Concordia Working Papers in Applied Linguistics

Proceedings of the International Symposium on the Acquisition of Second Language Speech
Concordia Working Papers in Applied Linguistics, 5, 2014 © 2014 COPAL

Perception of Allophonic Vowel Variation Among French L2 Learners

Jenna Nichols
University of Texas at Austin

Abstract

Native English speakers learning French and previously unexposed to Quebec French (QF) underwent a perception experiment targeting allophones of phonemes present in QF but not in their instructional variety of French. In a discrimination task, participants with higher levels of proficiency were found to have higher accuracy in identifying and interpreting this variation, and allophones with a large degree of overlap with English were found to be more accurately perceived. In an intelligibility task, items in QF presented much more difficulty than those spoken in the participants' instructional variety. The study has implications for the impact of the instructional variety of the L2 on learners' ability to interpret L2 variation.

This study examines L2 French learners' ability to perceive allophonic contrasts in French that are subject to regional variation. In particular, the targeted variable is the tense-lax allophone pairs produced in Quebec French (QF) but not in standard, hexagonal French (HF): [i] and [ɪ]; [u] and [ʊ]; and [y] and [ʏ] (Friesner, 2010). In general, the tense variants appear in stressed open syllables, while the closed counterparts are limited to stressed closed syllables (Dumas, 1987). With European French

generally being the standard variety taught in American classrooms (Auger & Valdman, 1999), L2 learners learning French in the United States would not likely have much previous exposure to this contrast. Perception tasks including a discrimination task and an intelligibility task were administered to 28 L2 French learners of various levels, with the goal to analyze the extent to which a student's experience or level of a certain variety of French has an impact on his or her ability to recognize and interpret regional variation.

Previous work has examined the way L2 French learners perceive and distinguish French phonemes, particularly those such as /y/ and the nasal vowels which do not exist as phonemes in English (Levy & Law, 2010; Delvaux, 2009), finding generally that higher proficiency in the language has positive results in terms of distinguishing (and producing) these foreign phonemes. Other work has examined vowel perception from a dialectological perspective. Baker and Smith (2010), for example, found that L2 French learners learning Quebec French were more accurate at distinguishing French high vowels than those learning European French due to the extra assibilation cue in Quebec French but not in European French, highlighting the significance of the variety of French that students are taught. This notion of exposure to more minority varieties of French also has consequences for native speakers. In one study, Swiss French speakers were shown to be more accurate at perceiving phonemic contrasts in vowel length than Parisian French speakers, who no longer maintain that contrast in their dialect (Grosjean, Carrard, Godio, & Grosjean, 2007). All taken together, these studies motivate the study of perception of regional varieties of French among L2 learners, whose degree of exposure to French, and in particular to a certain instructional variety, may impact perceptional ability.

METHODOLOGY

Stimuli

The experimental material used in the perception test was inspired by the matched-guise technique conceived by Lambert (1960), which set out to study language attitudes toward French and English speakers in Montreal. Lambert found that different judgments were produced for the same speakers (bilinguals) speaking each language. With this technique eliminating inter-speaker variation that could contain additional cues or variables to influence speaker judgments, it thus allows for results to be

attributed to the language itself and the values associated with it. Among many other studies, this technique has also been employed to examine reactions toward written and spoken Quebec French and European French by English-speaking Canadian students in Montreal (Remillard, 1973), finding more favorable reactions to European French even among these students.

In this study, the matched-guise approach was motivated by the desire to eliminate inter-speaker variation that is particularly salient in such short tasks as the identification or discrimination of one or two words at a time, where the task of normalizing speech by the listener would become exceedingly taxing and would likely create undesired consequences for the study. Furthermore, the matched-guise approach allows for more natural-sounding stimuli, as opposed to synthesized speech.

