

# **Patrons de coarticulation des voyelles françaises quantiques /i, a, u/ prononcées par des apprenants tchécoslovaques. Illustration du logiciel VisuVo.**

Nikola Maurová Paillereau

Laboratoire de Phonétique et Phonologie (UMR 7018), CNRS/ Paris 3, 19 rue des  
Bernardins, 75005 Paris, France  
nikola.paillereau@mac.com

## **RESUME**

---

Cette étude acoustique quantifie les écarts de prononciation des voyelles françaises quantiques /i, a, u/ produites par dix apprenants tchèques en isolation et dans des contextes symétriques labial pVp, dental tVt, palato-vélaire kVk et uvulaire RVR. Les productions de dix Françaises non-méridionales servent de référence. Les résultats, visualisés grâce aux schémas générés dans le logiciel VisuVo, montrent que 1. Les structures formantiques des voyelles produites par les apprenants diffèrent de la référence, notamment le F1 de [a] et le F2 de [i, a, u], et que 2. La réduction de l'aire des voyelles en contexte par rapport à celle des voyelles isolées est supérieure chez les apprenants comparées aux natives. L'ensemble des résultats indique que les voyelles quantiques ne sont pas réalisées par les deux groupes de locutrices avec des structures formantiques identiques, ce qui est en partie dû à des patrons de coarticulation différents.

## **ABSTRACT**

---

### **Coarticulatory patterns of French quantic vowels /i, a, u/ as pronounced by Czech learners. Illustration of the VisuVo software.**

This acoustic study quantifies pronunciation errors of French quantic vowels /i, a, u/ produced by ten Czech learners in isolation and in symmetrical labial pVp, dental tVt, palato-velar kVk and uvular RVR contexts. The productions of ten non-southern French serve as reference. The results, displayed through diagrams generated in the software VisuVo show that 1. The formant structure of vowels produced by learners differs from the reference, especially the F1 of [a] and F2 of [i, a, u] and, 2. The reduction of the area of vowels in context relative to that of isolated vowels is higher in learners than in native speakers. The overall results indicate that quantic vowels are not realized by the two groups of speakers with identical formant, which is partly due to different patterns of coarticulation.

---

**MOTS-CLES :** tchèque, français langue étrangère (FLE), voyelles quantiques, formants, coarticulation, apprentissage.

**KEYWORDS:** Czech, French as a Foreign Language (FFL), quantic vowels, formants, coarticulation, learning.

---

# 1 Objets et évaluation de l'apprentissage phonétique

L'apprentissage phonétique d'une langue étrangère (L2) est un processus complexe. Il englobe les phénomènes suprasegmentaux, c'est-à-dire le rythme, l'accentuation lexicale et l'intonation, de même que les segments, c'est-à-dire les voyelles et les consonnes. Au plan segmental, l'apprenant apprend à maîtriser les phonèmes de la L2, leurs combinaisons possibles au sein de la syllabe définies par les règles phonotactiques de la L2 ainsi que les variations allophoniques (Best and Tyler 2007). Ces dernières dépendent entre autre de la position prosodique et du contexte phonétique. Les variations allophoniques suivant la position par rapport à la syllabe accentuée sont relativement petites dans des langues où les sons sont produits avec une forte tension musculaire, comme en français (Delattre 1969) ou bien où la durée, étant contrastive, rencontre de faibles variations dues à l'accentuation, comme en tchèque (Flemming 2005). En revanche, les variations dues au contexte phonétique caractérisent toute parole naturelle mais leur ampleur varie d'une langue à l'autre (Manuel et Krakow 1984).

Si l'apprentissage phonétique d'une L2 est jugé difficile, l'évaluation exacte du niveau phonétique en L2 n'en est pas moins (Dickerson 1975). De nos jours, l'évaluation est essentiellement subjective, car les enseignants se fient le plus souvent à leur capacité à percevoir le progrès des apprenants. Même si cette étape est importante car l'effet que la parole de l'apprenant aura sur l'auditeur est primordial dans la communication interhumaine, l'évaluation peut se faire également de manière objective à l'aide des analyses de paramètres acoustiques et articulatoires (Ashby 1990). L'avantage de l'analyse instrumentale réside en la possibilité de tracer toute évolution du système phonétique de l'apprenant, ainsi que de définir et visualiser les zones de difficultés qu'il faut travailler tout particulièrement.

## 1.1 Patrons de coarticulation

Dans la chaîne parlée, les sons ne sont prononcés de manière isolée que très rarement (en français, les sons isolés apparaissent par exemple dans l'énonciation des interjections *Euh...* ou *Oh !*) (Wioland 2005). La plupart du temps, les voyelles et consonnes se trouvent en contexte phonétique immédiat qui influence leur nature. En effet, du fait de l'inertie variable des organes articulatoire, les gestes articulatoires lors de la production d'une suite de sons se chevauchent et le résultat peut être visible dans le signal acoustique sous forme de transitions formantiques (Fant 1973). Ce phénomène s'appelle la *coarticulation* car les sons adjacents sont en effet articulés ensemble.

La coarticulation est en partie déterminée par les lois acoustiques universelles. Les mouvements (transitions) formantiques des voyelles coarticulées sont quantifiés par rapport à la cible, qui est le plus souvent définie par des voyelles produites en isolation (Miller 1981) ou en contexte glottal (Stevens and House 1963). Ainsi, dans une suite Consonne-Voyelle (CV), la direction des transitions formantiques dépend du lieu d'articulation de la voyelle et de la consonne. Lorsqu'il s'agit d'une voyelle antérieure comme par exemple le /e/ français, où le deuxième formant (F2) relativement élevé est essentiellement affilié à la courte cavité antérieure, ce dernier va baisser au contact de la consonne labiale /p/, allongeant la cavité antérieure, rester stable au contact de la consonne dentale /d/, ayant le locus dental vers 1800 Hz (Delattre 1969) ou en revanche augmenter au contact de la consonne palato-vélaire /k/, ayant un locus vers 3000 Hz pour les voyelles antérieures (Kewley-Port 1982). Si la coarticulation s'expliquait uniquement par les lois acoustiques, les patrons de coarticulations seraient identiques dans toutes les langues et ne

poseraient aucun problème pour un apprentissage phonétique d'une L2. Or ceci n'est pas le cas. En effet, les langues varient quant à l'ampleur de la coarticulation. Manuel et Krakow (1984) expliquent ces différences de coarticulation par les différences de la taille des systèmes phonologiques. Les auteurs proposent que les langues ayant des inventaires phonologiques réduits rencontrent de plus amples variations allophoniques que les langues avec des inventaires phonologiques plus riches. Même si cette hypothèse n'est pas toujours confirmée (Maddieson et Wright 1996), les études de coarticulation confirment que les langues du monde varient de part de l'ampleur de la coarticulation (pour revue de littérature, voir Steinlen 2005). Ces résultats, appliqués au niveau de l'apprentissage phonétique d'une L2, expliquent pourquoi les apprenants doivent entre autre acquérir les patrons de coarticulation spécifiques à la L2.

