

---

## **APPRENTISSAGE DES VOYELLES ORALES DU FRANÇAIS PAR DES ENFANTS TUNISIENS NEO-APPRENANTS DU FRANÇAIS**

**D**ans ce chapitre nous étudions la perception et la production des voyelles orales du français par de jeunes apprenants tunisiens âgés de 8/9 ans et en première année d'apprentissage de français. Ce chapitre s'articule autour de trois questions centrales :

**Quelles difficultés en perception et en production trouvent les néo-apprenants tunisiens au début de leur apprentissage des voyelles orales du français ?**

Dans la section III.1 nous présentons les résultats du premier test T0, mené à 3 mois d'apprentissage. Ce test diagnostique nous permet de pointer les principales difficultés dont font preuve nos apprenants au début de l'apprentissage.

Le model SLM de Flege (1995) prédit qu'un locuteur L1 identifie les phonèmes L2 en termes de catégories L1. Il postule qu'un son « nouveau », inexistant en L1, devrait être facile à apprendre alors qu'il serait plus difficile de produire correctement un son « similaire ». Le model PAM de Best (1995) prédit que la discrimination des sons L2 est faite en fonction de la correspondance des sons L2 avec le système phonémique de la langue maternelle. L'auteur postule que la perception du système non natif s'appuie sur la détection des similarités articulatoires et phonétiques avec les unités phonologiques et les contrastes de la langue maternelle.

Ainsi, nous supposons que les apprenants au début d'apprentissage traiteraient les voyelles du français en fonction de leur système L1. Sous l'influence de leur système maternel, les apprenants classeraient les voyelles orales du français en fonction de leur proximité/divergence par rapport à leurs voyelles L1.

**Quelles progressions enregistrent les apprenants après 9 mois d'apprentissage ?**

Dans la section III.2 nous exposons les résultats du test T2, test survenant en fin d'année d'apprentissage. Notre objectif est d'étudier l'effet « Temps » sur l'apprentissage des apprenants et de comprendre comment évoluent leur perception et leur production des voyelles françaises après 9 mois d'apprentissage. En effet, nous estimons que l'influence de la langue maternelle est plus présente au début de l'apprentissage d'une nouvelle langue. Plus cet apprentissage progresse plus les apprenants acquièrent une meilleure perception et production. Nous nous attendons de ce fait à noter en fin d'année d'apprentissage une amélioration des performances en perception-production de nos apprenants.

### **Quels bénéfices peuvent apporter des entraînements phonétiques sur les performances en perception et en production chez nos jeunes apprenants ?**

Dans la section III.3 nous mettons en avant les résultats du test T1 qui intervient à 6 mois d'apprentissage et suite à une période d'entraînements phonétiques en perception/production des voyelles orales du français appliquée à un groupe expérimental de 8 apprenants parmi les 16 apprenants testés. Ce test nous sert à évaluer l'effet de l'entraînement phonétique sur les performances de nos jeunes apprenants. En effet, nous nous attendons à noter des améliorations significatives en perception et en production pour le groupe d'apprenants ayant suivi des entraînements en comparaison du groupe contrôle (sans entraînement).

## **I. Objectif et description générale de l'expérience longitudinale**

Dans ce chapitre, nous examinons les difficultés en perception et production des dix voyelles orales du français /i e ε a o u y ø œ/ pour des enfants apprenants tunisiens en première année d'apprentissage du français.

Nous avons mené trois tests en trois temps différents sur 9 mois d'apprentissage :

D'abord, **un premier test diagnostic T0** à 3 mois de cours de français (96 heures) afin de relever les difficultés en perception et production pour les dix voyelles orales en début d'apprentissage.

Par la suite, nous avons divisé nos apprenants en deux groupes. Un groupe A expérimental et un groupe B contrôle. Nous avons administré au groupe A **un entraînement phonétique** (pendant 7 semaines) à la suite duquel nous avons appliqué **un deuxième test T1**. Ce test nous permet d'étudier l'effet d'entraînements phonétiques sur les performances des apprenants en perception-production.

Enfin, **un troisième test T2** après 9 mois de cours de français (288 heures) a été administré.

Ce test a pour objectif d'observer la progression de l'apprentissage des apprenants et de comprendre comment évoluent leurs perception et production des voyelles françaises entre le début et la fin de l'année.

## **II. Procédure expérimentale :**

### **II.1. Locuteurs et cadre**

Seize élèves tunisiens<sup>20</sup>, (8 garçons et 8 filles) âgés de 8 à 9 ans, tous natifs de la ville de Sousse, appartenant au même établissement scolaire, ont participé à l'expérience. Ces enfants sont en première année de français<sup>21</sup>.

L'école qui a accepté de nous accueillir est située dans un quartier défavorisé. Les élèves ont peu de contact avec le français en dehors des heures de cours. Les tests ont tous eu lieu à l'école dans une salle de classe isolée et calme. Les apprenants ont passé le test individuellement.

### **II.2. Déroulement des Tests :**

L'expérience s'est déroulée sur 17 semaines comme illustré sur le tableau 5, ce qui nous a permis d'administrer aux enfants trois tests en trois temps différents :

- Un test T0 en début de l'expérience après 96 heures de cours de français. (L'équivalent de 3 mois).
- Un test T1, après des entraînements phonétiques pour la moitié des enfants (groupe A) et pas pour l'autre moitié (groupe B).
- Un test T2 en fin d'année après 288 heures de cours de français (l'équivalent de 9 mois) et des entraînements phonétiques pour les deux groupes d'enfants.

Nous avons divisé nos apprenants en deux groupes : groupe A et groupe B. Le groupe A reçoit des entraînements phonétiques en début de l'année (après le test T0 et avant le test T1) pendant 7 semaines et à la fréquence de deux séances par semaine. Pendant ce temps le groupe B ne suit pas d'entraînement. Enfin d'année (après le test T1 et avant le Test T2), le groupe B, jusqu'à alors simple groupe contrôle, suit les mêmes entraînements présentés au groupe A et pendant la même durée. Pendant ce temps nous avons arrêté les entraînements

---

<sup>20</sup> Annexe 3 pour la description des locuteurs.

<sup>21</sup> A ce niveau d'apprentissage scolaire, les apprenants apprennent seulement l'arabe standard et le français.

pour le groupe A. A la fin de l'année les deux groupes ont bénéficié des entraînements phonétiques.

Groupe A	Groupe B
<b>T0 : Test Diagnostic après 3 mois de cours de français (96 heures)</b>	
<b>Entraînement Phonétique</b> (Sur 7 semaines)	<b>Pas d'entraînement</b>
<b>T1 : Effet des entraînements phonétiques</b>	
<b>Pas d'entraînement</b>	<b>Entraînement Phonétique</b> (Sur 7 semaines)
<b>T2 : Test d'évaluation après 9 mois de cours de français + entraînements phonétiques (288 heures)</b>	

Tableau 5 : Calendrier des tests effectués sur les 17 semaines de l'expérience.

Les entraînements ont tous eu lieu à l'école. Il était convenu avec l'école que nous devons mener nous même les séances d'entraînements et qu'elles devaient impérativement avoir lieu au sein de l'école et pendant la séance du cours de français. Nous décrivons dans la section qui suit le déroulement des entraînements et nous présentons sur les Annexes 02 les documents que nous avons utilisés pour les exercices d'entraînements.

### II.3. L'entraînement phonétique :

#### **Quels objectifs ont les exercices d'entraînement phonétique que nous avons administrés à nos apprenants ?**

Nous avons posé comme objectif pour ces exercices d'entraînement auditif et articulaire, d'amener les apprenants à identifier correctement les voyelles orales du français, distinguer les contrastes du système des voyelles orales du français non ou mal perçus et à produire des voyelles françaises proches acoustiquement de celles des natifs. Nous nous sommes appuyés sur les problèmes interférentiels et les difficultés de perception/production relevées suite aux tests diagnostiques effectués à T0. Il s'agit en effet, non pas « *d'inculquer à l'apprenant de nouveaux traits inexistants dans son système phonologique, mais de lui apprendre à modifier les traits qu'il a l'habitude d'utiliser dans sa propre langue* » (Callamand,1981 :). Les exercices doivent ainsi, permettre à l'apprenant d'ajuster sa perception/production en modifiant en plus ou en moins des traits qu'il connaît déjà.

Ainsi, les exercices d'entraînement que nous avons proposés incitent l'apprenant à prendre conscience des particularités des nouveaux sons qu'il entend et qu'il produit. Il apprend à les distinguer des sons de son système maternel mais aussi à les distinguer les uns des autres.

### **Quelles formes prennent nos exercices d'entraînement phonétique ?**

L'entraînement est fondé sur la perception d'une part et la répétition d'autre part. Les entraînements sont essentiellement à l'oral. Nous avons évité au maximum de recourir au code écrit. Nous avons choisi de présenter et de faire produire les voyelles cibles essentiellement dans des mots monosyllabiques isolés (exercices d'identification et répétition). Nous avons recouru par ailleurs à des mini-dialogues où nous avons inséré les voyelles cibles dans des syllabes accentuées (exercices de production). Toutefois, nous ne concentrons pas l'attention sur le lexique, les entraînements n'ont pas une réelle visée communicative.

Ensuite, nous avons veillé à ce que les voyelles soient présentées dans des « contextes facilitants »<sup>22</sup>. Callamand (1981) présente dans son ouvrage « Méthodologie de l'enseignement de la prononciation » des fiches pédagogiques pour chaque son avec des consignes sur les traits à renforcer pour chaque son, les contextes les plus favorisants et des exemples de difficultés progressives. Nous avons choisi de nous appuyer sur ces fiches pour l'élaboration de nos exercices.

Les entraînements ont suivi un ordre croissant en difficulté. Nous avons adopté la progression proposée par Lauret (2007). Cette progression se présente comme suit : de la perception vers la production ; des structures syllabiques simples aux structures complexes ; du plus accentué au moins accentué et du plus favorisant au moins favorisant.

### **Comment se déroule une séance d'entraînement<sup>23</sup> ?**

Nous avons toujours débuté la séance par des exercices d'échauffement associant la voix et le corps. Des exercices qui associent perception et articulation où l'apprenant prend conscience du geste articulatoire du son nouveau toujours en opposition à un autre son. La gestuelle, aide l'apprenant à s'approprier l'articulation et apporte des précisions sur les contrastes articulatoires entre certaines voyelles.

---

<sup>22</sup> Ce sont des contextes qui favorisent et renforcent un ou plusieurs traits caractéristiques des sons.

<sup>23</sup> Annexe 02 : fiches descriptives du déroulement des séances d'entraînement.

Viennent par la suite les exercices d'entraînement auditif : Nous avons construit nos exercices d'identification, comme nous l'avons mentionné ci-dessus, à partir de paires minimales. A l'écoute de la paire, nous avons demandé si les mots sont identiques ou différents. Les apprenants lèvent des cartons verts pour « identiques » ou rouges pour « différents ». Ensuite, pour les exercices de discrimination, les apprenants écoutent une série de mots principalement monosyllabiques et doivent reconnaître la voyelle cible. Dès qu'ils entendent le mot contenant le son, ils lèvent le carton vert sinon c'est le carton rouge qui est levé. Parfois, nous essayons de rendre ces exercices plus ludiques, en demandant par exemple aux apprenants de se lever s'ils entendent le son et de rester assis s'ils ne l'entendent pas ou d'avancer d'un pas s'ils entendent par exemple un /y/ et de reculer d'un pas s'ils entendent /u/. Dans cet exercice on veille à ce que les stimuli aient les mêmes contours intonatifs et ne varient que par le son qu'on vise à assimiler. L'objectif reste de faire acquérir à l'apprenant les nouveaux sons en oppositions avec d'autres sons. Il est à noter que ces exercices se font sans aucun support orthographique et que la correction est immédiate.

A la suite des entraînements perceptifs, nous avons proposé aux apprenants des exercices de production sous forme de répétitions. Les sons cibles sont insérés dans des syllabes accentuées et en contexte favorisant. L'apprenant écoute donc, puis répète. L'exercice de répétition a pour objectif de favoriser « l'assouplissement de la prononciation » de l'apprenant. Lauret (2007) affirme que la répétition est une notion d'habitude et d'automatisme très importants en prononciation.

