level**
1	(Intercept)	0.965	0.303	0.0014	**
2	Sequence Played [CVN]	−1.432	0.418	<0.001	***
3	Sequence Played [CṼ]	−0.453	0.513	0.377	
4	Gate	−0.264	0.058	<0.01	*
5	Sequence Played [CVN]: Gate	0.424	0.082	<0.001	***
6	Sequence Played [CṼ]: Gate	0.384	0.111	<0.001	***
	B. Response: CṼ versus CVN				
	Predictors	***β***	**SE**	***p***	**Sig. level**
1	(Intercept)	−0.552	0.389	0.156	
2	Sequence Played [CVN]	0.949	0.508	0.062	.
3	Sequence Played [CṼ]	1.851	0.462	<0.001	***
4	Gate	−0.076	0.082	0.350	
5	Sequence Played [CVN]: Gate	−0.174	0.106	0.101	
6	Sequence Played [CṼ]: Gate	0.117	0.098	0.229	
	C. Response: CṼ versus CVC				
	Predictors	***β***	**SE**	***p***	**Sig. level**
1	(Intercept)	0.414	0.393	0.293	
2	Sequence Played [CVN]	−0.480	0.514	0.350	

Table 3 (continued)

	C. Response: CṼ versus CVC				
	Predictors	***β***	**SE**	***p***	**Sig. level**
3	Sequence Played [CṼ]	1.400	0.534	0.009	**
4	Gate	−0.339	0.080	<0.001	***
5	Sequence Played [CVN]: Gate	0.249	0.106	0.019	*
6	Sequence Played [CṼ]: Gate	0.502	0.116	<0.001	***

Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Positive *β*-coefficient is associated with an increased likelihood of selecting the bolded response over the nonbolded response. Reference for Sequence Played is CVC.Positive *β*-coefficient is associated with an increased likelihood of selecting the bolded response over the nonbolded response. Reference for Sequence Played is CVC.

Table 3C). Additionally, the likelihood of accurately selecting Response CVN for a CVN Sequence Played significantly increased (row 2, Table 3A). However, when the Sequence Played was CṼ, the results show a nonsignificant increase in the likelihood of selecting CVN over CVC (row 3, Table 3A), suggesting that learners did not select CVN significantly more than CVC when nasality was present in the signal.

The model also shows an interaction between Gate and Sequence Played, depending on the reference level of Sequence Played. When the Sequence Played is CVN (reference = CVC), an increase in Gate (i.e., more of the vowel being played) is associated with an increase in accurately selecting CVN over CVC (row 5, Table 3A). That is, as more of the vowel became available, learners were more accurate at selecting CVN for CVN stimuli. Additionally, when the Sequence Played is CṼ (reference = CVC), an increase in Gate is associated with an increase in accurately selecting CṼ over CVC (row 6, Table 3C), meaning accuracy in selecting CṼ increased as more of the vowel became available. Importantly, when the Sequence Played is CṼ, an increase in Gate is associated with a significant increase in selecting CVN over CVC (row 6, Table 3A). This indicates that the CVN became more of a competitor for CṼ sequences as more of the vowel was presented to the participants.

6 Discussion

The results from the present study showed that participants were highly accurate at identifying the CṼ stimuli, even from the earliest gates. This was unexpected given the previous literature. It was predicted that learners would behave as native English speakers and assign the nasality in the signal to an upcoming nasal consonant, making CVN and CṼ essentially homophonous and equally likely responses. The results show that the L2 learners in this study behaved as neither L1 English speakers nor L1 French speakers. Table 4 summarizes the English speakers' results from Lahiri and Marslen-Wilson (1991), French speakers' results from Ingram et al. (1997), and the present study's participants' results.

Table 4: Percentage of responses per stimulus type up to vowel offset for English natives (from Lahiri and Marslen-Wilson 1991), French natives (from Ingram et al. 1997), and L1 English L2 French learners from the present study.

			Response	
English speakers (Lahiri and Marslen-Wilson 1991)	Stimulus	CVC	CVN	
	CVC	**83.4**	16.6	
	CVN	59.3	**40.7**	
			Response	
French speakers (Ingram et al. 1997)	Stimulus	CVC	CVN	CṼ
	CVC	**93**	3	4
	CVN	86	**8**	6
	CṼ	49	3	**48**
			Response	
L1 English, L2 French learners (present study)	Stimulus	CVC	CVN	CṼ
	CVC	**49.5**	35.7	14.7
	CVN	33.5	**45.7**	20.8
	CṼ	6.4	17.4	**76.2**