En 1995, le Speech Learning Model – SLM, qui est un modèle d'acquisition de la parole, a été clairement décrit par Flege. Le SLM a servi de cadre pour de nombreuses études acoustiques sur la production des segments d'une L2. Cependant, même si Flege (1995) attire l'attention sur la nécessité pour un apprenant de maîtriser toutes les variations allophoniques afin de créer une représentation cognitive d'un phonème, peu de chercheurs s'intéressent aux effets de coarticulation en L2. A notre connaissance, aucune étude concernant l'effet systématique des consonnes sur les voyelles du FLE n'a été réalisée jusqu'à présent.

## 1.2 Les systèmes vocaliques du français et du tchèque

Le français, langue romane, est caractérisé par un système vocalique riche. Selon le pays et la région où le français est parlé comme langue maternelle (L1), le système vocalique est composé de trois à quatre voyelles nasales ainsi que de sept à onze voyelles orales (Vaissière 2006). Ainsi, le système maximal comporte les voyelles /i, e, ε, a, y, ø, œ, u, o, ɔ, α, œ̃, ê, ë, ã/ alors que le système minimal, caractéristique du français méridional, élimine les oppositions entre les voyelles mi-fermées/ mi-ouvertes e/ε, o/ɔ, ø/œ, de même que le /ɑ/ postérieur, pour opposer /an/ « âne » de /an/ « Anne » ou le /œ̃/ pour opposer /bRœ̃/ « brun » de /bRễ/ « brin » (Léon et Léon 2007). Le français standard - important à définir pour les besoins de la phonétique didactique - est parfois décrit comme le français « non méridional » (Pustka 2011). Le système vocalique du français standard contient dix voyelles orales et trois voyelles nasales qui doivent donc être présentées aux apprenants de français langue étrangère (FLE). Cependant, il existe une hiérarchie entre les voyelles du français qu'il faut prendre en compte dans l'apprentissage phonétique (Wioland 2005). Du fait de la distribution semi-complémentaire des voyelles moyennes et du nombre restreint de paires minimales basées sur leur opposition, les contrastes e/ε, o/ɔ et ø/œ ont un poids fonctionnel plus petit que les contrastes avec les voyelles fermées /i, y, u/ et la voyelle ouverte /a/.

Le tchèque, langue slave, est caractérisé par un système vocalique plus restreint. Traditionnellement, l'on postule cinq qualités vocaliques avec un contraste de durée, opposant les voyelles brèves /a, e, i, o, u/ et longues /a:, e:, i:, o:, u:/ (Kučera et George 1968). Les études instrumentales plus récentes pointent sur l'ouverture des voyelles fermées brèves (Skarnitzl et Volín 2012). Les auteurs tchèques contemporains décrivent ainsi le système phonologique des locuteurs du tchèque de Bohême (qui correspond au tchèque standard) avec le contraste entre les voyelles longues tendues et les voyelles brèves relâchée i:/i (Šimáčková et al. 2012) et parfois également le contraste entre u:/u (Duběda 2005). De la même façon, suite aux études instrumentales, les symboles pour indiquer une qualité mi-ouverte sont de plus en plus utilisés pour la transcription des voyelles moyennes antérieures /ε, ε:/ et parfois pour celle des voyelles moyennes postérieures du

tchèque /ɔ, ɔ:/ (Duběda 2005). Enfin, le système tchèque comporte également trois voyelles diphtongues /au, eu, ou/.

Nous avons récemment réalisé (Paillereau, à paraîtreB) une comparaison des formants vocaliques des voyelles orales du français et des voyelles monophthongues du tchèque produites isolément et en contextes symétriques labial, dental et palato-vélaire. Les résultats montrent que les deux langues varient par l'ampleur de la coarticulation, comme attendu. Contre attente cependant, les voyelles postérieures du français sont plus fortement coarticulées, notamment en contexte dental, que les voyelles tchèques, moins nombreuses. Ces résultats peuvent être appliqués en didactique de la prononciation du FLE pour les apprenants tchécophones où des difficultés au niveau de la coarticulation des voyelles peuvent être anticipées du fait du transfert de la L1 vers la L2.

### **1.3 VisuVo, logiciel de visualisation des formants vocaliques**

Dans notre thèse de doctorat (Paillereau 2015), nous avons élaboré le logiciel VisuVo pour diagnostiquer les difficultés que les apprenants rencontrent dans la production des voyelles d'une L2. Le logiciel est composé de deux parties : la partie Serveur qui permet d'analyser en moins de deux secondes les données brutes, en calculant la moyenne et l'écart type des formants. Ensuite, la partie Client, écrite en Javascript, qui génère instantanément des schémas de visualisation. Les graphiques présentés dans cet article permettent de rendre compte des patrons de coarticulation en traçant l'évolution des formants vocaliques suivant la position dans le mot (initiale, médiane et finale) et le contexte consonantique symétrique (consonnes de différents lieux d'articulation : labiale /p/, dentale /t/, palato-vélaire /k/ et uvulaire /R/).

L'avantage du logiciel VisuVo par rapport aux logiciels comme Praat (Boersma et Weenink 2015) réside en la rapidité de la visualisation des résultats et en son interactivité. Il est possible de comparer différentes variables, comme les voyelles françaises et tchèques telles qu'elles sont produites par des natifs, ou bien des voyelles françaises produites par des natifs et par des apprenants, ou bien observer les différences individuelles en comparant les voyelles produites par les différents apprenants. Selon les objectifs, il revient ainsi à l'utilisateur d'organiser ses données dans un fichier Excel, qui sert d'entrée au logiciel VisuVo.