Les apprenants répètent d'abord des phrases courtes (trois mots maximum) puis des mini-dialogues sous la forme d'un jeu de rôle. Le jeu de rôle est un bon exercice pour favoriser l'assimilation de caractéristiques rythmiques, dynamiques, accentuelles et mélodiques. L'apprenant parle plus fort, il joue un rôle, il n'a pas peur de l'erreur. En ce qui concerne les consignes données lors des entraînements perceptifs et articulatoires, elles sont données en français et sont accompagnées par des gestes. Par exemple : Je montre l'oreille pour dire « écoutez ». Nous avons précisé à chaque fois le nombre d'items à écouter pour les exercices d'identification et de discrimination. En ce qui concerne les exercices de production, les consignes sont également, courtes et claires.

Finalement, le traitement de l'erreur et la correction sont généralement immédiats.

### **Quelles difficultés rencontrées lors de l'entraînement phonétique ?**

Nous avons été confrontée dans ce type d'exercice à un grand problème de lexique. En effet, nous étions très limitées quant au vocabulaire maîtrisé par nos jeunes apprenants ; et comme

nous voulions que nos exercices d'entraînements s'intègrent parfaitement dans leur programme scolaire, nous avons été contraintes à nous limiter à un lexique familier aux apprenants.

En outre, les apprenants ont eu beaucoup de difficultés quant à l'exercice de jeu de rôle. En effet, cet exercice même s'il reste assez court, avec deux échanges, demande une attention par rapport au sens. De ce fait, le travail sur la prononciation ne peut se faire qu'une fois le dialogue est compris et mémorisé par les apprenants.

## **II.4. Le contenu des tests T0, T1, T2 :**

### **A. La partie perception :**

Le test de perception vise à évaluer les performances des apprenants dans l'identification des voyelles orales du français. Il dure environ 10 minutes pour chaque apprenant. Les apprenants écoutent d'abord la voyelle cible isolée et par la suite, ils écoutent une série de trois mots monosyllabiques CV ou CVC (les contextes consonantiques sont variés et ne sont pas contrôlés). Cette série de mots renferme un mot cible contenant la voyelle cible et deux mots « compétiteurs » contenant une voyelle s'opposant à la voyelle cible dans le même contexte consonantique (triplets minimaux), comme dans l'exemple donné dans le tableau 6. L'apprenant entend d'abord la voyelle cible /i/ ensuite il entend /Ri/ – /Re/ – /Ry/. Il doit identifier le mot contenant la voyelle cible qu'il a entendue.

Quatre séries de triplet (mot cible, compétiteur1, compétiteur2) ont été conçus pour chaque voyelle. Les voyelles des mots compétiteurs s'opposent à la voyelle cible sur 1 ou 2 traits comme indiqué dans le tableau 8 récapitulant toutes les voyelles testées (cibles) et les traits manipulés pour leurs compétiteurs. Nous avons choisi les compétiteurs parmi les voyelles pouvant porter à confusion pour les apprenants. En effet, notre choix pour les voyelles compétitrices était guidé essentiellement par les résultats obtenus suite à la classification par le modèle des voyelles tunisiennes en fonction des voyelles françaises (chapitre1.III). Nous avons estimé les prédictions du modèle comme des confusions possibles qui peuvent survenir lors de la perception des voyelles françaises sous l'influence de la L1. De surcroît, nous avons pris en considération les réponses des enseignants ayant participé à notre enquête (cf. Partie théorique, Chap2.II) sur les confusions qu'ils rencontrent chez les apprenants.

Voyelle cible	Compétiteur1	Compétiteur2
/i/ Le mot cible « riz » /Ri/	/e/ Le mot compétiteur « ré » /Re/	/y/ Le mot compétiteur « rue » /Ry/

Tableau 6 : Exemple de voyelle testée et des ces compétitrices.

Les stimuli ont été produits par une locutrice native du français parisien et enregistrés en chambre sourde au LPP (UMR 7018. Paris). Ils sont présentés dans un ordre aléatoire, le même ordre pour tous les apprenants. Les apprenants notent leurs réponses sur des feuilles de réponse présentées sur le tableau 7. Nous donnons la consigne aux apprenants oralement et en français. Nous nous assurons que l'apprenant a bien compris la consigne en lui donnant un exemple (l'exemple ne fait pas partie du test). A la fin du test de perception nous obtenons un total de 640 réponses (10 voyelles /a i u e ε o ɔ œ ø y/ X 4 exemplaires X 16 locuteurs).



J'écoute  et je mets une <input type="checkbox"/> sous le numéro qui convient :			
	①	②	③

Tableau 7 : Modèle de la feuille de réponse pour le test de perception

Nous présentons ci-dessous un tableau récapitulatif présentant toutes les voyelles testées (cibles) et les traits manipulés pour leurs compétiteurs. L'annexe 03 présente le tableau renfermant tous les triplets présentés pour le test.

Voyelles Cibles	Traits acoustiques / compétiteurs				
	Aperture	Labialité	Antériorité	Antériorité et Labialité	Antériorité et Aperture
/a/	/ε/ / /e/	/œ/	-	/ɔ/	-
/i/	/e/	/y/	-	/u/	-
/u/	/o/ / /ɔ/	-	/y/	-	-
/e/	/i/ / /ε/	-	/o/	/ø/	-
/ε/	/e/ / /a/	-	-	/œ/	-
/o/	/u/ / /ɔ/	-	/ø/	/e/	-
/ɔ/	/o/	-	/œ/	/ε/	-
/œ/	/ø/	/a/	/ɔ/	/ε/	-
/ø/	/œ/ / /y/	-	/o/	/e/	/ɔ/



/y/	/ø/	/i/	/u/	-	-
-----	-----	-----	-----	---	---

Tableau 8 : voyelles cibles et traits acoustiques manipulés pour les compétiteurs de chaque voyelle cible.

## **B. La partie production**

Le test de production consiste en une tâche de répétition immédiate. Il dure environ 3 minutes. Il vise à tester les performances en production des dix voyelles orales du français par nos apprenants tunisiens.

Le corpus est présenté sous forme d'enregistrement de la production d'une locutrice native du français ou les cibles à répéter ont été produites par une locutrice native du français parisien et enregistrées en chambre sourde au LPP (UMR 7018, Paris). La voyelle cible en contexte /pV/ est insérée dans une phrase cadre : « je dis pV trois fois pV ». Par exemple pour la voyelle /i/, la phrase cadre est « je dis pi trois fois pi pi pi ». L'apprenant entend la phrase cadre via un casque puis répète immédiatement. Nous avons retenu pour l'analyse les trois dernières voyelles de la phrase cadre (/i/ dans « pi pi pi »). A la fin du test nous obtenons 480 voyelles testées (3 répétitions X 10 voyelles X 16 apprenants). L'enregistrement est effectué avec un micro-casque, via le logiciel Praat. La fréquence d'échantillonnage est de 44,1 kHz, 16 bits.

## **II.5. Procédure d'analyse des données :**

### **A. Données du test de perception**

Les données de perception ont été extraites manuellement des feuilles de réponses remplies par les apprenants lors des trois tests de perception. Dans chaque test, chaque voyelle a obtenu 64 réponses (16 apprenants X 4 réponses par voyelles).

Nous avons par la suite établi des tableaux de confusions nous permettant d'évaluer la capacité des apprenants à identifier les voyelles cibles de manière pertinente et d'analyser les confusions. En effet, lorsque l'apprenant identifie un phonème cible, nous relevons deux types de réponses : soit il identifie le mot cible contenant la voyelle cible, soit il se trompe et choisit un compétiteur. En face de ces deux réponses de l'apprenant, nous avons deux jugements possibles : soit le mot choisi contient réellement la voyelle cible (c'est une identification correcte), soit le mot choisi contient un compétiteur (c'est une erreur d'identification). Cela donne alors quatre situations différentes :

Situations	Description
Les vrais positifs (vp)	L'apprenant identifie à raison le mot comme contenant le phonème cible.
Les faux positifs (fp)	L'apprenant identifie à tort le mot comme contenant le phonème cible.
Les vrais négatifs (vn)	L'apprenant identifie à raison le mot comme ne contenant pas le phonème cible.
Les faux négatifs (fn)	L'apprenant identifie à tort comme ne contenant pas le phonème cible.

A partir de ces situations nous avons calculé les deux mesures suivantes :

D'une part le **Rappel** correspondant à la proportion des réponses correctes données par les apprenants parmi l'ensemble des réponses correctes possibles. Le rappel est calculé selon la formule suivante :

$$\text{Rappel} = \frac{\text{vrai positif}}{\text{vrai positif} + \text{faux négatif}} \quad (\text{Kuhn}, 2008).$$

Par exemple, si parmi les 64 /a/ cibles dans le test, l'apprenant identifie 50 /a/, ces 50 /a/ sont les « vrais positifs » (vp) et les 14 /a/ qui n'ont pas été identifiés par l'apprenant correspondent au « faux négatifs » (fn). Le score de Rappel calculé selon la formule ci-dessus est de 78%.

**Cette mesure nous permet d'évaluer la capacité de l'apprenant à identifier les voyelles cibles. Nous estimons que si l'apprenant obtient un score de Rappel élevé cela correspond à une bonne performance dans l'identification de la voyelle cible.**

D'autre part, **La Précision** correspondant à la proportion des réponses réellement correctes (réponses pertinentes) parmi l'ensemble des réponses qui sont considérées comme correctes par l'apprenant. La Précision est calculée selon la formule suivante :

$$\text{Précision} = \frac{\text{vrai positif}}{\text{vrai positif} + \text{faux positif}} \quad (\text{Kuhn}, 2008).$$

Par exemple, si l'apprenant identifie 100 voyelles comme étant la voyelle /a/ et que parmi ces 100 voyelles 64 seulement sont réellement des /a/, les 36 autres résultats considérés à tort comme des /a/ sont de « faux positifs » et le score de Précision est de 64% (selon la formule ci-dessus). **Cette mesure nous permet d'évaluer la capacité de l'apprenant à donner uniquement les réponses pertinentes. En effet, un score de Précision élevé indique que**

**l'apprenant est capable d'identifier la voyelle cible et d'exclure les voyelles compétitrices (réponses non pertinentes).**

## **B. Données du test de production**

Les voyelles produites par les apprenants dans la séquence « pVpVpV » à la fin de la phrase cadre ont été extraites du fichier son, segmentées et annotées manuellement à partir du spectrogramme (bande large : fenêtre de 5 millisecondes) sous le logiciel Praat (Boersma et Weenink, 2016). Le début et la fin des voyelles ont été déterminés comme l'onset et l'offset du F2 de ces voyelles. Les mesures acoustiques des quatre premiers formants ont été faites avec un script Praat qui nous a permis d'extraire des valeurs formantiques correspondant à celles se situant au milieu de la voyelle. En cas d'erreurs de détection automatique, les valeurs formantiques sont vérifiées et corrigées manuellement.

Chaque voyelle a été par la suite caractérisée par les valeurs des formants F1, F2, F3 et F4 pour /i/. Nous avons défini aussi, les écarts F1/F2, F2/F3, et F3/F4 pour /i/ (voyelle focale sur F3-F4). Nous avons par la suite établi une moyenne par apprenant sur les trois voyelles produites (pVpVpV) pour chacune des valeurs extraites. Nous avons ensuite, soumis nos données d'apprenants à une analyse de classification effectuée via le logiciel R (Librairie MASS, Venables et Ripley, 2002) selon les étapes suivantes :

- Constitution de l'échantillon d'entraînement : Afin de construire un échantillon d'entraînement à partir des voyelles orales natives, nous avons enregistré 3 enfants francophones natifs du même âge que nos apprenants (voir annexe 03 pour tableau détails sur les locuteurs). L'échantillon contient les dix voyelles orales du français : /a/, /i/, /u/, /e/, /ɛ/, /o/, /ɔ/, /œ/, /ø/ et /y/.
- Nous avons par la suite, établi un modèle de classification, se basant sur l'échantillon d'entraînement et ayant pour variables discriminantes les valeurs moyennes F1, F2, F3 et F4 z-scorées<sup>24</sup>. Ce modèle a été ensuite testé sur les données formantiques de nos apprenants. Le modèle va ainsi classer chacune des voyelles produites par les apprenants dans une des dix catégories vocaliques cibles en appui sur leurs valeurs formantiques. Les résultats obtenus sous forme de

---

<sup>24</sup> Par exemple pour calculer la valeur Z score de F1 de /i/ pour un apprenant nous avons d'abord calculer la moyenne F1 toutes voyelles confondues pour tous les apprenants. Nous divisons par la suite la valeur en Hz de F1 /i/ pour l'apprenant X par cette moyenne. Le résultat obtenu correspond à la valeur z-scorées de F1 de /i/ pour l'apprenant X.

tableau de confusions nous permettent de voir dans quelles catégories des voyelles françaises a été classée chacune des voyelles produites par les apprenants.