The results from the present study show a very different pattern when compared to the native speaker results seen in both English and French. If the learners had behaved as predicted by the SLM(-r) (Flege 1995; Flege and Bohn 2021), we would have expected the L2 sound Ṽ to be mapped onto the position-specific allophone of VN in the L1. As mentioned above, this would render both [CṼ] and [CVN] as underlyingly /CVN/. The expected pattern of results would have been equally likely CVN and CṼ responses to CṼ stimuli. Separating participants out by proficiency level as well as the finding from the Accuracy Model did suggest that as proficiency increased, accuracy in selecting CṼ increased. The low-proficiency participants *did* exhibit much more competition with CVN and CṼ (see Figure 5). This suggests that for low-proficiency participants, nasality in the signal *was* interpreted as potentially cuing up a nasal consonant, rendering CVN and CṼ equally likely. However, for high-proficiency participants, CṼ was overwhelmingly selected in response to CṼ stimuli. This is indicative of a change in the mapping of the L2 sounds, however not in a native-like way in which the specified feature [NASAL] is the cue to the contrast.

One possible explanation for these findings is found in the acoustic-phonetic differences between nasal vowels in English and nasal vowels in French. As mentioned earlier, nasal vowels in French are not *nasalized oral vowels*, but rather formant values shift causing /ɛ̃, ɑ̃, ɔ̃/ to surface more as [æ̃, ɔ̃, õ], respectively. The SLM(-r) and the PAM-L2 suggest that new category formation can occur if the L2 sound is sufficiently distinct from the L1 category in phonetic characteristics or in category goodness rating, respectively. It is possible that surface nasal vowels in French are phonetically different enough from the English nasalized oral vowel along the spectral dimension that advanced learners have posited a new category on the basis of spectral cues rather than nasality. This would explain why CṼ stimuli were accurately selected from the earliest of gates and did not change as a function of Gate (Table 3B). Considering the reliance on spectral cues in English vowel contrasts (Neary 1989), an interesting continuation of this line of research would be to test advanced learners of French using manipulated stimuli that vary systematically in their nasality and their spectral changes to see which of the cues they rely on.

The present study raises important questions about the role of phonetic cues in phonological contrast and specifically how these contrasts are mapped from L2 phones to L1 categories. Phonetically, the nasality cue is utilized in both languages. In English, it is utilized by way of coarticulatory nasality to signal an upcoming nasal consonant. In French, it is utilized contrastively to distinguish between phonemes at the lexical level. It is the phonological realization of the same phonetic cue that presents learners with so much difficulty. The advanced learners of French in the present study may have found a way around that difficulty by utilizing a different cue – vowel quality – to create a new category.

The feature in question is [NASAL], which is proposed to be specified when the segment is nasal but unspecified when the segment is oral (following Lahiri and Marslen-Wilson 1991). The interesting yet complicating factor of this study is that this is *second language* perception. Models for second language speech perception have relied on mapping *sounds* from L2 to L1 categories (either phonetic per SLM (-r) or phonological per PAM-L2). The present study suggests a mapping of the *feature* [NASAL], which is manifested in the L1 as activating an upcoming segment and in the L2 as activating the feature of the vowel itself or an upcoming consonant. The expectation was that English-speaking learners of French would interpret the feature [NASAL] as they would in their first language – as a specified feature for consonants – activating all the lexical items that either match or simply do not mismatch the extracted cue. The results suggested that low-proficiency learners *did* in fact process the feature [NASAL] as they would in their first language, activating CVN much more frequently than the high-proficiency learners. The high-proficiency learners, however, were extremely accurate at identifying the CṼ stimuli. The suggested interpretation of these results was a new category created on the basis of the vowel quality difference between oral and nasal in French. Within the framework of feature extraction and mapping to the lexicon, this would mean more than one feature is at play. While [NASAL] in a CṼ signal may exhibit a no-mismatch with CVN and CVC, learners could be extracting an additional feature from the vowel that *mismatches* with the CVN and CVC, resulting in fewer of these responses to CṼ stimuli. For example, the vowel /ɛ̃/, as mentioned above, surfaces with a lower and more back realization and could be transcribed as [æ̃]. This was evidenced in the acoustic analysis of the stimuli in this study (see Figure 3) where /ɛ̃/ surfaced with a higher F1 and lower F2, indicating a lower and more retracted pronunciation. Perhaps the surface realization of /ɛ̃/ caused the learners to extract a height or advancement feature that *mismatched* with the underlying oral vowel /ɛ/ in CVC and CVN.