## **2 Objectifs**

Le but de l'article est d'examiner les effets systématiques des consonnes de différents lieux d'articulation – labial /p/, dental /t/, palato-vélaire /k/ et uvulaire /R/ - sur les voyelles quantiques du français /i, a, u/ produites par des apprenants tchécophones et de comparer les données à la référence. Les voyelles quantiques /i, a, u/ sont de l'intérêt particulier car elles sont définies par Stevens (1989) comme acoustiquement stables. Cependant, de nombreuses études ont montré que les structures formantiques des voyelles /i, a, u/ varient d'une langue à l'autre (Steinlen 2005, Gendrot et Adda-Decker 2007).

Les apprenants tchécophones du niveau avancé produiront-ils les voyelles françaises /i, a, u/ avec des structures formantiques différents de ceux de la « référence » ? Les écarts en production dépendront-ils des contextes consonantiques dans lesquelles les voyelles apparaissent ? La réalisation des patrons de coarticulation sera élucidée à l'aide des graphiques générés dans VisuVo.

## 3 Méthodologie

### 3.1 Corpus

Le corpus est celui de Landron et al. (2010). Les voyelles analysées dans cette étude sont le /a/, /i/ et /u/ du français. Elles apparaissent :

- En isolation, séparées par des petites pauses, dans des phrases cadre du type : « Bébé, il a dit <é> comme dans bébé ».
- En contexte consonantique symétrique, dans des logatomes trisyllabiques CVCVCVC où C correspond aux consonnes de quatre lieux d'articulation différents : labial /p/, dental /t/, palato-vélaire /k/ et uvulaire /R/. Les logatomes sont insérés dans des phrases cadre telles que « Le mot *kaukaukauke* peut bien coller. »

Toutes les phrases ont été présentées aux locuteurs quatre fois dans un ordre prédéfini de manière à ce qu'une même phrase ne soit jamais répétée à la suite. Avant les enregistrements, les locuteurs ont d'abord suivi une phase d'entraînement avec des consignes écrites et des exemples sonores préenregistrés.

### 3.2 Locuteurs

Les phrases ont été lues par dix Françaises natives dont les productions ont servi à définir la « référence » acoustique. Des femmes plutôt que des hommes ont été choisies pour des raisons pratiques : la plupart des étudiants du français à l'université, comparées à la « référence », sont des femmes. Les locutrices natives sont âgées entre 21 ans et 48 ans (M âge = 28,5 ; E.T. = 7,5). Aucune locutrice n'est née ni n'a grandi dans le sud de la France.

Le groupe test est composé de dix locutrices tchèques, âgées entre 25 ans et 28 ans (M âge = 25,7 ; E.T. = 1,3). Toutes les locutrices sont nées et ont grandi dans la région de Bohême et elles ont commencé à apprendre le français après l'âge de onze ans. Au moment des enregistrements, les locutrices étaient étudiantes en Master de FLE à l'université de Bohême et possédaient un bon niveau global du français (B2/C1).

### 3.3 Conditions d'enregistrement et traitement des données

Les enregistrements ont été réalisés dans une pièce calme (chambre sourde pour les locutrices natives et une pièce silencieuse pour les apprenants tchéophones) avec une carte son Edirol UA 25, reliée à un ordinateur Apple et un microphone serre-tête AKGC 520 L. L'utilisation du microphone serre-tête permet d'assurer une distance constante entre les lèvres et le microphone tout le long de l'enregistrement. La fréquence d'enregistrement était fixée à 44 100 Hz avec une résolution de 16 bits. Nous avons utilisé Audacity (<http://audacity.sourceforge.net>) pour enregistrer les phrases lues et Praat pour l'analyse acoustique.

Seulement les parties du signal acoustique comportant les voyelles isolées et les voyelles insérées dans les logatomes trisyllabiques ont été soumises à l'analyse. Les fichiers son ont été segmentés et étiquetés automatiquement et vérifiés manuellement. Le début et la fin vocalique coïncident avec l'apparition des formants et notamment du F2. Les valeurs des quatre premiers formants ont été relevées de manière semi-automatique avec la formule To Formant (burg) basée sur une analyse LPC<sup>1</sup>, à environ un tiers, à la moitié et à deux tiers de la durée vocalique. La fenêtre d'analyse spectrale étant fixée à 25 ms (spectre à bandes larges) permet une bonne analyse temporelle. En cas de mauvaise détection des formants - souvent due à la voix craquée<sup>2</sup> ou au *voice decay time*<sup>3</sup> - retournant des valeurs aberrantes, les données ont été vérifiées et corrigées manuellement. Toutes les données brutes ont été stockées dans un fichier Excel qui a servi d'entrée au logiciel VisuVo pour la visualisation des données.

La voyelle postérieure /u/, caractérisée par une proéminence perceptive des basses fréquences (Delattre 1951), est définie par la valeur des deux premiers formants. La voyelle antérieure /a/ est définie par la valeur des trois premiers formants (comme les autres voyelles antérieures du français dont F1 à F3 sont visualisés dans VisuVo afin de rendre compte du contraste entre les voyelles étirées et arrondies, même si ce contraste n'est pas pertinent pour le /a/). Enfin la voyelle /i/ est définie par la valeur des quatre premiers formants. Dans la production du /i/ français, l'énergie est concentrée au niveau des troisième et quatrième formants, qui sont rapprochés et par conséquent amplifiés (Vaissière 2009). Cette caractéristique acoustique du /i/ n'est cependant pas partagée par toutes les langues, comme démontré par Gendrot et Adda-Decker (2007) qui ont comparé des inventaires vocaliques de huit langues différentes, y compris le français.

## 4 Résultats

### 4.1 Comparaison de la structure formantique de [i, a, u]

Afin de comparer la structure formantique selon le contexte et le groupe de locutrices, des analyses de variance ANOVA à trois facteurs ont été effectuées sur les mesures de F1 et F2 de toutes les voyelles, sur le F3 de /i/ et de /a/ de même que sur le F4 de /i/. Le facteur Contexte possède cinq niveaux : isolé (iso), labial (/p/), dental (/t/), palato-vélaire (/k/) et uvulaire (/R/) et le facteur Locuteurs possède deux niveaux : locutrices françaises natives (FR) et apprenantes tchécoslovaques (TCH). Le facteur Voyelle possède trois niveaux dans l'analyse des formants F1 et F2 : /i/, /a/ et /u/ ; deux niveaux pour l'analyse du F3 : /i/ et /a/ ; et enfin le F4 n'est analysé que dans le cas de /i/.