As such, the material for the test stimuli was procured from recordings of two native female Quebec French speakers, who had had sufficient experience to hexagonal French to duplicate the same recordings in that variety. The two speakers recorded a list of monosyllabic French words at the end of the carrier phrase "*Je dis le mot...*" ("I say the word…"). The carrier phrase served only to ensure intonational uniformity, and ultimately only the target words were included in the experiment. The list targeted mostly high vowels but included mid and low vowels as well, to be included later as distractors. The test words included a variety of consonantal environments and were split roughly half-and-half between open and closed syllables, including minimal pairs and near-minimal pairs (consisting of open-closed pairs to elicit the tense-lax QF contrast) where possible. The stimuli also underwent a verification process, by which the vowel formants in the words were measured and compared between dialects of the same speaker. In this way, it was assured that, for example, closed-syllable /i/ was realized [ɪ] in the speaker's QF but [i] in her HF. Significant differences between the dialects' closed syllables were recorded for all the high vowels, while the open syllable vowels were not significantly different between dialects. From there, the target words were extracted and scaled to equal intensity; finally, each word was adjusted to equal duration by appending silence for a total length of 1.5 seconds.

To form the test stimuli for the discrimination task, the words were concatenated into pairs for four types of test questions by varying the language variety, phoneme, and type of syllable (open/closed). This resulted in the following types of responses, where "vowel" indicates realized vowel pronunciation, i.e. allophone: (a) Q: Same Vowel, Same Word; (b) S: Same Vowel, Different Word; (c) P: Different Vowel, Same

Word; (d) L: Different Vowel, Different Word. Category Q consisted of a word from either dialect and the same word repeated in the same dialect. Category S consisted of pairs of words with the same vowel in the same dialect (i.e. same allophone). The words differed only by their consonant and were constant in syllabic structure (i.e. both open or both closed). Category P consisted of the same closed-syllable word pronounced once in HF and once in QF, whereby varying the allophone. Category L consisted of minimal phonemic pairs, where the consonants, dialect and syllabic structure were kept constant. In all cases, the two words could be spoken by the same speaker or by different speakers. These categories are summarized with examples in Table 1.

Table 1. Summary of Test Stimuli Categories

Category	**Vowel**	**Word**	**Dialect**	**Syllable**	**Consonants**	**Example**
Q	Same	Same	QF or HF, constant	Open or closed, constant	Constant	QF *bulle* [bʏl] (‘bubble’) QF *bulle* [bʏl]
S	Same	Diff.	QF or HF, constant	Open or closed, constant	Different	HF *mou* [mu] (‘soft’) HF *loup* [lu] (‘wolf’)
P	Diff.	Same	QF and HF, one each	Both closed	Constant	QF *cil* [sɪl] (‘eyelash’) HF *cil* [sil]
L	Diff.	Diff.	QF or HF, constant	Open or closed, constant	Constant	QF *rue* [ʁy] (‘street’) QF *roue* [ʁu] (‘wheel’)

For the discrimination task, thirty test pairs were created to correspond to each category, including non-experimental items with low and mid vowels. This resulted in 120 trials for the discrimination task, with the word pairs for all trials roughly balanced in terms of speaker order, vowel

phoneme, and dialect (where applicable). These items were verified auditorily for sound quality.

For the intelligibility task, the stimuli creation process was much less difficult, since only single words were to be chosen. From a list of all of the single-word recordings from both speakers in both varieties, a randomized list was generated. The first 50 were taken as a base list, but some words were exchanged in order to include more near-minimal pairs.

Procedure

The experiment was administered via headphones on a computer using the stimuli presentation software SuperLab. First, the participants went through a short orientation activity, which introduced them to the discrimination task instructions and keyboard input method. Answer choices were listed and explained along with audio examples of corresponding word pairs in English.

All participants completed the discrimination task first. For each of the 120 trials played, participants selected a response on the keyboard (Q, S, P, or L), with the choices visible on the screen (e.g. Same Word, Same Vowel) throughout all trials. Participants were instructed to go at their own pace, and the next trial was played immediately after selecting an answer choice. The order of the stimuli was random for each participant. Upon completion of the discrimination task, participants completed the intelligibility task, which was completed with pencil and paper. Participants were instructed to write the French word that they heard, and this task was also self-paced. The order of the fifty items was randomized once and played in this same order for all participants. Finally, at the end of both tasks, the participants completed a background questionnaire detailing their experiences with French, including formal study and study abroad experience. The total time for all four parts of the study (orientation, discrimination, intelligibility, and questionnaire) typically took 30-40 minutes.