Les scores de « classifications correctes » ainsi rendus par le modèle nous renseignent sur les performances en production de nos apprenants. Nous parlons de « classifications correctes » lorsque le modèle fait correspondre la voyelle produite par l'apprenant avec la catégorie de voyelle cible : par exemple lorsqu'un /a/ produit par un apprenant est classé comme /a/ par le modèle. Par contre, si le /a/ testé est classé par le modèle dans une autre catégorie, /e/ ou /ø/ par exemple, nous parlons dans ce cas de « confusions ».

### **III. Résultats des Tests :**

#### **III.1. Perception et production des voyelles françaises chez des enfants tunisiens néo-apprenant du français. Test T0 :**

##### **A. Résultats en perception des voyelles françaises à T0 :**

###### *a. Analyse des résultats T0 en perception par statut de voyelle*

Rappelons que nous avons classé les voyelles orales du français en fonction de leur similarité/divergence par rapport aux voyelles tunisiennes en trois catégories de statut : les voyelles « identiques, id » /i,a,u/, les voyelles « similaires, Sim » /e, ε, o/ et les voyelles « nouvelles, Nv » /y, ɔ, œ, ø/. Nous avons appliqué les mesures de Rappel et de Précision (cf. définition dans la partie précédente) aux scores obtenus pour chaque catégorie de statut. Ces mesures nous ont permis de conclure sur la capacité des apprenants à identifier les voyelles cibles de manière pertinente et à exclure les réponses non pertinentes. Nous avons par ailleurs, analysé les confusions relevées entre catégories de statut.

###### **a.1. Le Rappel :**

Sur le tableau 9, nous avons classé les scores obtenus dans chaque catégorie de statut en relevant d'une part, le nombre des voyelles cibles identifiées par les apprenants (vrai positif=vp) et d'autre part, le nombre des voyelles cibles non identifiées par les apprenants (faux négatif =fn). Nous avons par la suite calculé le Rappel pour chaque catégorie de statut.

<b>Catég de statut</b>	<b>Voyelles cibles identifiées (vp)</b>	<b>Voyelles cibles non identifiées (fn)</b>	<b>Nombre total des voyelles cibles à identifier</b>	<b>Le Rappel <math>\frac{vp}{vp + fn}</math></b>
<b>Identiques (id)</b>	173	19	192	<b>90%</b>

<b>Nouvelles (Nv)</b>	151	105	256	<b>59%</b>
<b>Similaires (sim)</b>	115	77	192	<b>60%</b>

Tableau 9 : les scores en nombre des voyelles cibles identifiées (vrai positif =vp) et non identifiées (faux négatif = fn) ainsi que le % du Rappel pour chaque catégorie de statut calculés à T0.

La comparaison des scores des vrais positifs entre les trois catégories de statut montre une différence significative entre voyelles identiques, similaires et nouvelles ( $F(1,160) = 31$ ;  $p < 0.0001$ ). Les scores des vrais positifs sont significativement plus bas pour les voyelles nouvelles ou similaires par rapport aux voyelles identiques qui obtiennent un score de rappel de 90%. ( $p < 0.001$  pour id vs. Nv,  $p < 0.001$  id vs. Sim). Les voyelles nouvelles ne sont pas mieux identifiées que les voyelles similaires et pour ces deux catégories, environ 60% des voyelles ont été correctement identifiées ( $p = 0.85$  pour voyelles Nv vs. voyelles Sim).

**Les valeurs de Rappel ainsi obtenues attestent que les apprenants réussissent bien à identifier les voyelles cibles « identiques ». Leur capacité à identifier les voyelles cibles est moindre lorsque celles-ci sont des voyelles « nouvelles » ou « similaires ». Les scores de Rappel quasi égaux entre ces deux catégories de statut, montrent par ailleurs, que les voyelles « nouvelles » n’ont pas été mieux perçues que les voyelles « similaires ».**

### a.2. La Précision

Sur le tableau 10, nous avons classé les scores de Précision obtenus dans chaque catégorie de statut en relevant d’une part, le nombre des voyelles cibles identifiées par les apprenants (vrai positif = vp) et d’autre part, le nombre des voyelles identifiées à tort comme des voyelles cibles.

<b>Catég de statut</b>	<b>Voyelles cibles identifiées (vrais positifs = vp)</b>	<b>Voyelles identifiées à tort comme voyelles cibles (faux positifs = fp)</b>	<b>La Précision <math>\frac{vp}{vp + fp}</math></b>
<b>Identiques</b>	173	53	<b>76%</b>
<b>Nouvelles</b>	151	93	<b>62%</b>
<b>Similaires</b>	115	70	<b>62%</b>

Tableau 10 : les scores à T0 en nombre des voyelles cibles identifiées (vp) et des voyelles identifiées à tort comme voyelles cibles (fp) ainsi que le score de Précision pour chaque catégorie de statut.

La comparaison des scores de Précision entre les trois catégories de statut montre un effet significatif du statut ( $F(1,20) = 6,2$  ;  $p = 0.002$ ). Les voyelles identiques avec 76% enregistrent ainsi un score de Précision significativement plus élevé que les deux autres catégories ( $p < 0.001$  pour id vs. Nv et  $p = 0.009$  pour id vs. Sim). Les voyelles nouvelles et similaires présentent des scores égaux (62% respectivement avec  $p = 0.3$  Nv vs. Sim).

**Les apprenants semblent donc avoir moins de difficulté à exclure les réponses non pertinentes lorsqu'il s'agit de voyelles existantes dans leur système phonologique L1 (id) que lorsqu'il s'agit de voyelles non existantes dans leur système phonologique L1 (Sim et Nv).**

Malgré ces disparités entre catégories de statut, il est à remarquer que nous n'enregistrons pas de scores inférieurs à 59% (ni en Rappel ni en Précision). Ceci traduit des performances générales assez bonne pour des apprenants en début d'apprentissage.

Il nous semble de ce fait important de poursuivre notre analyse en examinant les confusions possibles qui surviennent entre catégories de statut.

### **a.3. Les confusions**

En appui sur les postulats du modèle SLM de Flege d'une part et sur les patterns d'assimilation perceptive du modèle PAM de Best d'autre part, nous avons établi cinq contrastes non natifs (opposition ou confusion entre catégorie de statut) et nous avons défini les prédictions de discrimination pour chacun de ces contrastes. Nous répertorions dans le tableau 11 ci-dessous, ces contrastes ainsi que leurs prédictions.

<b>Confusions<sup>25</sup></b>	<b>Equivalence dans le modèle PAM</b>	<b>Prédiction de discrimination</b>
<b>nouvelles &lt;&gt; nouvelles</b>	<b>Le type Uncategorizable-Uncategorizable (UU) :</b> Les deux sons non-natifs ne sont pas assimilés à des catégories natives spécifiques.	De mauvaise à très bonne, selon la proximité entre les sons et aux catégories natives.
<b>nouvelles &lt;&gt; identiques</b>	<b>Le type Uncategorized-Categorized (UC) :</b> L'un des sons non-natifs est assimilé à une catégorie native, alors que l'autre ne l'est pas.	Discrimination très bonne.
<b>nouvelles &lt;&gt; similaires</b>		
<b>similaires &lt;&gt; identiques</b>	<b>Le type Category-Goodness (CG) :</b> Les deux sons non-natifs membre du contraste sont assimilés à la même catégorie native,	De moyenne à très bonne en fonction du rapprochement de chacun de ces sons du bon

<sup>25</sup> Le symbole <> signifie que la confusion est dans les deux sens. Par exemple « Nv<>id » signifie identique (id) prise pour nouvelle (Nv) ET nouvelle (Nv) prise pour identique (id).

	mais différent dans l'écart par rapport à « l'idéal natif » (ex. l'un est acceptable, l'autre est déviant).	exemplaire.
<b>similaires &lt;&gt; similaires</b>	<b>Le type Single-Category (SC) :</b> Les deux sons non-natifs sont assimilés à la même catégorie native, mais sont loin de « l'idéal natif », (les deux sont acceptables ou déviants à part égale).	Discrimination mauvaise.

Tableau 11 : Les contrastes inter-catégories de statut testés, leurs équivalences dans le modèle PAM et leurs prédictions de discrimination perceptive.

Nous représentons sur la figure 33 les taux de confusion relevés entre catégories de statut. Le taux est représenté ici en pourcentage<sup>26</sup>.

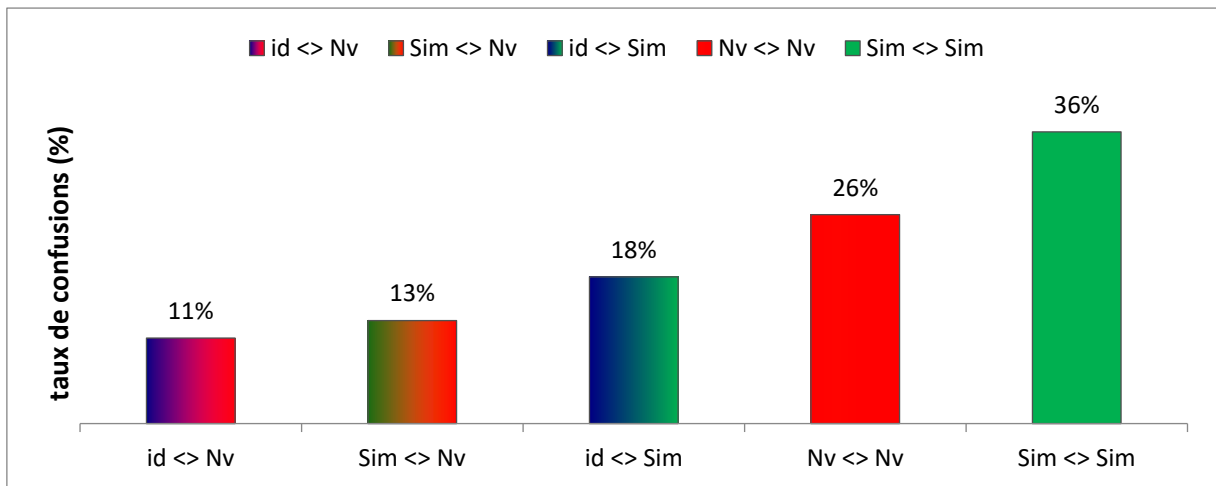


Figure 33 : taux de confusions entre catégories de statut. (Identiques = id (en bleu) ; nouvelles = Nv (en rouge) et similaires = Sim (en vert))

Les contrastes entre les paires opposant id<>Nv et Sim<>Nv, pour lesquels nous avons prédit une discrimination très bonne, montrent très peu de confusions avec 11% et 13% respectivement. Les apprenants montrent de ce fait moins de difficultés à différencier une voyelle nouvelle inexistante dans leur système L1 d'une voyelle existante dans leur système L1 à l'identique ou comme variante allophonique (i.e. Sim). Les paires id<>Sim présentent 18% de cas de confusions. Les prédictions pour ces paires sont de moyenne à très bonne. Avec des confusions peu nombreuses nous estimons que leur discrimination est bonne. Les

<sup>26</sup> Ici et puis pour tous les taux de confusion nous avons divisé le nombre de confusions relevés entre paire de catégorie de statut par le nombre de confusions maximum pour chaque paire de catégorie de statut. id<>Nn n= 320 ; id<>sim n= 192 ; Nv<>Nv n=208 ; sim<>Nv n=336 ; sim<>sim n=80.

paires opposants deux voyelles appartenant à la même catégorie de statut, c'est-à-dire les paires : Nv<>Nv et Sim<>Sim, sont celles qui enregistrent les taux de confusions les plus élevés avec 26% et 36% respectivement. Ces derniers scores répondent à nos prédictions dans la mesure où nous avons estimé pour ces paires des discriminations plus ou moins mauvaises (avec un nombre de confusions ici plus important). Ces deux dernières paires semblent être plus difficiles à discriminer que les précédentes.

**En résumé, les apprenants font plus de confusions lorsqu'ils ont à discriminer deux voyelles inexistantes dans leur système L1 ou entre deux voyelles qu'ils considèrent en L1 comme variantes d'un même phonème.**

***b. Analyse des résultats T0 en perception par voyelle au sein des catégories de statut :***

Dans cette section nous regardons au sein de chaque catégorie de statut les performances pour les voyelles individuelles afin de tester l'homogénéité des performances dans ces catégories de statut.