Les Tableaux 1 et 2 indiquent les valeurs moyennes (en Hz) des formants et les écart types (en italique) pour chaque voyelle et chaque contexte séparément. Le Tableau 1 indique les données des dix locutrices françaises natives et le Tableau 2 les données des dix apprenants tchécoslovaques. Les valeurs ayant servi pour le calcul de la moyenne sont les trois valeurs prises pour les voyelles isolées et la valeur prise à la moitié de la durée vocalique pour les voyelles coarticulées.

<sup>1</sup> Linear Predictive Coding est une estimation du signal basée sur des coefficients de prédiction qui caractérisent des formants.

<sup>2</sup> Caractérisée par une irrégularité de la période et de l'amplitude des vibrations des plis vocaux.

<sup>3</sup> Période de temps où les plis vocaux continuent à vibrer mais le son vocalique n'est plus audible.

Le résultat de l'analyse de variance montre la significativité de l'effet global ainsi que de l'effet de chacun des trois facteurs sur les formant F1 à F3 ( $p < 0,05$ ). En revanche, l'effet du facteur Locuteur sur la valeur de F4 n'est pas significatif ( $p = 0,47$ ). Le test à posteriori de Fisher indique à quels groupes est dû l'effet de l'interaction des trois facteurs au niveau de chaque formant. Pour le F1, la significativité est due à la différence ( $p < 0,05$ ) dans la production de /a/ (dans tous les contextes sauf palato-vélaire), de /i/ (en contexte labial et uvulaire) et de /u/ (en contexte uvulaire).

FR	Iso			pVp			tVt			kVk			RVR		
	i	a	u	i	a	u	i	a	u	i	a	u	i	a	u
F1	272 26	805 132	<b>287</b> 30	<b>275</b> 28	843 93	<b>299</b> 31	<b>296</b> 29	817 113	<b>298</b> 27	<b>277</b> 23	774 113	<b>295</b> 26	431 46	852 104	333 47
F2	2524 272	1300 178	770 72	2360 266	1583 110	845 72	2417 266	1770 85	1176 174	2517 215	1873 117	<b>825</b> 81	<b>2499</b> 228	1412 128	<b>726</b> 81
F3	3787 229	<b>2834</b> 168		<b>3513</b> 256	<b>2797</b> 162		<b>3455</b> 242	<b>2856</b> 176		3686 167	<b>2642</b> 185		<b>3318</b> 209	2785 227	
F4	<b>4423</b> 196			<b>4367</b> 208			<b>4334</b> 195			<b>4359</b> 161			<b>4238</b> 311		

TABLEAU 1 : Valeurs formantiques moyennes (en Hz) des voyelles françaises /i, a, u/ produites en isolation (iso) et en syllabe finale pVp, tVt, kVk et RVR par 10 Françaises natives (valeurs relevées à la moitié de la durée vocalique, 4 répétitions). L'écart type (en Hz) est en italique. Les formants identiques aux voyelles des apprenants sont en gras.

TCH	Iso			pVp			tVt			kVk			RVR		
	i	a	u	i	a	u	i	a	u	i	a	u	i	a	u
F1	253 29	839 87	<b>286</b> 33	<b>264</b> 30	806 67	<b>303</b> 37	<b>270</b> 22	764 78	<b>300</b> 43	<b>268</b> 24	775 70	<b>306</b> 39	379 65	799 90	364 44
F2	2663 249	1479 222	708 93	2616 226	1476 166	727 73	2621 196	1635 181	1074 201	2697 220	1665 180	<b>777</b> 91	<b>2516</b> 130	1564 163	<b>787</b> 142
F3	3711 237	<b>2800</b> 119		<b>3553</b> 153	<b>2766</b> 163		<b>3441</b> 166	<b>2913</b> 160		3597 213	<b>2594</b> 207		<b>3254</b> 253	2810 153	
F4	<b>4359</b> 209			<b>4364</b> 269			<b>4383</b> 261			<b>4379</b> 232			<b>4264</b> 231		

TABLEAU 2 : Valeurs formantiques moyennes des voyelles françaises /i, a, u/ produites en isolation (iso) et en syllabe finale pVp, tVt, kVk et RVR par 10 apprenants tchéophones (valeurs relevées à la moitié de la durée vocalique, 4 répétitions). L'écart type (en Hz) est en italique. Les formants identiques aux voyelles des Françaises natives sont en gras.

Au niveau du F2, la significativité est due aux voyelles produites dans la plupart des contextes. En effet, seules les voyelles /u/ (en contexte palato-vélaire et uvulaire) et /i/ (en contexte uvulaire) sont produites avec la même valeur moyenne du F2 par les deux groupes de locutrices. La différence significative de F3 des voyelles /i/ et /a/ selon le contexte et les locutrices est due à la voyelle /i/ produite en isolation et en contexte palato-vélaire, ainsi qu'à la voyelle /a/ produite en contexte uvulaire. Enfin, la voyelle /i/ a été produite avec la même valeur moyenne de F4 par les deux groupes de locutrices dans tous les contextes étudiés. Ces résultats sont en partie reportés dans les Tableaux 1 et 2 où les formants moyens identiques ( $p > 0,05$ ) pour une même voyelle produite par les deux groupes de locutrices sont en gras. Les Figures 3 à 5 montrent les patrons de coarticulation des voyelles françaises /i, a, u/ produites en isolation (traits horizontaux traversant l'image) et dans des logatomes CVCVCVC où C correspond respectivement à la consonne labiale, dentale, palato-vélaire et uvulaire. Les moyennes affichées des voyelles coarticulées sont celles des valeurs prises à un tiers, à la moitié et à deux tiers de la durée vocalique. Dans notre étude, nous nous intéressons aux moyennes des valeurs prises à la moitié de la durée des voyelles coarticulées finales. Pour les voyelles isolées, une valeur moyenne globale est calculée à partir des trois valeurs. Les voyelles des

Françaises natives sont marquées avec un rond et celles des apprenants tchèques avec une croix. L'écart-type affiché est de un.