Participants

Twenty-eight native English speakers (12 male, 16 female) completed the experiment, with a mean age of 33. According to their self-reports on the questionnaire, all of their French experience was limited to the United States or France in the case of study abroad; none had studied abroad in Quebec or reported exposure to that variety. Participants also self-

assessed their proficiency in French, and this assessment was corroborated by their formal study and reported experience with the language. For example, beginners typically reported 1-2 semesters of French study, while the advanced or near-native learners had numerous years of study and experience abroad. Out of all 28 participants, there were 6 beginners, 11 intermediate (4 early, 7 late), and 11 advanced or near-native.

RESULTS

The discrimination task (intelligibility task results are treated in the Discussion) was first scored for overall accuracy (correct response out of the four choices) for all trials including at least one high vowel (Table 2). A one-way ANOVA found a significant effect for proficiency with 5 levels, including beginners, early intermediate, late intermediate, advanced, and near-native ($F_{(4,23)} = 3.93$, $p = 0.014$). However, the categories were subsequently collapsed into beginners, intermediate (including early and late), and advanced (including near-native) to simplify further analysis. Under this model, proficiency was also found to be significant ($F_{(2,25)} = 7.19$, $p = 0.003$), with advanced learners outperforming intermediates and beginners, who performed at around chance.

Table 2. Accuracy by Proficiency Level

Proficiency	**Mean Score % (std. dev.)**
Beginner	52.1 (13.6)
Intermediate	66.2 (9.1)
Advanced	73.8 (11.9)

Results were then broken down by question type (see Table 1). Although the lax allophonic variants of QF (i.e. the most unfamiliar variants) were found in all question types, they were specifically targeted in type P (Different Vowel, Same Word), which always consisted of a closed syllable with the lax QF variant contrasting with the corresponding tense HF variant. It was therefore hypothesized that participants would fare the worst on this question type, being unfamiliar with this property of QF. A 4x3 (Question Type x Proficiency) repeated-measures ANOVA found significant effects for both Question Type ($F_{(3,75)} = 10.89$, $p < 0.001$) and Proficiency ($F_{(2,25)} = 7.97$, $p = 0.002$. Although there was no overall interaction found between proficiency and question type, a post-hoc Tukey test found that accuracy for beginners was significantly lower than

that of intermediate and advanced learners (p = 0.043 and p = 0.001 respectively). Overall, it was found as expected that type P was met with the lowest accuracy scores across all proficiency levels, at 52.9%. This was followed by type Q (Same Vowel, Same Word) at 64.9%, type S (Same Vowel, Different Word) at 72.5%, and type L (Different Vowel, Different Word) at 77.1%. This is represented visually in Figure 1. Although it is somewhat surprising that participants did not demonstrate higher accuracy the Same/Same question type (Q), potential explanations for this will be treated in the discussion.