**b.1. Le Rappel :**

Le tableau 12 présente pour chaque voyelle cible le nombre d'items cibles identifiées (vp) et le nombre d'items cibles non identifiées (fn). Les taux de Rappel sont représentés sur la figure 34.

Statuts	Voyelles	Items cibles identifiées (vp)	Items cibles non identifiées (fn)
Voyelles identiques	a	64	0
	i	55	9
	u	54	10
Voyelles similaires	e	45	19
	ɛ	35	29
	o	35	29
Voyelles nouvelles	ø	41	23
	y	41	23
	ɔ	37	27
	œ	32	32

Tableau 12 : les scores calculés à T0, en nombre des voyelles cibles identifiées (vp) et non identifiées (fn) pour chacune des voyelles des trois catégories de statut.



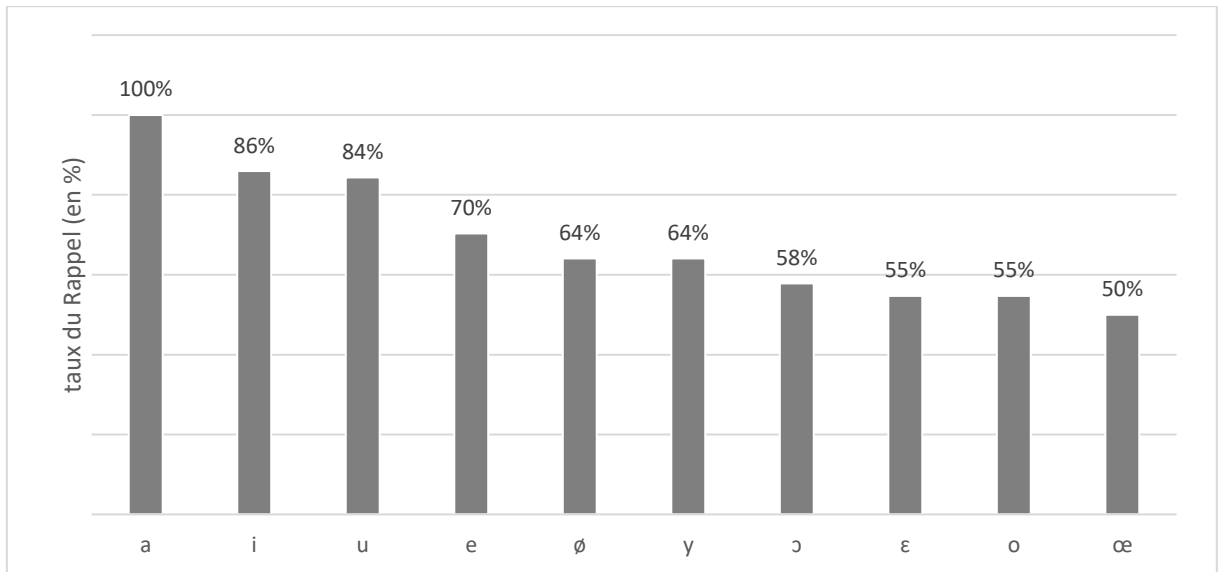


Figure 34 : Taux du Rappel calculé pour chaque voyelle cible à T0. (Pour chaque voyelle n=64)

La voyelle /a/ n'a pas posé de problème d'identification pour les apprenants qui arrivent à identifier tous les items cibles. Les voyelles /i/ et /u/, enregistrent, elles aussi de bons scores avec 86% et 84% respectivement.

Remarquons d'ailleurs, que les voyelles ayant enregistré des scores supérieurs à 80% sont les voyelles appartenant à la catégorie de statut « identique » ayant leurs équivalentes dans le système phonologique de la L1 des apprenants.

En revanche, les voyelles ayant enregistré des scores moins élevés (entre 50% et 70% d'identifications correctes) sont celles qui n'ont pas d'équivalent phonologique dans le système L1 des apprenants à savoir les voyelles « nouvelles » et les voyelles « similaires ». Pourtant, il est intéressant ici de souligner la disparité entre voyelles au sein de ces deux catégories. Par exemple alors que 70% de similaires /e/ ont été bien identifiées, seulement 55% des similaires /o/ et /ε/ l'ont été. De même pour les nouvelles, nous observons 64% de bonne identification pour /y/ et /ø/ contre moins de 60% pour /ɔ/ et /œ/ (avec 58% et 50% respectivement).

### **b.2. La Précision**

Sur le tableau 13, nous avons présenté pour chacune des voyelles des trois catégories de statut, le nombre des items cibles identifiées (vrai positif= vp) et le nombre des items identifiées à tort comme des cibles. Nous avons par la suite calculé la Précision pour chacune des voyelles cibles. La figure 35 présente les scores de Précision pour chaque voyelle cible.

Statuts	Voyelles	Items cibles identifiés (vp)	Items identifiés à tort comme cibles (fp)
Voyelles identiques	A	64	18
	U	54	23
	I	55	12
Voyelles similaires	E	45	32
	ε	35	12
	O	35	26
Voyelles nouvelles	Ø	43	26
	Œ	32	17
	ɔ	37	30
	Y	40	20

Tableau 13 : les scores calculés à T0 en nombre des voyelles cibles identifiées (vrai positif =vp) et des voyelles identifiées à tort comme voyelles cibles (faux positifs = fp).

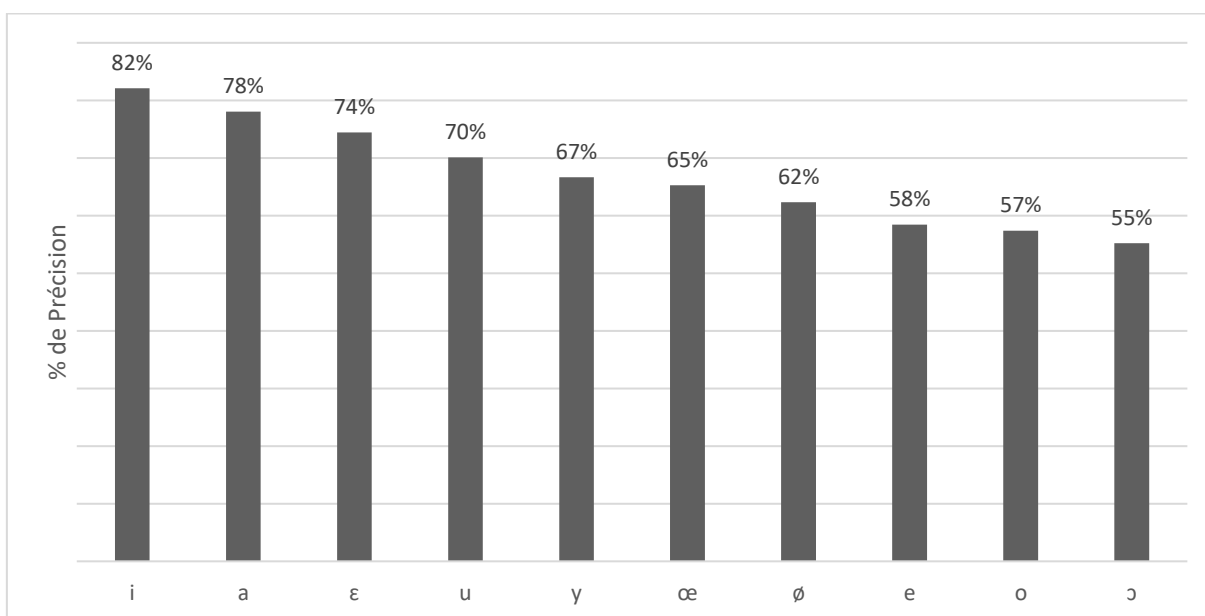


Figure 35 : pourcentage de Précision calculé pour chaque voyelle cible à T0.

L'observation des scores de Précision montre une disparité entre voyelles appartenant à la même catégorie de statut. En effet, Les voyelles /a, i, u/, voyelles identiques et la voyelle similaire /ε/ sont celles qui présentent les scores de Précision les plus importants (plus de 70%). Les voyelles nouvelles /y, ø, œ/ présentent des scores moins élevés avec des taux de Précision entre 60% et 70%. Enfin, les similaires /e, o/ et la nouvelle /ɔ/ présentent des scores en dessous des 60%.

Par ailleurs, la comparaison des résultats en Rappel et ceux en Précision montrent que certaines voyelles ayant enregistré des scores de rappel élevés présentent une Précision plus faible : comme les voyelles /a/, /u/ et /e/. Certaines, présentent au contraire une bonne

Précision contre un taux de rappel faible comme les voyelles /ε/ et /œ/. D'autres, ne présentent pas de grandes différences entre leurs scores de Rappel et ceux de précision comme les voyelles /i/, /y/, /ø/, /o/ et /ɔ/.

### **b.3. Les confusions**

La figure 36 illustre les confusions enregistrées pour chaque paire de voyelles. Nous avons classé nos paires en fonctions du statut des voyelles membre.

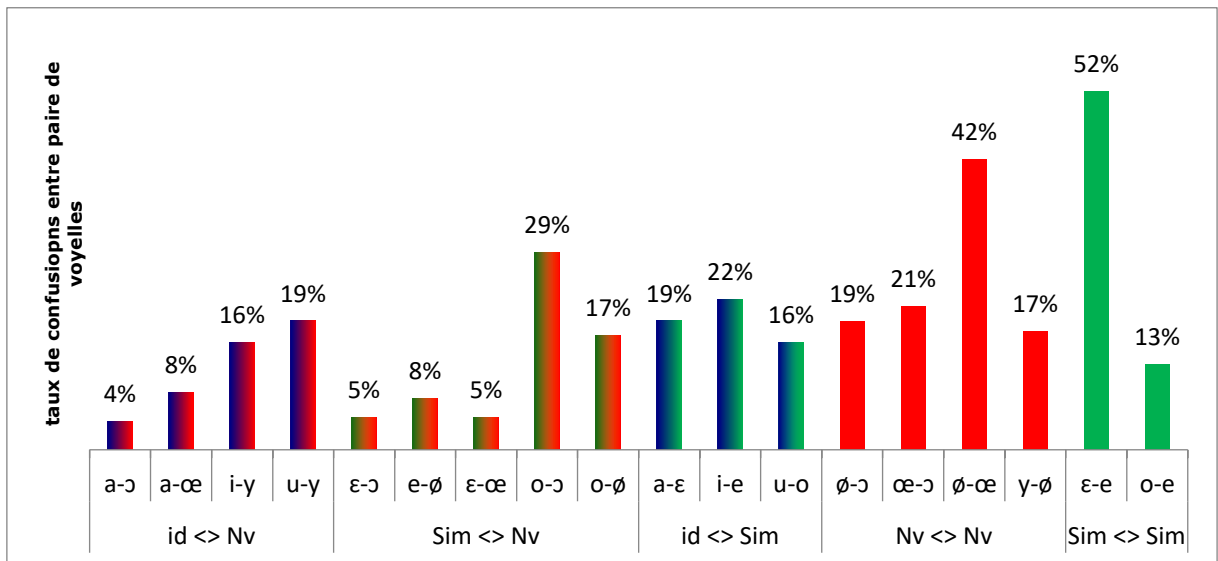


Figure 36 : taux de confusions entre paires de voyelles. Les paires sont classées en fonctions des catégories de statut : Identiques = id (en bleu) ; nouvelles = Nv (en rouge) et similaires = Sim (en vert).

Nous remarquons d'abord que parmi les paires mettant en compétition une voyelle identique et une voyelle nouvelle, les confusions les plus fréquentes se font entre les identiques /i/ et /u/ et la voyelle nouvelle /y/. Les paires où la voyelle /a/ est en compétition avec une voyelle nouvelle présentent très peu de difficultés avec des taux de confusions inférieurs à 10%. Cette voyelle est la voyelle la plus ouverte en dialecte tunisien et ne partage son degré d'aperture avec aucune autre voyelle. Les apprenants ne présentent pas de difficultés à la distinguer d'une voyelle arrondie.

Ensuite, les paires /o-ɔ/ (29%) et /o-ø/ (17%) présentent les scores de confusions les plus élevés parmi les paires mettant en compétition une voyelle similaire et une voyelle nouvelle. Les trois autres paires de la même série (/ε-ɔ/, /e-ø/ et /ε-œ/) présentent des scores de confusions inférieurs à 10%. Les paires id<>Sim, présentent toutes les trois des scores de confusions assez proches avec 19% pour /ε-a/, 22% pour /i-e/ et 16% pour /u-o/.