#### 4.1.1 Structure formantique du [i]

La Figure 3 illustre l'évolution des formants moyens F1 à F4 du [i] selon le contexte et les locutrices. En isolation, les apprenants produisent le [i] par rapport à la référence avec un F1 plus bas (253 Hz contre 272 Hz), un F2 plus élevé (2663 Hz contre 2524 Hz) et un F3 plus bas (3711 Hz contre 3787 Hz) (pour l'ensemble des résultats, se référer aux Tableaux 1 et 2). Lorsque [i] apparaît en contexte labial, dental et palato-vélaire, les Tchécophones le produisent avec un F2 plus élevé (de 256 Hz, 204 Hz et 180 Hz respectivement) que la référence. En contexte uvulaire, le [i] des apprenants est réalisé avec un F1 plus bas que la référence (379 Hz contre 431 Hz). Cette différence est probablement due à une différence dans l'articulation de la consonne /R/ que les Tchèques réalisent parfois comme apicale (Nováková, à paraître).

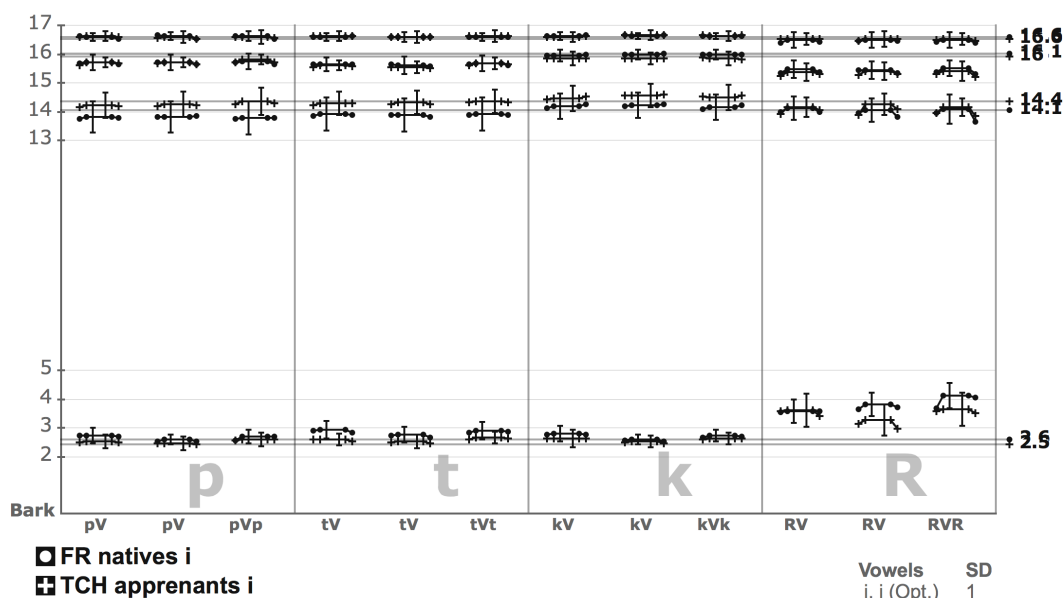


FIGURE 3: Formants moyens F1, F2, F3 et F4 (en Bark) du [i] français produit par 10 Tchèques (croix) et par 10 Françaises (rond) en isolation (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʀ]. L'écart-type est de un.

#### 4.1.2 Structure formantique du [a]

La Figure 4 illustre l'évolution des formants moyens F1 à F3 du [a] selon le contexte et les locutrices. En isolation, les apprenants produisent le [a] par rapport à la référence avec un F1 et un F2 plus élevés (respectivement 839 Hz contre 805 Hz et 1479 Hz contre 1300 Hz). En contexte labial et dental, la tendance s'inverse car les apprenants produisent le [a] avec un F1 et F2 plus bas que la référence. Les valeurs sont respectivement de 806 Hz et 1476 Hz contre 843 Hz et 1583 Hz en contexte labial et de 764 Hz et 1635 Hz contre 817 Hz et 1770 Hz en contexte dental. En contexte palato-vélaire, la différence concerne le deuxième formant uniquement : les apprenants produisent un [a] avec un F2 de 208 Hz plus bas que la référence. Ce résultat soutient une forte



antériorisation acoustique du [a] français produit en contexte consonantique (Paillereau, à paraîtreA) qui est ici une source de problèmes pour les apprenants. Enfin, en contexte uvulaire, le [a] des apprenants est réalisé avec un F1 plus bas (799 Hz contre 852 Hz) et un F2 plus élevé (1564 Hz contre 1412 Hz) que la référence, ce qui peut de nouveau s'expliquer par l'articulation apicale de certaines occurrences du /R/ produit par les Tchèques.

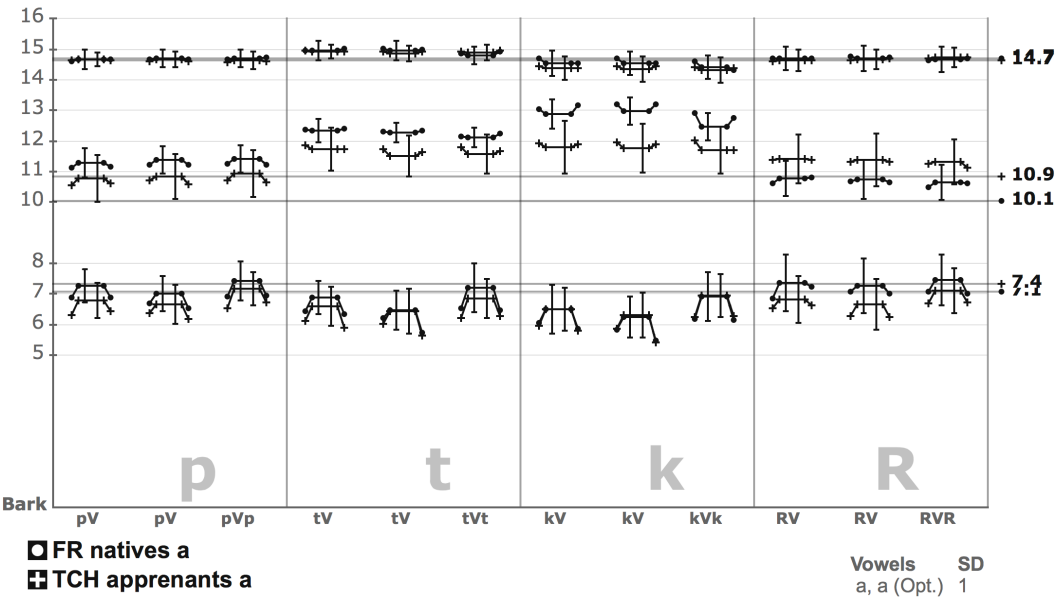


FIGURE 4: Formants moyens F1, F2 et F3 (en Bark) du [a] français produit par 10 Tchèques (croix) et par 10 Françaises (rond) en isolation (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʀ]. L'écart type est de un.