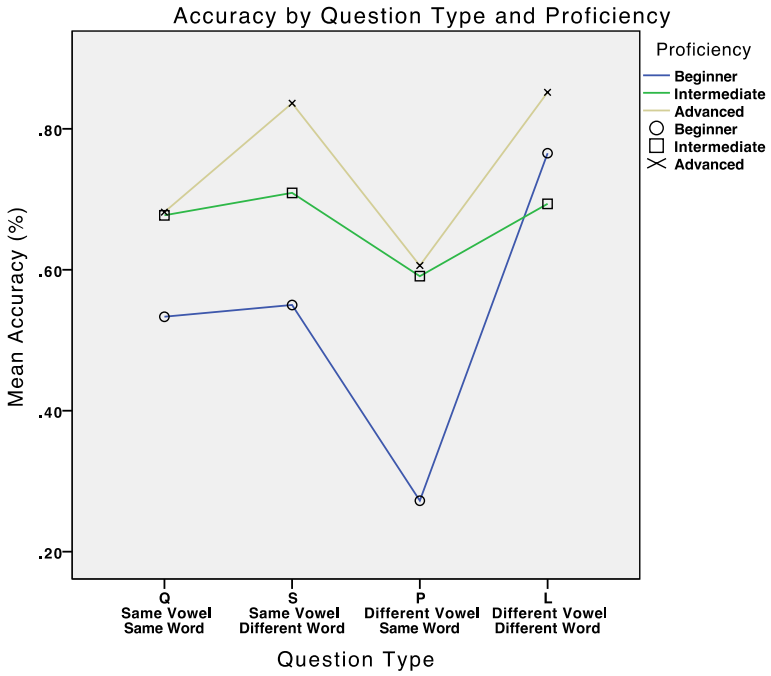

Figure 1. Accuracy by question type and proficiency

Finding that type P was indeed met with the most difficulty, further analysis sought to examine the types of errors made by the participants, i.e. which responses were chosen for these trials. Overall, the correct answer choice was selected close to half the time (49.3%), while the answer selected next most often was type Q (Same Vowel, Same Word) at 24.4%. Beginners were more likely than Intermediate or Advanced learners to choose this choice, as illustrated in Figure 2. Across all participants, types L (Different Vowel, Different Word) and S (Same Vowel, Different Word) were chosen 13.3% and 6.3% of the time, respectively. For type P, we also note that overall, participants fared better

at the word level (sum of P and Q responses totaling 79.0%) than at the vowel level (sum of P and L responses totaling 66.9%).

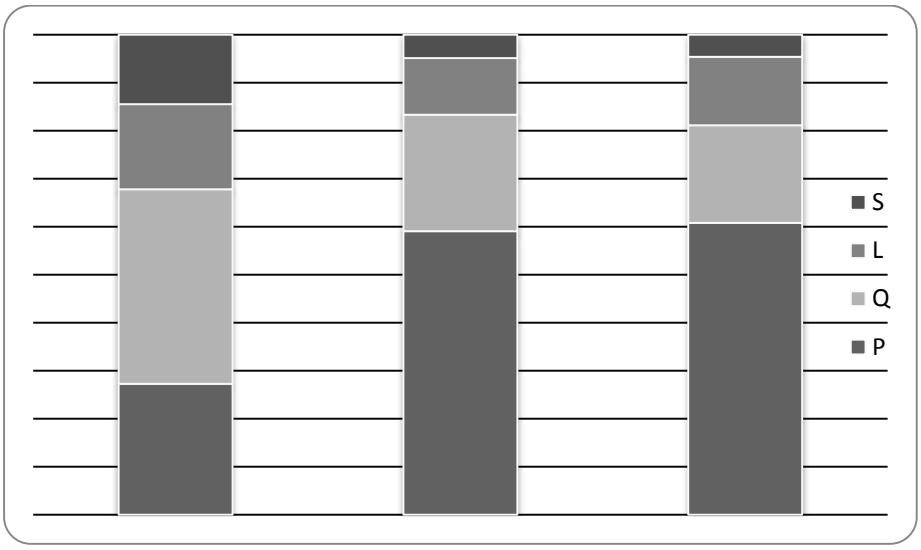

Figure 2. Answers selected by participants for the allophonic condition (P).

Next, a 3x3 (Phoneme x Proficiency) repeated-measures ANOVA was used to test whether participants performed with higher accuracy on any of the three phonemes /i/, /u/, or /y/ across all question types. It was hypothesized that trials involving allophones of /i/ would be found easier by the participants, as both /i/ and /ɪ/ are phonemes of English and therefore could be more easily distinguished. Indeed, significant effects were found for both Phoneme ($F_{(2,50)} = 12.75$, $p < 0.001$) and Proficiency ($F_{(2,25)} = 6.88$, $p = 0.004$), with trials involving /i/ being met with more accuracy than for /u/ or /y/, as illustrated in Figure 3. Advanced learners neatly outperformed intermediates, who in turn performed more accurately than beginners across all phonemes.