En outre, les paires mettant en compétition deux voyelles nouvelles montrent une disparité

des scores qui paraît dépendre des voyelles membres de la paire. En effet, nous remarquons des taux de confusions relativement importants pour la paire /œ-ø/ avec 42% suivie de la paire /œ-ɔ/ avec 21% de confusions et finalement les paires /ø-ɔ/, /ø-y/ qui présentent, des taux de confusions moins élevés avec 19%, 17% respectivement. De même pour les paires Sim<>Sim, le taux des confusions diffèrent considérablement selon les voyelles en compétition. Nous notons d'ailleurs, que la paire /e-ε/ est celle qui enregistre le score de confusions le plus élevé (52%).

**En résumé, L'observation des scores de confusion en considérant les voyelles au sein de chaque catégorie, montre d'une part, une disparité au sein des catégories et révèle que ces scores dépendent des voyelles en compétition. D'autre part, les apprenants font preuve de difficultés essentiellement sur les paires de voyelles moyennes /œ-ø/, /o-ɔ/ et /e-ε/.**

## **B. Résultats en production des voyelles françaises à T0**

Nous allons dans cette partie analyser les voyelles produites par nos apprenants lors du test diagnostique T0 à 3 mois d'apprentissage du français. Nous avons d'abord, analysé les résultats de classification acoustique correctes rendues par un modèle établi sur R. Par la suite, nous avons étudié les structures formantiques des voyelles produites par nos apprenants à T0 et nous les avons comparées à celles des natifs du même âge. Nous essayons ainsi de déterminer les différences acoustiques entre apprenants et natifs et comprendre ainsi les difficultés que rencontrent les apprenants lors de leur production des voyelles françaises.

Nous nous attendons à ce qu'en début d'apprentissage les apprenants rencontrent des difficultés en production notamment en ce qui concerne la production des voyelles nouvelles /y, ø, œ, ɔ/ et des voyelles similaires /e, ε, o/.

### ***a. Analyse des classifications acoustiques faites par le modèle à T0***

La figure 37 présente les scores en pourcentages des « classifications correctes » obtenus pour chaque catégorie de statut. Il s'agit des pourcentages des voyelles produites par les apprenants qui ont été correctement classifiées comme étant de bons exemplaires de la voyelle cible.

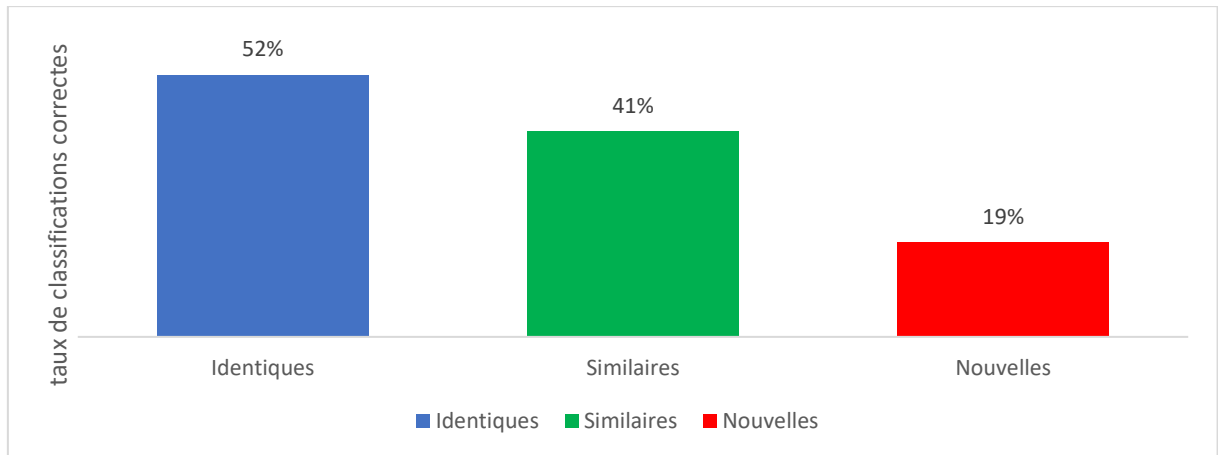


Figure 37 : taux des classifications correctes pour chaque catégorie de statut d'après le modèle. Voyelles identiques : en bleu, N=140 / voyelles nouvelles : en rouge, N=187/ Voyelles similaires : en vert, N=144.

Les taux de classifications correctes tels que rendus par le modèle, montrent qu'en général les scores de classification correctes ne sont pas très bons. Les scores pour les catégories « identiques » et « similaires » révèlent qu'une sur deux voire moins, des cibles sont classées correctement. On observe un score plus faible pour les « nouvelles » avec seulement 19% de classifications correctes.

**Toutefois, avec des scores plus élevés, les voyelles « identiques » et les voyelles « similaires » semblent avoir été produites plus proches des voyelles des locuteurs natifs que les voyelles « nouvelles ».**

Nous avons de ce fait regardé les scores des catégorisations acoustiques faites par le modèle pour chaque voyelle au sein de chacune des catégories de statut.

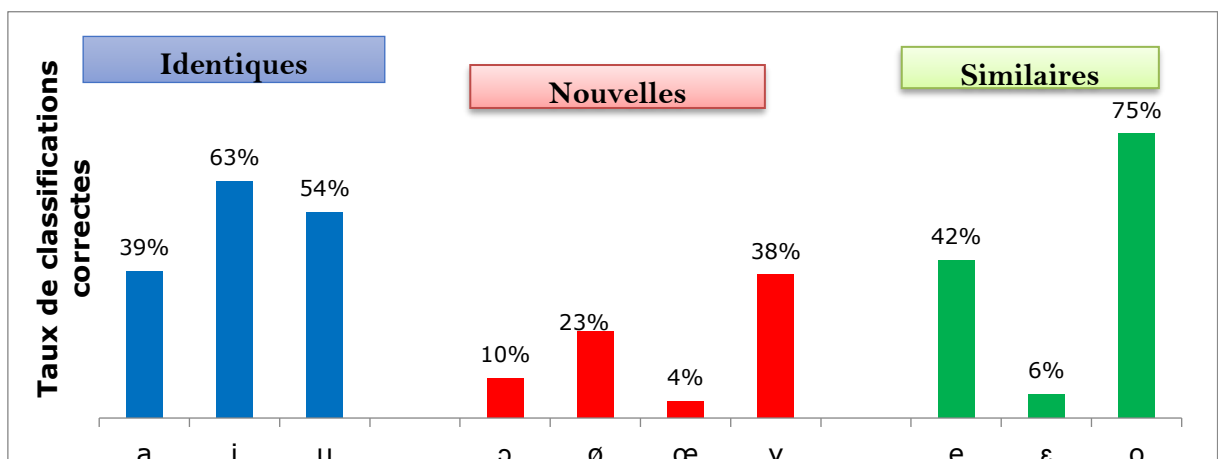


Figure 38 : taux de classifications correctes pour chaque voyelle cible. Les voyelles sont ici regroupées par catégorie de statut : en bleu les voyelles identiques, en vert les voyelles similaires et en rouge les voyelles nouvelles. Analyse basée sur F1,F2,F3.

Les résultats présentés sur la figure 38 montrent de façon générale, une disparité entre

voyelles appartenant à la même catégorie de statut. En effet, si les voyelles identiques /i/ et /u/ marquent des scores de classifications correctes assez élevés avec 63% et 54% respectivement, la voyelle /a/ de même statut n'est catégorisée comme telle qu'à 39%. Pour les voyelles « nouvelles » nous relevons les mêmes disparités, avec d'une part /y/ qui enregistre 38% et /ø/ 23% et d'autre part, les deux voyelles /ɔ/ et /œ/ qui ne sont quasiment jamais bien identifiées acoustiquement comme telles par le modèle, avec 10% et 4% respectivement. Finalement, même constats pour les voyelles « similaires », où 75% des /o/ et 42% des /e/ ont été classés comme de bons exemplaires contre seulement 6% des /ɛ/.

**Ces résultats montrent que les performances de production des apprenants au sein de la même catégorie de statut ne sont pas du même niveau.**

### *b. Analyse acoustique comparative Apprenants vs Natifs*

Nous avons procédé à une analyse des valeurs formantiques moyennes des voyelles françaises produites par nos apprenants. Nous avons ensuite comparé ces valeurs à celles de locuteurs natifs du même âge que nos apprenants.

La figure 39 présente les valeurs formantiques moyennes de F1, F2, F3 pour les dix voyelles cibles et F4 pour /i/ (voyelle focale F3-F4) chez nos 16 apprenants d'une part et celles des natifs d'autre part. Les résultats de confusions rendus par le modèle nous ont guidés sur le choix des confusions à examiner acoustiquement.

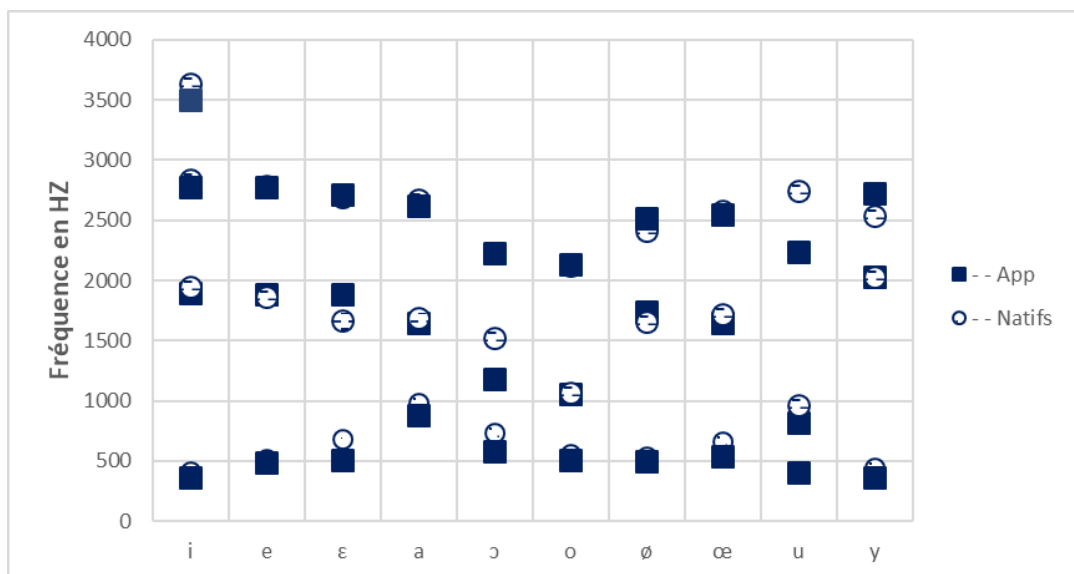


Figure 39 : Valeurs formantiques moyennes (F1, F2, F3) et F4 pour /i/ en Hz des voyelles cibles produites par nos 16 apprenants (carré plein) et par les natifs (rond discontinu).

Une première observation des valeurs formantiques moyennes des voyelles cibles produites par nos apprenants et celles produites par les natifs montre un rapprochement des valeurs moyennes pour certaines voyelles et une différence notable pour d'autres.

Le tableau 14 représente les résultats d'une Anova à un seul facteur testant l'effet « origine du locuteur » (Natifs vs. Apprenants) sur les formants F1, F2, F3 et F4 pour /i/.