7 Conclusion

In summary, the present study investigated second language speech perception of French nasal vowels utilizing a gating task modeled after Lahiri and Marslen-Wilson (1991). It was found that less proficient learners performed as predicted such that their first language influenced their perception of the nasal vowels. The mapping of L2 Ṽ onto L1 VN category was supported by low-proficiency participants' equal CṼ and CVN responses to CṼ stimuli. For high-proficiency learners, however, results show a pattern distinct from either native English speakers or native French speakers. It was posited that these high-proficiency learners were forming a new category for Ṽ, but that they might be relying on spectral cues in the oral–nasal vowel contrast rather than the nasal feature itself to form this new category. Further study is needed to delve into this issue and tease apart the nasal cues and the spectral cues in the learners' perception.

Appendix

Table 5: Stimuli divided up into Group 1 (CVN ~ CVC) and Group 2 (CVN ~ CṼC).

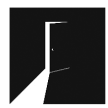
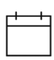
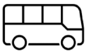
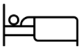
Sequence	CVC		CṼ (or CṼC)		CVN	
	Word	**Image**	**Word**	**Image**	**Word**	**Image**
Triplet 1 *[ɔ] ~ [ɔ̃]*	poste [pɔst] ‘post office’		pont [pɔ̃] ‘bridge’		pomme [pɔm] ‘apple’	
Triplet 2 *[ɔ] ~ [ɔ̃]*	sobre [sɔbʀ] ‘sober’		sombre [sɔ̃bʀ] ‘somber’		somme [sɔm] ‘sum’	2+2
Triplet 3 *[ɔ] ~ [ɔ̃]*	toque [tɔk] ‘chef’s hat’		tond [tɔ̃] ‘I mow’		tonne [tɔn] ‘metric ton’	
Triplet 4 *[a] ~ [ɑ̃]*	bac [bak] ‘baccalaureate’		banque [bɑ̃k] ‘bank’		banne [ban] ‘wicker basket’	
Triplet 5 *[a] ~ [ɑ̃]*	flaque [flak] ‘splash’		flan [flɑ̃] ‘flan’		flamme [flam] ‘flame’	
Triplet 6 *[a] ~ [ɑ̃]*	date [dat] ‘date’		dents [dɑ̃] ‘teeth’		dame [dam] ‘lady’	
Triplet 7 *[ɛ] ~ [ɛ̃]*	veste [vɛst] ‘jacket’		vin [vɛ̃] ‘wine’		veine [vɛn] ‘vain’	
Triplet 8 *[ɛ] ~ [ɛ̃]*	pêche [pɛʃ] ‘peach’		pin [pɛ̃] ‘pine tree’		peine [pɛn] ‘punishment’	
Triplet 9 *[ɛ] ~ [ɛ̃]*	reste [ʀɛst] ‘remaining’		rince [ʀɛ̃s] ‘I rinse’		reine [ʀɛn] ‘queen’	

Table 6: Filler stimulus list and picture associations.

Sequence	CiC		CyC		CuC	
	Word	**Image**	**Word**	**Image**	**Word**	**Image**
Triplet 1	bise [biz] 'kiss'		bus [bys] 'bus'		bouc [buk] 'goat'	
Triplet 2	lit [li] 'bed'		lu [ly] 'read'		loup [lu] 'wolf'	
Triplet 3	rie [ʁi] 'laugh'		rue [ʁy] 'street'		roux [ʁu] 'red-headed'	
Triplet 4	gris [gʁi] 'grey'		grue [gʁy] 'crane'		groupe [gʁup] 'group'	
Triplet 5	chic [ʃik] 'chic'		chute [ʃyt] 'fall'		choux [ʃu] 'cabbage'	

References

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48.

Beddor Patrice S. & Winifred Strange. 1982. Cross-language study of perception of the oral–nasal distinction. *The Journal of the Acoustical Society of America* 71. 1551–1561.

Best, Catherine T. & Michael D. Tyler. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In Murray. J. Munro & Ocke-Schwen Bohn (eds.), *Second Language Speech Learning: The Role of Language Experience in Speech Perception and Production*, 13–34. Amsterdam: John Benjamins.

Boomershine, Amanda, Kathleen Currie Hall, Elizabeth Hume & Keith Johnson. 2008. The impact of allophony versus contrast on speech perception. In Peter Avery, B. Elan Dresher and Keren Rice (ed.), *Contrast in Phonology: Theory, Perception, Acquisition*, 145–172. Berlin, New York: De Gruyter Mouton.

Boersma, Paul & David Weenink. 2022. Praat: Doing phonetics by computer [Computer program]. Version 6.2.14, http://www.praat.org/. (10 Oct 2020)

Brysbaert, Marc. 2013. LEXTALE_FR: A fast, free, and efficient test to measure language proficiency in French. *Psychologica Belgica* 53. 23–37.