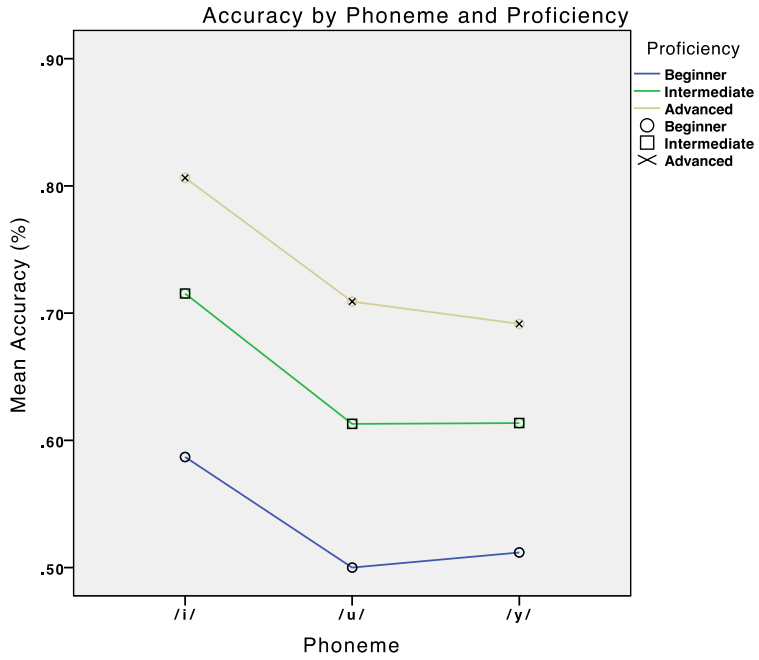

Figure 3. Accuracy by phoneme and proficiency

Finally, we might compare the errors for each phoneme within the allophonic condition; here, for example [i] would always contrast with [ɪ] and [u] with [ʊ]. We see that the response choice breakdown differs considerably between /i/ and /u/ (also representing /y/, which had very similar results), as shown in Table 3. While the proportion of S and L responses is similar between the two phonemes, Q and P are quite different. Regardless of phoneme, participants largely agreed that the two words in the trial were the same but had more difficulty differentiating [u]-[ʊ] than [i]-[ɪ].

Table 3. Response Choices by Phoneme in Allophonic Condition

Category (V/W)	**Response (%) /i/**	**Response (%) /u/**
Q (Same/Same)	20.0	27.2
S (Same/Diff)	7.1	6.6
P (Diff/Same)	57.6	48.2
L (Diff/Diff)	15.0	17.9

DISCUSSION AND CONCLUSION

The results of the discrimination task verify the difficulty that learners of French in the United States have when confronted with an unknown variety of French; however, experience with French (here operationalized as proficiency) allows learners to better perceive and interpret allophonic variation. As expected, it was found that the question type explicitly targeting allophonic variation (Different Vowel, Same Word) across the two varieties of French was met with the least accuracy. Surprising, however, was that participants fared relatively poorly on the Same Vowel, Same Word (Q) category, in which the same word is repeated by a same or different speaker (although even by the same speaker, the tokens are different). For this category, it was more often the vowel than the word that was labeled as different (27.1% of the time versus 13.9% of the time, respectively). Due to the nature of the task which asks participants to make judgments about very subtle vowel variation, it may be the case that participants, especially the beginners who are less familiar with French phonology, perceived the small natural intra- and interspeaker differences between tokens as distinct enough to merit a "different" vowel reponse, and confusion on the word level could be the result of mishearing the initial or final consonants in isolation.

Within the allophonic condition, it was also demonstrated the beginners, compared to intermediate and advanced learners, perceive this regional variation differently. With a higher proportion of Q (Same Vowel, Same Word) responses, beginners were less likely to notice allophonic variation than intermediate or advanced learners. It may be that they are more focused on understanding the words themselves to be able to be attune to such subtle differences.

In other instances, the phoneme status of the sounds in the L1 may have an impact on allophone perception. The allophonic pair [i]-[ɪ] was labeled as the same less often than the pairs [u]-[ʊ] or [y]-[ʏ]. Because both of these sounds are phonemes in English it is possible that allowed them to more easily perceived as different. While the same may be said of [u]-[ʊ], its patterning with [y]-[ʏ] may be due to the relatively lower frequency of minimal pairs in English as compared with [i]-[ɪ] and that English /u/ is considerably more fronted than /u/ in French (Clopper, Pisoni, & de Jong, 2005), leading [i]-[ɪ] to be a closer perceptual match with English. However, this distinction did not necessarily result in the two sounds being perceived as phonemes (i.e. a distinction at the word level), indicating that L2 learners have relatively good intuitions about L2

phonology.