<b>/V/</b> <sub>app</sub> <b>VS.</b> <b>/V/</b> <sub>natif</sub>	<b>F1</b>	<b>F2</b>	<b>F3</b>	<b>F1-F2</b>	<b>F2-F3</b>	<b>F3-F4</b>	<b>F4</b>
<b>/i/</b> <sub>app</sub> <b>vs.</b> <b>/i/</b> <sub>natifs</sub>	F (1,70) =4 ; p=0.41	F (1,70) =2 ; p=0.9	F (1,70) =1,06 ; p<0.001*	F (1,70) =8,5 ; p=0.36	F (1,70) =0,39 ; p=0.84	F (1,70) =1 ; p<0.001*	F (1,70) =0,01 ; p<0.001*
<b>/e/</b> <sub>app</sub> <b>vs.</b> <b>/e/</b> <sub>natif</sub>	F (1,70) =3,5 ; p=0.064	F (1,70) =0,34 ; p<0.001*	F (1,70) =0,01 ; p<0.001*	F (1,70) =0,07 ; p=0.76	F (1,70) =0,33 ; p=0.56		
<b>/ɛ/</b> <sub>app</sub> <b>vs.</b> <b>/ɛ/</b> <sub>natif</sub>	F (1,70) =76 ; p<0.001*	F (1,70) =13 ; p<0.001*	F (1,70) =43 ; p<0.001*	F (1,70) =45 ; p<0.001*	F (1,70) =6,73 ; p=0.012*		
<b>/a/</b> <sub>app</sub> <b>vs.</b> <b>/a/</b> <sub>natif</sub>	F (1,70) =13 ; p=0.001*	F (1,70) =0,16 ; p<0.001*	F (1,62) =1 ; p<0.001*	F (1,70) =32 ; p<0.001*	F (1,70) =11 ; p=0.001*		
<b>/o/</b> <sub>app</sub> <b>vs.</b> <b>/o/</b> <sub>natif</sub>	F (1,70) =9 ; p=0.03*	F (1,70) =0,39 ; p<0.001*	F (1,70) = 0,1 ; p<0.001*	F (1,70) =1 ; p<0.001*	F (1,70) =0,49 ; p=0.48		
<b>/ɔ/</b> <sub>app</sub> <b>vs.</b> <b>/ɔ/</b> <sub>natif</sub>	F (1,70) =42 ; p<0.001*	F (1,70) =52 ; p<0.001*	F (1,70) =12 ; p=0.01*	F (1,70) =12 ; p=0.001*	F (1,70) =0,39 ; p=0.84		
<b>/u/</b> <sub>app</sub> <b>vs.</b> <b>/u/</b> <sub>natif</sub>	F (1,70) =18 ; p<0.001*	F (1,70) =7 ; p=0.007*	F (1,70) =34 ; p<0.001*	F (1,70) =7 ; p=0.007*	F (1,70) =18,4 ; p<0.001*		
<b>/ø/</b> <sub>app</sub> <b>vs.</b> <b>/ø/</b> <sub>natif</sub>	F (1,70) =1 ; p<0.001*	F (1,70) =5 ; p=0.27	F (1,70) =2 ; p<0.001*	F (1,70) =5 ; p=0.19	F (1,70) =0,25 ; P<0.001*		
<b>/œ/</b> <sub>app</sub> <b>vs.</b> <b>/œ/</b> <sub>natif</sub>	F (1,70) =4 ; p=0.38	F (1,70) =30 ; p<0.001*	F (1,70) =3,05 ; p=0.85	F (1,70) =0,39 ; P<0.001*	F (1,70) =1 ; p<0.001*		
<b>/y/</b> <sub>app</sub> <b>vs.</b> <b>/y/</b> <sub>natif</sub>	F (1,70) =7 ; p=0.006*	F (1,70) =3 ; p=0.63	F (1,70) =1 ; p<0.001*	F (1,70) =0,01 ; p<0.001*	F (1,70) =7 ; p=0.01*		

Tableau 14 : Résultats de l'Anova à un seul facteur « origine du locuteur » (Natifs/Apprenants) sur les trois premiers formants et les écarts F1-F2, F2-F3 pour les dix voyelles du français et sur F4 et l'écart F3-F4 pour la voyelle /i/. (En gris les différences significatives)

Nous avons dans ce qui suit, regardé ces résultats voyelle par voyelle. Les Figures 40 à 49 représentent les valeurs formantiques moyennes en Hz des voyelles cibles réalisées par les apprenants et par les natifs (moyennés).

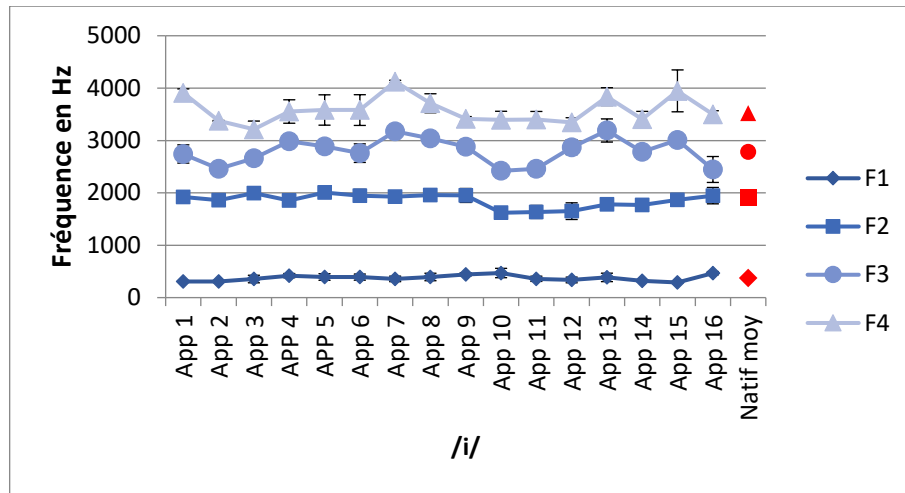


Figure 40 : Valeurs formantiques moyennes en Hz des /i/ réalisés par apprenant, en bleu (moy de 4 réalisations par apprenant) et par les natifs, en rouge (moyenne sur 3 enfants natifs).

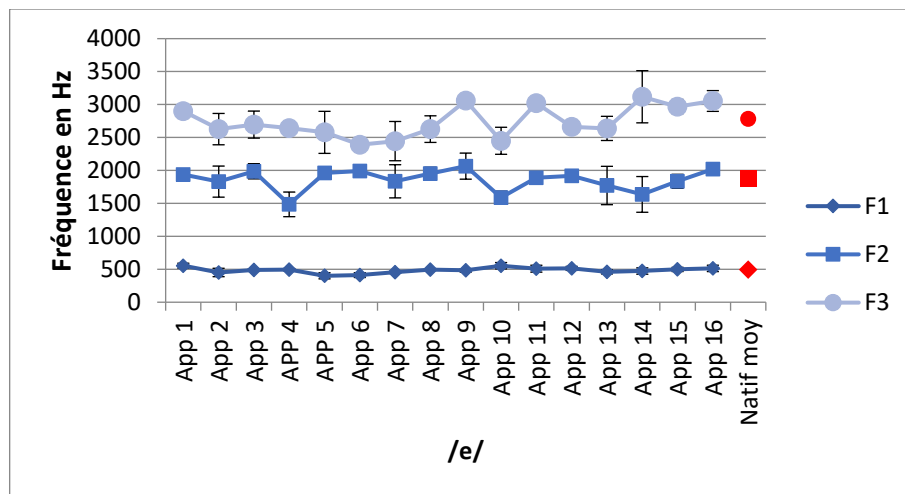


Figure 41 : Valeurs formantiques moyennes en Hz des /e/ réalisés par apprenant, en bleu (moy de 4 réalisations par apprenant) et par les natifs, en rouge (moyenne sur 3 enfants natifs).

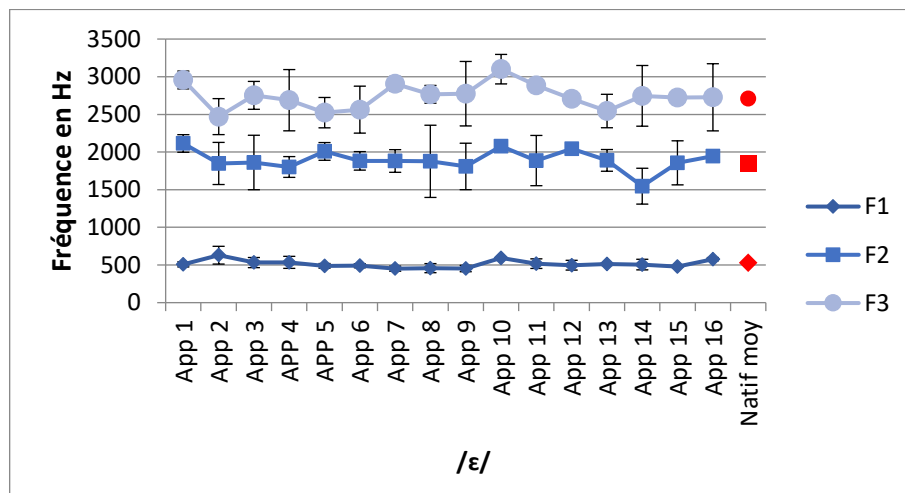


Figure 42 : Valeurs formantiques moyennes en Hz des /ε/ réalisés par apprenant, en bleu (moy de 4 réalisations par apprenant) et par les natifs, en rouge (moyenne sur 3 enfants natifs).



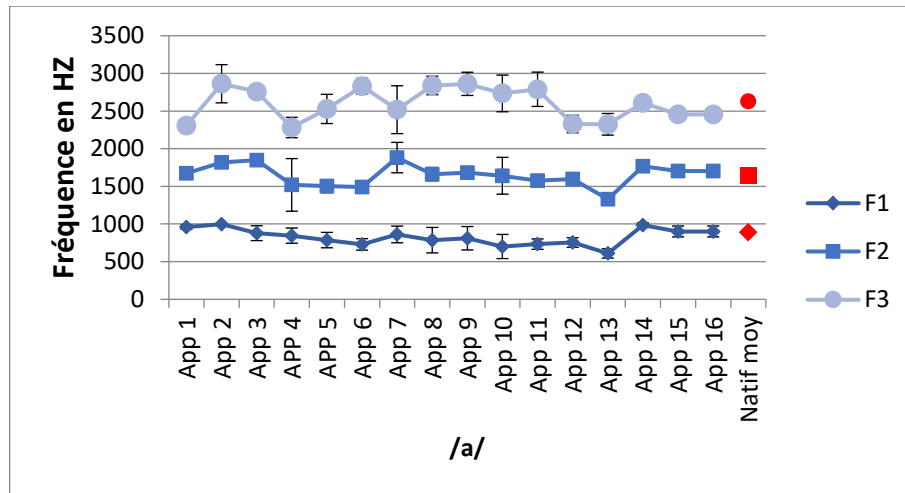


Figure 43 : Valeurs formantiques moyennes en Hz des /a/ réalisés par apprenant, en bleu (moy de 4 réalisations par apprenant) et par les natifs, en rouge (moyenne sur 3 enfants natifs).

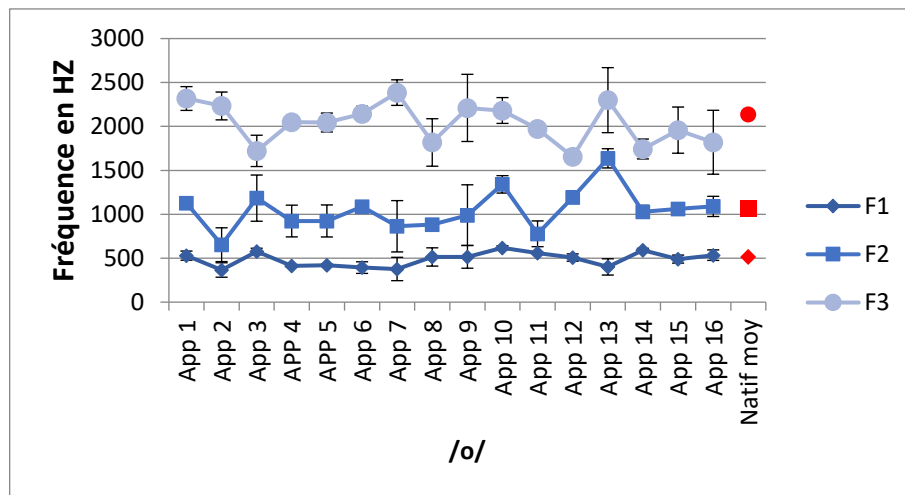


Figure 44 : Valeurs formantiques moyennes en Hz des /o/ réalisés par apprenant, en bleu (moy de 4 réalisations par apprenant) et par les natifs, en rouge (moyenne sur 3 enfants natifs).