#### 4.1.3 Structure formantique du [u]

L'évolution des formants F1 et F2 du [u] français selon le contexte et les locutrices est illustrée à la Figure 5. En isolation, les apprenants produisent le [u] français avec un F2 plus bas que la référence (708 Hz contre 770 Hz). Ce résultat est surprenant compte tenu du fort caractère focal caractérisant le [u] français, qui est un parfait exemple des voyelles cardiales telles que définies par Daniel Jones (Vaissière 2009). Cette différence va dans le même sens en contexte labial et dental dans lequel les Tchèques produisent le [u] avec un F2 plus bas que les Françaises (il est de 727 Hz contre 845 Hz pour le contexte labial et de 1074 Hz contre 1176 Hz pour le contexte dental). En contexte palato-vélaire, le [u] des apprenants est réalisé avec exactement la même structure formantique que le [u] des Françaises natives ( $p > 0,05$  pour F1 et F2). Enfin, en contexte uvulaire, la différence concerne le F1 qui est de 31 Hz plus élevé chez les apprenants que chez les natives.

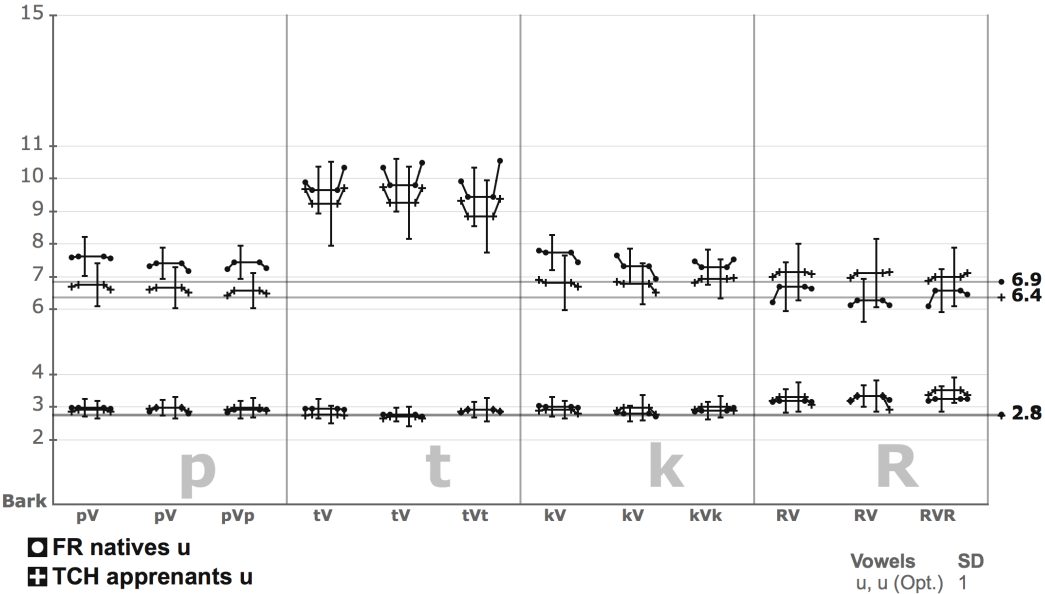


FIGURE 5: Formants moyens F1 et F2 (en Bark) du [u] français produit par 10 Tchèques (croix) et par 10 Françaises (rond) en isolation (traits horizontaux) et dans des logatomes CVCVCVCV où C = [p, t, k,  $\kappa$ ]. L'écart-type est de un.

4.2 Aire du triangle vocalique /i, a, u/

Afin d'examiner l'effet du contexte sur la réduction vocalique, l'aire de l'espace vocalique des locutrices françaises natives et des apprenants tchèques est également comparée. La taille de cette aire, exprimée en  $\text{Hz}^2$ , est calculée selon la formule de Heron. Dans un premier temps est calculée la moitié du périmètre du triangle  $s$  :

$$s = (a+b+c)/2,$$

puis est calculée l'aire  $A$  du triangle :

$$A = \sqrt{s (s-a) (s-b) (s-c)}.$$

Le Tableau 6 indique la taille de l'aire (en  $\text{Hz}^2$ ) en fonction du contexte dans lequel les voyelles sont prononcées (en isolation et en syllabe finale pVp, tVt, kVk et RVR). Puisque la valeur absolue de la taille de l'aire n'a pas de sens en soi mais doit être interprétée en fonction de la taille de l'aire des voyelles isolées qui servent de cible (Steinlen 2005), le Tableau 6 indique également la différence de taille de l'aire (en  $\text{Hz}^2$  et en %) des voyelles en contexte et de celle des voyelles isolées.

	FR			TCH		
	Aire ( $\text{Hz}^2$ )	Diff. $\text{Hz}^2$ Iso-CVC	Diff. % Iso-CVC	Aire ( $\text{Hz}^2$ )	Diff. $\text{Hz}^2$ Iso-CVC	Diff. % Iso-CVC
Iso	458.261			553.279		
pVp	420.936	- 37.325	- 8%	489.689	-63.590	-11%
tVt	322.634	- 135.627	- 30%	367.319	- 185.960	- 34%
kVk	414.666	- 43.595	- 9%	467.112	- 86.167	- 16%
RVR	426.480	- 31.781	- 7%	370.230	- 183.049	- 33%

TABLEAU 6 : La taille des aires (en  $\text{Hz}^2$ ) et la différence des aires des triangles acoustiques (en  $\text{Hz}^2$  et en %) calculées à partir des voyelles /i, a, u/ produites par 10 Françaises natives (FR) et 10 apprenants tchécophones (TCH) en isolation (iso) et en syllabes finales pVp, tVt, kVk et RVR.