Carignan, Christopher. 2014. An acoustic and articulatory examination of the 'oral' in 'nasal': The oral articulations of French nasal vowels are not arbitrary. *Journal of Phonetics* 46. 23–33.

Cohn, Abigail C. 1990. Phonetic and Phonological Rules of Nasalization. *UCLA Working Papers in Phonetics* 76. 1–224.

Cohn, Abigail C. 2007. Phonetics in phonology and phonology in phonetics. *Working Papers of the Cornell Phonetics Laboratory* 16. 1–31.

Cutler, Anne, Andrea Weber & Takashi Otake. 2006. Asymmetric mapping from phonetic to lexical representations in second language listening. *Journal of Phonetics* 34. 269–284.

Dow, Michael. 2020. A phonetic-phonological study of vowel height and nasal coarticulation in French. *The Journal of French Language Studies* 30. 239–274.

Elff, Martin. 2022. mclogit: Multinomial logit models, with or without random effects or overdispersion. R package version 0.9.6, <https://CRAN.R-project.org/package=mclogit>. (12 August 2023)

Flege, James E. 1995. Second-language speech learning: Theory, findings, and problems. In Winifred Strange (ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, 229–273. Timonium, MD: York Press.

Flege, James & Ocke-Schwen Bohn. 2021. The revised speech learning model (SLM-r). In Ratree Wayland (ed.), *Second Language Speech Learning: Theoretical and Empirical Progress*, 3–83. Cambridge: Cambridge University Press.

Herd, Wendy, Allard Jongman & Joan Sereno. 2013. Perceptual and production training of intervocalic /d, ɾ, r/ in American English learners of Spanish. *The Journal of the Acoustic Society of America* 133. 4247–4255.

Ingram, John C. L., See-Gyoon Park & Tom Mylne. 1997. Studies in cross-language speech perception. *Asian Pacific Journal of Speech, Language, and Hearing* 2. 1–23.

Kotzor, Sandra, Allison Wetterlin, Adam C. Roberts, Henning Reetz & Aditi Lahiri. 2022. Bengali nasal vowels: lexical representation and listener perception. *Phonetica* 79. 115–150.

Lahiri, Aditi & Allard Jongman. 1990. Intermediate level of analysis: features or segments? *Journal of Phonetics* 18. 435–443.

Lahiri, Aditi & William Marslen-Wilson. 1991. The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition* 38. 245–294.

Martinez, Ruth. M., Heather Goad & Michael Dow. 2021. L1 phonological effects on L2 (non-)naïve perception: A cross-linguistic investigation of the oral-nasal vowel contrast in Brazilian Portuguese. *Second Language Research* 37. 1–37.

Marquez Martinez, Miguel. A. 2016. *The acquisition of French nasal vowels: From first language allophony to second language phonological contrast*. Indiana University dissertation.

Neary, Terrance. 1989. Static, dynamic, and relational properties in vowel perception. *The Journal of the Acoustical Society of America* 85. 2088–2113.

Pallier, Christophe, Angels Colomé & Núria Sebastián-Gallés. 2001. The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science* 12. 445–449.

R Core Team. 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

Reetz, Henning. 2020. Formants_5_0_0 [Praat script]. Retrieved from https://github.com/Henning Reetz/Praat-scripts/. (3 April 2023)

Sampson, Rodney. 1999. *Nasal Vowel Evolution in Romance*. Oxford: Oxford University Press.

Solé, Maria-Josep. 1992. Phonetic and phonological processes: The case of nasalization. *Language and Speech* 35. 29–43.

Steriade, Donca. 1999. Lexical conservatism. In Linguistic Society of Korea (LSK), *Linguistics in the Morning Calm: Selected Papers from SICOL-'97 [Seoul International Conference on Oceanic Linguistics]*. 157–179. Seoul: Hanshin Publishing.

Styler, Will. 2017. On the acoustical features of vowel nasality in English and French. *The Journal of the Acoustical Society of America* 142. 2469–2482.

Styler, Will. 2018. NasalityAutomeasure [Praat script]. Retrieved from https://github.com/stylerw/styler_praat_scripts (3 April 2023)

Tyler, Michael. 2021. Perceived phonological overlap in second-language categories: The acquisition of English /r/ and /l/ by Japanese Native Listeners. *Languages* 6. 1–23.

Walker, Douglas C. 2001. *French Sound Structure*. Canada: University of Calgary Press.

Weber, Andrea & Anne Cutler. 2004. Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language* 50. 1–25.