The results of the intelligibility task corroborate the difficulty that learners of French in the United States have when confronted with this unknown QF variety. This task was evaluated qualitatively, for the nature of the write-in response resulted in a vast array of responses and errors, including large variations in orthography especially for the QF items. In this way, it was evident that this task was found difficult by many participants. Items containing a perhaps unexpected QF lax variant were especially prone to inventive spellings or changed vowel quality, such as the spelling *"pleusse"* for QF *plus* ([plʊs]) or *"cesse"* for QF *six* ([sɪs]). Comparison across varieties also proved informative, with participants writing different spellings for the same words pronounced in QF and HF, for example *"quitte"* for the HF trial and *"kit"* for the QF trial, demonstrating that the words were not recognized at the same. The near-minimal (allophonic) pair for *cil* also had distinctions of this sort. In HF, many participants wrote "*s'il*" (or *"cil"*, or *"sil"*), and this often translated into *"seule"*, *"sille"*, or *"celle"* in QF. Evidence of influence from English was also evident, largely for the QF items, such as *"kit"* for QF *quitte* or the very common spelling *"fight"* for the QF distractor word *fête*. Overall, it was evident that the QF words presented more confusion than the HF ones, particularly in closed syllables. However, even the more advanced learners were prone to different judgments for the near-minimal allophonic pairs, but they were less "inventive" in terms of orthography. That is, advanced learners were more likely to substitute a known word such as *"seule"* for *cil*, while participants at the lower levels were more prone to writing something like *"sille"*, and this is reflective of both the phonological and lexical knowledge for which the advanced learners are at an advantage.

To conclude, these discrimination and intelligibility tasks both reflect the difficulty faced by L2 French learners when confronted with unknown variety of French, in this case Quebec French, which has important phonological differences from the "standard" hexagonal French to which they are virtually exclusively exposed. With more phonological and lexical knowledge, intermediate and advanced learners are more equipped to interpret and be more attune to nuanced distinctions in allophonic variation, and overlap with the L1 also allows more enhanced perception of this particular allophonic phenomenon. This study is an introductory look at the implications of the instructional variety of French in L2 perception, and future work should compare results from these learners with students in Canada who ostensibly have more access to

Quebec French. Further work can also examine the perception of other diatopic phonological variables and explore the impact of instructional L2 variety on more global perception tasks such as dialect perception and classification.

REFERENCES

Auger, J., & Valdman, A. (1999). Letting French students hear the diverse voices of francophony. *The Modern Language Journal, 83*(3), 403-412.

Baker, W., & Smith, L. C. (2010). The Impact of L2 Dialect on Learning French Vowels: Native English Speakers Learning Québécois and European French. *The Canadian Modern Language Review, 66*(5), 711-738.

Clopper, C. G., Pisoni, D. B., & de Jong, K. (2005). Acoustic characteristics of the vowel systems of six regional varieties of american english. *The Journal of the Acoustical Society of America, 118*(3), 1661-1676.

Delvaux, V. (2009). Perception of the vowel nasalization contrast in French. *The Journal of French Language Studies, 19*(1), 25-59.

Dumas, D. (1987). *Nos façons de parler.* Sillery, Qc.: Presses de l'Université du Québec.

Friesner, M. (2010). Une prononciation «tsipéquement» québécoise?: La diffusion de deux aspects stéréotypés du français canadien. *The Canadian Journal of Linguistics, 55*(1), 27-53.

Grosjean, F., Carrard, S., Godio, C., Grosjean, L., & Dommergues, J-Y. (2007). Long and Short Vowels in Swiss French: Their Production and Perception. *Journal of French Language Studies, 17*(1), 1-19.

Lambert, W. E., Hodgson, R. C., Gardner, R. C., & Fillenbaum, S. (1960). Evaluational reactions to spoken language. *Journal of Abnormal and Social Psychology, 60*(1), 44-51.

Levy, E. & Law, F. (2010). Production of French Vowels by American-English Learners of French: Language Experience, Consonantal Context, and the Perception-Production Relationship. *The Journal of the Acoustical Society of America, 128*(3), 1290-1305.

Remillard, L., Tucker, G. R., & Bruck, M. (1973). The role of phonology and lexicon in eliciting reactions to dialect variations. *Anthropological Linguistics, 15*(9), 383-397.