Le Tableau 6 met en évidence une réduction de l'aire des voyelles produites en contexte par rapport à l'aire des voyelles isolées, comme attendu. Le degré de réduction n'est cependant pas identique pour les deux groupes de locuteurs. En contexte labial, l'aire vocalique des apprenants est réduite de 11% alors que celle des natives est réduite de 8%. En contexte dental et palato-vélaire, elle est réduite de 34% et 16% respectivement chez les apprenants et de 30% et 9% respectivement chez les natives. Enfin, en contexte uvulaire l'aire entre les voyelles /i, a, u/ est réduite de 33% dans la production des apprenants et de 7% seulement dans la production des Françaises natives. La différence de réduction d'aire en contexte uvulaire est illustrée à la Figure 7 qui est un triangle F1/F2 des voyelles /i, a, u/ produites en isolation (trait plein) et en syllabe finale RVR (trait pointillé) par des Françaises natives (à gauche) et des tchécophones (à droite). Ainsi, dans tous les contextes consonantiques étudiés, l'aire vocalique est davantage réduite dans la production de /i, a, u/ par des apprenants que par des natives.

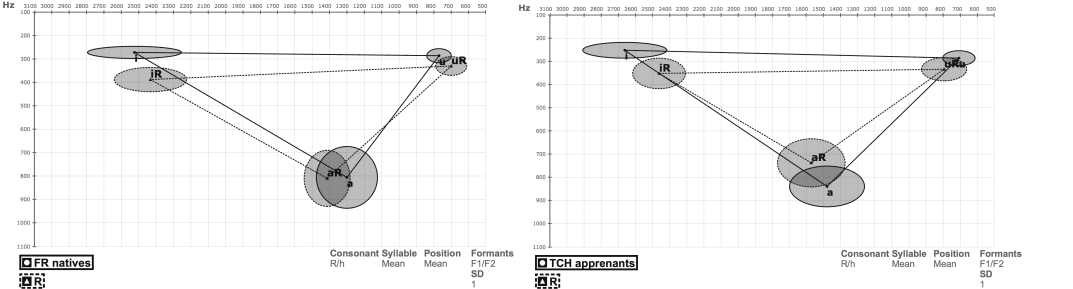


FIGURE 7: L'aire du triangle acoustique (F1/F2) des voyelles /i, a, u/ produites en isolation (trait plein) et en syllabe finale RVR (trait pointillé) par 10 Françaises natives (FR, à gauche) et par 10 apprenants tchécophones (TCH, à droite). L'écart type est de un.

## 5 Conclusion

L'étude acoustique des voyelles quantiques françaises /i, a, u/ a mis en évidence que les apprenants tchécophones ne produisent pas ces dernières avec les mêmes structures formantiques que les Françaises natives, mis à part le [u] en contexte palato-vélaire. Ce résultat est contradictoire à la stabilité acoustique de ces voyelles clamée par la théorie quantique de Stevens (1989) mais soutient de nombreuses études inter-langues (par exemple Gendrot et Adda-Decker 2007). Les écarts en production concernent essentiellement le F1 de [a], en partie corrélé à l'aperture du conduit vocal, et le F2 de toutes les voyelles, en partie corrélé au lieu d'articulation antérieur-postérieur de la voyelle. Ainsi, les apprenants du français doivent viser un F1 plus élevé du [a] en contexte, qui peut être atteint par une plus grande ouverture de la cavité buccale, et un F2 plus élevé des voyelles /u/ et /a/ en contexte dental, qui peut être atteint par une articulation plus antérieure. Dans la production du [a] isolé, les apprenants doivent en revanche parvenir à rapprocher le F1, qui est élevé, et le F2, qui est bas. Le caractère focal du [a] isolé caractérise en effet le français (Vaissière 1985). Dans la production du [i], les apprenants doivent viser un troisième formant maximal qui peut être atteint par une articulation très antérieure. Il est important de rappeler l'existence de compensations articulatoires qui font qu'une même structure formantique peut être obtenue avec des configurations articulatoires différentes (Maeda, 1989). Les stratégies articulatoires présentées ci-dessus peuvent

donc être proposées aux apprenants mais ne doivent pas être considérées comme les seules configurations possibles.

Nous avons également mis en évidence que les zones de difficultés dépendent des contextes consonantiques. Dans certains contextes, les écarts des apprenants par rapport à la référence sont plus grands que dans d'autres contextes. En prenant l'exemple du /u/ qui est le plus parlant, les résultats montrent que les apprenants produisent cette voyelle de manière native en contexte palato-vélaire alors qu'ils devraient viser une valeur du F2 plus élevée de la voyelle produite en isolation et en contextes labial et dental.

Les résultats concernant la production des voyelles en contexte uvulaire suggèrent que les apprenants produisent certaines occurrences du /R/ avec un lieu d'articulation apical, comme pour le /r/ tchèque, et non pas uvulaire, qui est le lieu d'articulation standard de ce phonème en français (Galazzi, à paraître). Malgré l'absence de l'étude articulatoire, ce phénomène est indiqué par un abaissement du F1 et une augmentation du F2 des voyelles /u/ et /a/ en contexte /R/ par rapport à leur cible. Si l'articulation était uvulaire, comme chez les Françaises natives de cette étude, le F1 devrait augmenter et F2 devrait baisser selon les lois acoustiques (Stevens, 1998).

La comparaison des aires vocaliques montre qu'une plus grande réduction de l'espace concerne les voyelles /i, a, u/ produites en contexte dental. Cependant, une réduction d'un tiers environ est également observée pour les voyelles produites en contexte /R/ par les apprenants, mais non pas par les Françaises natives, ce qui peut de nouveau s'expliquer par une réalisation différente de la consonne.

Enfin, il faut bien garder à l'esprit que même si les données acoustiques apportent des résultats concernant des zones de difficulté des apprenants, une vérification perceptive des voyelles produites devrait également être conduite. En effet, quoique les tests statistiques montrent une différence significative entre les formants des voyelles produites par des apprenants et des locutrices natives, il est possible que ces différences ne soient pas toujours perçues. Les graphes de visualisation des résultats montrent que les voyelles /i, a, u/ sont dans la plupart des contextes réalisées par des apprenants avec des formants à un écart type de la référence, ce qui correspondrait selon des critères appliqués dans certaines études acoustiques aux productions à priori fidèles (Birdsong, 2003).