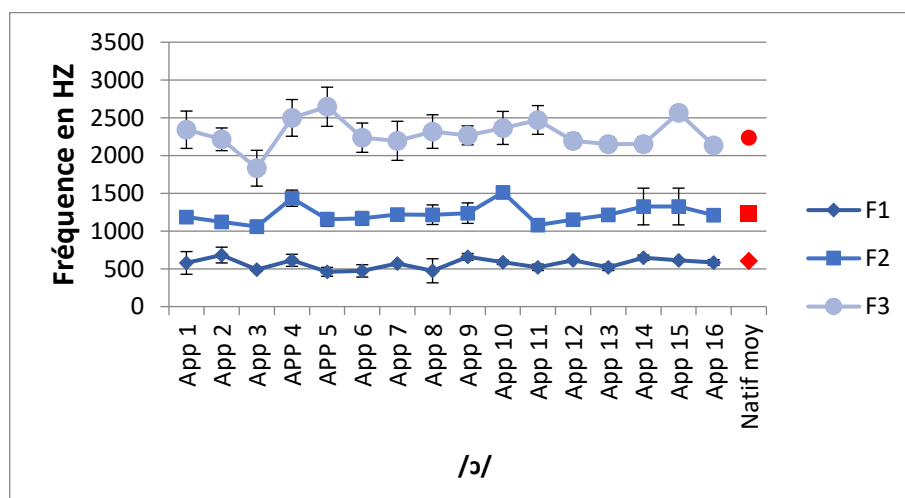


Figure 45 : Valeurs formantiques moyennes en Hz des /ɔ/ réalisés par apprenant, en bleu (moy de 4 réalisations par apprenant) et par les natifs, en rouge (moyenne sur 3 enfants natifs).

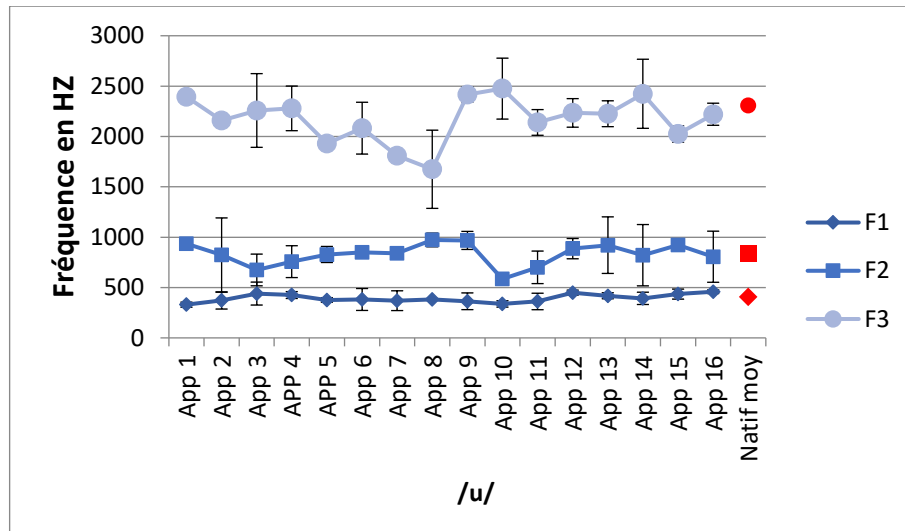


Figure 46 : Valeurs formantiques moyennes en Hz des /u/ réalisés par apprenant, en bleu (moy de 4 réalisations par apprenant) et par les natifs, en rouge (moyenne sur 3 enfants natifs).

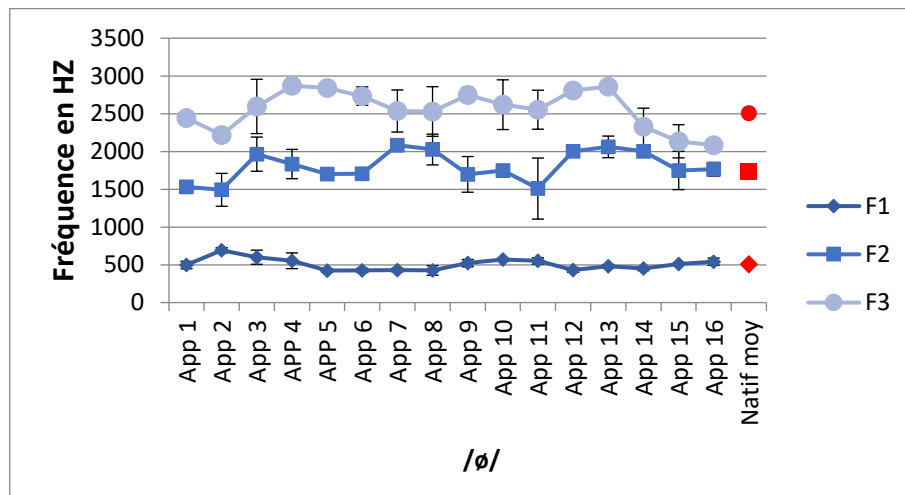


Figure 47 : Valeurs formantiques moyennes en Hz des /ø/ réalisés par apprenant, en bleu (moy de 4 réalisations par apprenant) et par les natifs, en rouge (moyenne sur 3 enfants natifs).

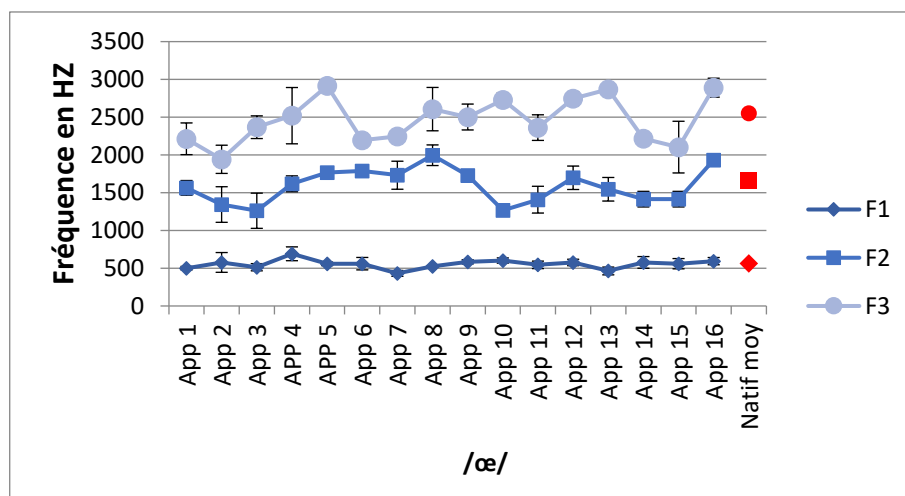


Figure 48 : Valeurs formantiques moyennes en Hz des /œ/ réalisés par apprenant, en bleu (moy de 4 réalisations par apprenant) et par les natifs, en rouge (moyenne sur 3 enfants natifs).

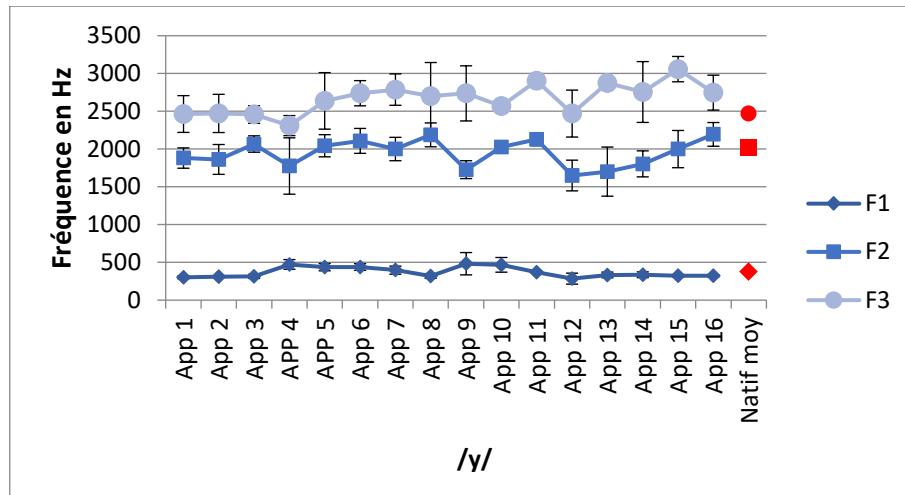


Figure 49 : Valeurs formantiques moyennes en Hz des /y/ réalisés par apprenant, en bleu (moy de 4 réalisations par apprenant) et par les natifs, en rouge (moyenne sur 3 enfants natifs).

**La voyelle /i/ :** Le /i/ en français comme décrit par Vaissière (2011) se caractérise par un rapprochement F3-F4 (focalisation permettant de créer « un pic spectral proéminent » dans les hautes fréquences), un F2 relativement affaibli en intensité et un F1 assez bas.

La représentation graphique des valeurs moyennes des quatre premiers formants de la voyelle /i/ produite par nos apprenants (figure 39) montre un rapprochement F3-F4 notable et des valeurs F1 assez basses.

Les analyses statistiques faites sur ces valeurs et répertoriées sur le tableau 14 ne montrent pas de différences significatives entre les deux groupes de locuteurs pour les deux premiers formants. En revanche, les différences notées sur les valeurs des formants F3 et F4 s'avèrent elles significatives. De plus la distance F3-F4 est significativement différente entre les deux groupes de locuteurs.

La figure 40, illustrant les valeurs moyennes des /i/ produits pour chaque apprenant montre que certains apprenants vont produire des /i/ avec un rapprochement F3-F4 (par exemple les App 3 et 4), d'autres avec un F3 qui se détache nettement du F4 (par exemple les App 1 et 7) avec risque d'une production qui tend vers [e] et d'autres avec un rapprochement F2-F3 (par exemple les App 2, 10 et 16) avec des productions qui tendent vers [y].

Ces variations du formant F3 peuvent être à l'origine des 37% des /i/<sub>app</sub> classés par le modèle dans d'autres catégories que celle correspondant à la voyelle cible. En effet, les réalisations /i/<sub>app</sub> considérées par le modèle comme éloignées de la cible native ont été classées dans trois catégories de voyelles différentes avec 17% classées comme /y/, 13% comme /ø/ et 8% comme /e/.

**La voyelle /e/ :** en français, le /e/ se distingue du /i/ par un F1 plus haut, un F2 plus bas et un F3 qui se détache du F4 et se situe à mi-chemin entre F2 et F4 (Calliope, 1989 ; Vaissière,

2006).

La comparaison entre les valeurs moyennes des /e/ produits par apprenant (tableau 14), montre des différences significatives sur les formants F2 et F3 et pour l'écart F3-F4. Sur la figure 41, on remarque certaines réalisations dans le groupe apprenants où le F2 est plus bas que celui des natifs (App 10 par exemple), d'autres avec un rapprochement notable des F2-F3 (App 4, App 6, App 7 et App 8) ou encore avec un F3 plus élevé que celui des natifs (App 15 et App 16).

Ces différences au niveau des formants F2 et F3 peuvent être à l'origine des classification du modèle de 13% des /e/ dans la catégorie /i/, 19% dans la catégorie /ø/ et de 23% dans la catégorie /y/.

**La voyelle /ε/ :** En français, la mi ouverte /ε/ se distingue de la mi fermée /e/, par un F1 plus élevé et des F2 et F3 plus bas.

La comparaison des valeurs formantiques moyennes des /ε/ réalisés par le groupe apprenants et ceux réalisés par le groupe des natifs révèlent des différences significatives au niveau des trois premiers formants ainsi que pour les écart F1-F2 et F2-F3 (Tableau 14). La figure 42 montre chez certains apprenants des réalisations avec des F2, F3 élevés (App 1, App 10, App11) en comparaison des natifs ou d'autres avec un rapprochement de F2 et F3.

Ces structures formantiques significativement différentes entre apprenants et natifs expliquent les 94% des /ε/<sub>app</sub> classés par le modèle dans d'autres catégories que celle de /ε/<sub>natif</sub>. En effet, le modèle a classé 46% des /ε/<sub>app</sub> dans la catégorie /e/, 21% dans la catégorie /y/ et 12% dans la catégorie /i/.

**La voyelle /a/ :** Voyelle la plus ouverte en français (Vaissière, 2006) qui se caractérise par un F1 élevé et un F2 à mi-chemin entre F1 et F3.

La comparaison des valeurs formantiques moyennes des /a/ réalisés par le groupe apprenants et ceux réalisés par le groupe des natifs révèlent des différences significatives au niveau des trois premiers formants (Tableau 14). En effet, nous observons (figure 43) un F1 moins élevé que celui des natifs chez plusieurs apprenants (Apps 4,5,6,7,8,9,10,11,12 et 13). Nous remarquons notamment des réalisations avec les deux premiers formants plus bas que ceux des natifs.

Ces réalisations peuvent être à l'origine des classements par le modèle de cette voyelle comme /ε/ (32%), comme /ɔ/ (16%) et comme /œ/ (9%) toutes trois voyelles mi-ouvertes, proches de /a/ sur F1.