## Références

- ASHBY, M.G. (1990). Prototype categories in phonetics. *Speech Hearing and Language :Work in Progress U.C.L.* 4, 21-28.
- GALAZZI, E. (à paraître). Du locuteur natif à l'étranger expert: quel(s) modèle(s) de prononciation pour les apprenants de FLE dans la société globalisée? In M. Bořek-Dohalská, K. Suková Vychopňová (Eds.), *Didactique de la phonétique et phonétique en didactique du FLE*. Praha : Karolinum.
- BEST, C. T., TYLER, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. J. Munro & O. S. Bohn (Eds.), *Second language speech learning: The role of language experience in speech perception and production*, 13-34. Amsterdam: John Benjamins.

- BIRDSONG, D. (2003). Authenticité de prononciation en français L2 chez des apprenants tardifs anglophones : analyses segmentales et globales. *Acquisition et interaction en langue étrangère*, 17-36.
- BOERSMA, P., & WEENINK, D. (Producer). (2015). Praat: doing phonetics by computer.
- DELATTRE, P. (1951). The Physiological Interpretation of Sound Spectrograms. *PMLA* 66(5), 864-875.
- DELATTRE, P. (1969). An Acoustic and Articulatory Study of Vowel Reduction in Four Languages. *International Review of Applied Linguistics in Language Teaching* 7(4), 295-324.
- DICKERSON, L. B. (1975). The learner's interlanguage as a system of variable rules. *TESOL Quarterly* 9, 401-407.
- DUBĚDA, T. (2005). *Jazyky a jejich zvuky : univerzálie a typologie ve fonetice a fonologii*. Praha: Karolinum.
- FANT, G. (1973). *Speech sounds and features*. Cambridge: MIT Press.
- FLEGE, J. E. (1995). Second language speech learning. Theory, findings, and problems. In W. Strange (ed) *Speech Perception and Linguistic Experience : Theoretical and Methodological Issues in Cross-Language Speech Research*, 233-272. Timonium, MD: York Press Inc.
- FLEMMING, E. (2005). *A phonetically-based model of phonological vowel reduction*. ms. MIT.
- GENDROT, C., ADDA-DECKER, M. (2007). Impact of duration and vowel inventory size on formant values of oral vowels: an automated formant analysis from eight languages. *Actes de Inter. Con. Phon. Sciences*, Saarbrücken: 1417-1420.
- KUČERA, H., GEORGE, K. M. (1968). *A comparative quantitative phonology of Russian, Czech and German*. New York: American Elsevier Publishing Company.
- KEWLEY-PORT, D. (1982). Measurement of formant transitions in naturally produced stop consonant vowel syllables. *Journal of the Acoustic Society of America* 72, 379-389.
- LANDRON, S., PAILLEREAU, N., NAWAFLEH, A., EXARE, C., ANDO, H., GAO, J. (2010). Le corpus PhoDiFLE : un corpus commun de français langue étrangère pour une étude phonétique des productions de locuteurs de langues maternelles plurielles. *Cahiers de Praxématique* 54-55, 73-86.
- LÉON, P. R., LÉON, M. (2007). *La prononciation du français*. Paris: Armand Colin.
- MADDIESON, I., WRIGHT, R. (1996). Small vowel systems and phonetic variability – evidence from Amis. *Austronesian Formal Linguistics Association* 3, 305-309.
- MAEDA, S. (1989). Compensatory Articulation during Speech: Evidence from the Analysis and Synthesis of Vocal-Tract Shapes using an Articulatory Model. In W. J. Hardcastle & A. Marchal (Eds.), *Speech Production and Modelling*: Kluwer Academic Publishers.

- MANUEL, S. Y., KRAKOW, R. A. (1984). Universal and language particular aspects of vowel-to-vowel coarticulation. *Haskins Status Report on Speech Research* 77-78, 69-78.
- MILLER, L. J. (1981). Effects of speaking rate on segmental distinctions. In P. D. Eimas & J. L. Miller (Eds.), *Perspectives on the Study of Speech*. NJ: Erlbaum, 39-73.
- NOVAKOVA, S. (à paraître). La phonétique en cours de FLE : l'actualité tchèque. In M. Bořek-Dohalská, K. Suková Vychopňová (Eds.), *Didactique de la phonétique et phonétique en didactique du FLE*. Praha : Karolinum.
- PAILLEREAU, N. (2015). *Perception et production des voyelles orales du français par des futures enseignantes tchèques de Français Langue Etrangère (FLE)*. Thèse non publiée, Sorbonne-Nouvelle, Paris.
- PAILLEREAU, N. (à paraîtreA). Do isolated vowels represent vowel targets in French? An acoustic study on coarticulation. Actes du *5e Congrès Mondial de Linguistique Française*, Tours, France.
- PAILLEREAU N. (à paraîtreB). "Identical" Vowels in L1 and L2? Criteria and Implications for L2 Phonetic Teaching/Learning. *EUROSLA Yearbook 2016*. Amsterdam: John Benjamins.
- PUSTKA, E. (2011). L'accent méridional : représentations, attitudes et perceptions toulousaines et parisiennes. *Varia* 69, 117-152.
- SKARNITZL, R., VOLÍN, J. (2012). Referenční hodnoty vokálních formantů pro mladé dospělé mluvčí standardní češtiny. *Akustické listy* 18, 7-11.
- STEINLEN, A. K. (2005). *The influence of consonants on native and non-native vowel production. A cross-linguistic study*. Tübingen: GNV.
- STEVENS, K. N., HOUSE, A. S. (1963). Perturbation of Vowel Articulations by Consonantal Context: An Acoustical Study. *Journal of Speech and Hearing Research* 6(2), 111-128.
- STEVENS, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics* 17, 3-46.
- STEVENS, K. N. (1998). *Acoustic Phonetics*. Cambridge: MIT Press.
- ŠIMACKOVA, S., PODLIPSKÝ, V. J., CHLADKOVA, K. (2012). Czech spoken in Bohemia and Moravia. *Journal of the International Phonetic Association* 42(2), 225-232.
- VAISSIÈRE, J. (1985). The use of allophonic variations of /a/ in automatic continuous speech recognition of French. *Journal of the Acoustical Society of America* 77(1), S12-S12.
- VAISSIÈRE, J. (2006). *La phonétique*. Paris: Presses Universitaires de France.
- VAISSIÈRE, J. (2009). Articulatory modeling and the definition of acoustic-perceptual targets for reference vowels. *The Chinese Phonetics Journal* 2, 22-33.
- WIOLAND, F. (2005). *La vie sociale des sons du français*. Paris: L'Harmattan.