**La voyelle /o/ :** Cette voyelle rappelons-le est dans la L1 des apprenants une variante possible de /u/. En français les voyelles /o/ et /u/, ainsi que la voyelle /ɔ/ se caractérisent par une faible distance F1-F2 (Vaissière, 2006) et un F2 en dessous des 1200 Hz (Calliope, 1989 ; Georgeton et al, 2012).

L'analyse des valeurs moyennes des /o/ produits par les apprenants et de ceux produits par les natifs révèle des différences significatives entre les deux groupes de locuteurs sur les trois premiers formants ainsi que pour l'écart F1-F2 (Tableau14). L'observation des patrons individuels des apprenants (Figure 48) montre d'une part, certaines réalisations avec un F2 très bas (App2, App 11) et d'autre part des réalisations avec un F2 élevé (App3, App10, App12 et App13). Regardons ici sur la figure 50, la réalisation du /o/ par l'apprenant 2 (3 répétitions). Une réalisation qui tend vers [u] avec un écart F2-F1 très faible (focalisation).

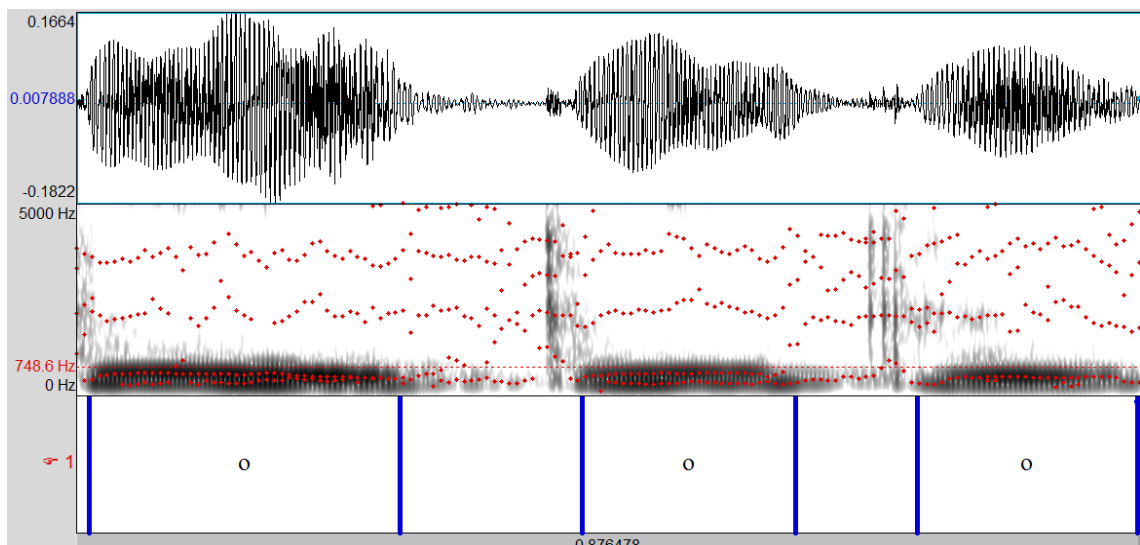


Figure 50 : Spectrogramme (sur Praat) des trois réalisations de la voyelle /o/ par l'apprenant 2.

Par ailleurs, le modèle de classification classe 25% des /o/<sub>app</sub> dans des catégories différentes que celle du /o/<sub>natif</sub> dont 13% classées comme /u/.

**La voyelle /ɔ/ :** Cette voyelle se caractérise en français par un écart faible F1-F2. Elle se distingue des deux autres voyelles postérieures par la valeur F2 la plus élevée qui varie selon la littérature entre les 1000Hz-1400Hz (1180Hz d'après Calliope (1989) ; 1144Hz d'après Georgeton et al (2012) et 1347Hz d'après Gendrot et al (2005)).

La comparaison entre les valeurs formantiques moyennes relevées pour cette voyelle telle qu'elle a été produite par les apprenants d'une part et par les natifs de l'autre, présentées précédemment sur la figure 45 et le tableau 14, s'avère significativement différente aux niveaux des F1, F2, F3 et de l'écart F1-F2.

Le modèle classe d'ailleurs, 90% des /ɔ/app comme différents des /ɔ/natif avec 67% des réalisations dans la catégorie /o/.

**La voyelle /u/ :** Cette voyelle postérieure arrondie se caractérise en français par un rapprochement (focalisation) des deux premiers formants (F1-F2) en dessous de 1000hz avec un F1 très bas (Calliope, 1989 ; Vaissière, 2006 ; Georgeton et al, 2012).

Les valeurs formantiques correspondant aux /u/app et aux /u/natif montrent des différences significatives pour les trois premiers formants et pour l'écart F1-F2 (tableau 14). Ces différences peuvent être la raison des classements par le modèle de quasiment un /u/app sur deux comme correspondant à la catégorie /o/ (46%).

**La voyelle /ø/ :** Les voyelles acoustiquement centrales /ø/ et /œ/ se caractérisent en français par des formants « approximativement équidistants » et par conséquent par des écarts F1-F2 et F2-F3 semblables (Vaissière, 2006). La voyelle mi-fermée /ø/ se distingue tout de même de la mi-ouverte /œ/ par un F1 plus bas.

La figure 47 présente les valeurs formantiques moyennes des voyelles /ø/ réalisées par les apprenants d'une part et par les natifs tous confondus de l'autre. L'analyse de variance menée sur ces valeurs formantiques (tableau 14) ne présente pas de différences significatives entre les /ø/app et des /ø/natif pour le formant F2 ainsi que pour l'écart F1-F2. En revanche, elle présente des différences significatives pour les formants F1 et F3 ainsi que pour l'écart F2-F3. Ces variations peuvent expliquer le classement par le modèle de certains /ø/app dans les catégories /y/ (21%), /i/ (21%) et /e/ (21%).

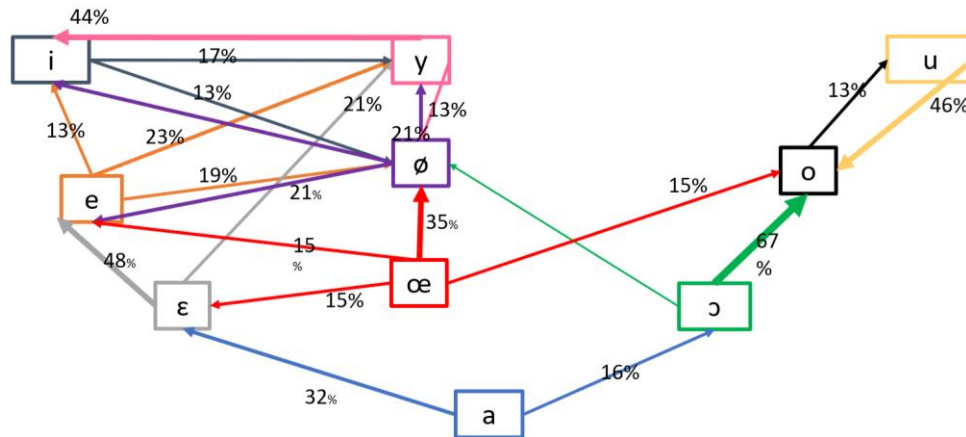
**La voyelle /œ/ :** cette voyelle a été classée par le modèle dans plus de sept catégories de voyelles différentes. En effet, seulement 4% des /œ/app ont été classés comme correspondant aux /œ/natif. La comparaison entre les /œ/ produits par chacun des apprenants d'une part et par les natifs d'autre part montre des différences significatives pour le formant F1 et pour les écarts F1-F2 et F2-F3. Ces variations peuvent être à l'origine des classements par le modèle des /œ/app comme des /ø/ (à 35%) avec un F1 plus bas, comme des /y/ (10%) avec un rapprochement F2-F3 ou encore comme des /e/ (15%) avec un F1 bas et un F3 élevé.

**La voyelle /y/ :** Cette voyelle antérieure arrondie se caractérise en français par un rapprochement des formants F2-F3. Vaissière (2006) insiste sur le rôle du F3 pour distinguer le /y/ du /i/ (allongement de la cavité labiale pour /y/ et par conséquent abaissement du F3 de 3000 à 2000 Hz, Vaissière 2006 :75).

La comparaison entre les valeurs formantiques moyennes des /y/ telles qu'elles ont été produites par apprenant et par les natifs, illustrées précédemment sur la figure 49 et le tableau

14, montre pour les formants F1 et F3 et pour l'écart F2-F3 des valeurs significativement différentes. Ceci peut être à l'origine du classement du modèle des /y/<sub>app</sub> à 44% comme correspondant à des /i/<sub>natif</sub> contre 38% comme des /y/<sub>natif</sub>.

Nous résumons ci-dessous, sous la forme d'un trapèze vocalique, les principales confusions en production relevés par le modèle. (Chaque voyelle est représentée par une couleur et les flèches présentent le sens de la confusion)



### C. Conclusion sur les résultats du test T0 en perception et en production des voyelles françaises chez des enfants tunisiens néo-apprenants du français.

Dans cette partie nous avons présenté les résultats du test T0, test diagnostique effectué à 3 mois d'apprentissage sur la perception et la production des dix voyelles orales du français par 16 enfants tunisiens néo-apprenants du français. Les résultats de ce test ont révélé que :

Les apprenants enregistrent de meilleures performances en perception et en production pour les voyelles « identiques ». Ils ont de bonnes performances pour les voyelles « similaires » plutôt en perception qu'en production. En revanche, les apprenants éprouvent des difficultés dans la perception et production des voyelles « nouvelles ».

Ceci étant, nous avons souligné une disparité entre les voyelles appartenant à la même catégorie de statut : en effet, en ce qui concerne les voyelles identiques, alors même que la voyelle /a/ enregistre en perception les meilleurs scores d'identifications correctes, c'est la voyelle /i/ qui présente en production le score de bons exemplaires le plus élevé. La voyelle /u/ marque en perception les scores les moins importants mais présente en production de bons scores, meilleurs que ceux enregistrés pour /a/. Ensuite, dans la catégorie des voyelles « similaires », les voyelles /e/ et /o/ sont les mieux identifiées perceptivement et les mieux classées acoustiquement comme de bons exemplaires par le modèle, face à des scores

moindres pour la voyelle /ɛ/. En outre, la voyelle /y/ est celle parmi les voyelles nouvelles qui semble poser le moins de difficultés aussi bien en perception qu'en production. Par contre, les voyelles /ɔ/, /œ/ et /ø/ présentent de meilleurs scores en perception qu'en production avec plus d'un item sur deux correctement identifié perceptivement mais moins de 25% d'items classés comme de bons exemplaires acoustiques en production.

Par ailleurs, les résultats en perception révèlent que les voyelles « identiques » ont été plus confondues avec des « similaires » qu'avec des « nouvelles ». Les voyelles « similaires » et les voyelles « nouvelles » se confondent principalement avec des voyelles de leurs catégories respectives. L'observation de ces confusions en considérant les voyelles au sein des paires montre que les scores en confusions dépendent des voyelles elles-mêmes plus que des catégories de statut de ces voyelles. Nous remarquons par exemple que les paires des voyelles moyennes sont celles qui enregistrent les scores de confusions les plus élevés avec 52% pour /e-ɛ/, 42% pour /œ-ø/ et 29% pour /o-ɔ/. Les apprenants enregistrent également des scores de confusions dépassant les 20% sur les paires /ɛ-a/, /i-e/, /u-y/, /ø-ɔ/ et /œ-ɔ/ ce qui traduit des difficultés de distinctivité entre paires appartenant à des régions acoustiques proches. Enfin, l'analyse acoustique des va

leurs formantiques moyennes des voyelles françaises telles produites par nos apprenants dévoile les chevauchements produits par les apprenants lors de leurs réalisations de voyelles proches acoustiquement. Nous relevons entre autres, des réalisations de /a/ qui tendent vers [ɛ], des /u/ vers [o], des /e/ vers [ø] et [y], des /ɛ/ vers [e] et [y], des /ø/ vers [e] ou encore des /y/ qui tendent vers [i]. Nos résultats en acoustique viennent en outre, appuyer les résultats rendus par le modèle de classification où nous observons par exemple plus d'un tiers des /y/<sub>app</sub>, /u/<sub>app</sub> et /a/<sub>app</sub> classés respectivement comme des /i/, /o/ et /ɛ/ natifs par le modèle de classement.