



La recherche de l'optimale corrective pour la création d'un laboratoire de langue en ligne

Laure Fesquet

Fonetix, France

laure@fonetix.fr

<https://orcid.org/0009-0000-1526-5907>

Sébastien Palusci

Fonetix, France

sebastien@fonetix.fr

<https://orcid.org/0009-0000-7895-3651>

Reçu le 15-10-2025 / Évalué le 10-11-2025 / Accepté le 01-12-2025

Résumé

Cet article examine dans quelle mesure la notion d'optimale corrective, formulée par Petar Guberina dans le cadre de la Méthode verbo-tonale d'intégration phonétique (MVT), peut servir de principe directeur pour la conception d'un laboratoire de langue en ligne destiné à l'entraînement autonome à la prononciation en français langue étrangère. À partir d'une démarche de praticiens s'appuyant sur la recherche et sur plusieurs années d'expérimentation de terrain, nous explicitons les choix pédagogiques et technologiques retenus pour proposer des activités visant à agir sur la perception de l'apprenant, en particulier sur les paramètres prosodiques, préalable au travail segmental. Nous montrons comment la recherche d'optimales correctives a guidé l'élaboration des énoncés, des tâches proposées et de l'interface, et comment ce cadre a également permis la création de parcours adaptés aux difficultés phonétiques propres à différentes langues maternelles. Enfin, l'article discute les perspectives de développement, notamment scientifiques.

Mots-clés : prononciation, méthode verbo-tonale, optimale corrective, prosodie, apprentissage en autonomie

Searching for the Corrective Optimum in the Design of an Online Language Laboratory

Abstract

This article examines how the notion of corrective optimum, formulated by Petar Guberina within the Verbo-tonal Method of phonetic integration (MVT), can serve as a guiding principle for the design of an online language laboratory dedicated to autonomous pronunciation training in French as a foreign language. Drawing on a practitioner-based approach informed by research and several years of classroom experimentation, we detail the pedagogical and technological decisions made to develop

activities that target learners' auditory perception, particularly prosodic parameters, as a prerequisite for segmental accuracy. We show how the search for corrective optima informed the selection of utterances, task design and interface, and how this framework supported the development of learning paths tailored to phonetic difficulties associated with different native languages. Finally, the article discusses future perspectives, including scientific developments.

Keywords: pronunciation, verbo-tonal method, corrective optimum, prosody, autonomous learning

Introduction

Pour qui s'intéresse au travail de la prononciation en langue étrangère, le nom de Petar Guberina, fondateur de la Méthode verbo-tonale d'intégration phonétique (désormais MVT), est généralement associé aux principes qui président à cette approche : travail sur la perception et non sur l'articulation d'une part, rôle central joué par la prosodie d'autre part.

Pourtant, à la relecture des articles dans lesquels il développe les différentes modalités de travail visant à l'amélioration de l'oral tant des déficients auditifs que des apprenants de langue étrangère, on perçoit très vite l'originalité de son approche, que l'on pourrait résumer ainsi : comprendre les mécanismes qui entraînent les difficultés de production orale et trouver l'aide, on parlerait aujourd'hui d'étayage, qui permettra de surmonter les difficultés de perception qui en sont la cause. Il semble d'ailleurs que P. Guberina considérait que c'était là son apport majeur : la notion d'optimale comme « recherche continue des moyens les plus adaptés afin de fournir [...] à l'apprenant les stimuli les plus efficaces et les plus susceptibles d'être intégrés » (Roberge, 2003 : 15). Partant d'un problème « de terrain », il convoque pour y répondre de nombreuses disciplines allant de la phonétique acoustique à la neuropsychologie, transformant en pratiques de classe et en outils (les SUVAG) les conclusions de ses recherches.

Ce même souci de recherche des optimales correctives du fondateur infuse, nous semble-t-il, dans les pratiques et réflexions de l'ensemble des verbo-tonalistes. Nous inscrivant de par nos études comme notre pratique dans ce courant, l'idée d'explicitier les choix des solutions pédagogiques dans la conception d'un « laboratoire de langue numérique » avec comme

grille de lecture cette notion d'optimale nous permettra de rendre visible l'héritage de Petar Guberina dans notre approche.

Après avoir présenté brièvement le contexte de création du laboratoire de langue virtuel, nous verrons les éléments directement issus des apports de la MVT, notamment pour la correction des sons. Nous présenterons ensuite des choix pédagogiques qui, s'ils pourront sembler s'éloigner des fondamentaux de la MVT répondent selon nous au même souci de fournir aux apprenants les stimuli optimaux pour compenser leurs difficultés de perception dans l'apprentissage du français langue étrangère.

1. Contexte d'émergence du laboratoire de langue Fonetix¹

1.1. L'aboutissement d'une démarche collective d'enseignants

Afin de comprendre ce qui nous a amenés à créer un espace d'entraînement autonome et l'objectif poursuivi, il convient de revenir sur nos parcours et sur les principes pédagogiques qui ont façonné notre manière d'aborder la prononciation.

Formés à la Méthode Verbo-Tonale à l'université Toulouse 2 par Michel Billières, nous avons par la suite eu l'opportunité de mener des ateliers bi-hebdomadaires de correction phonétique à l'Alliance Française de Toulouse, pendant de nombreuses années. Issue de la MVT, notre pratique de la remédiation phonétique est centrée sur la perception plutôt que sur un guidage des mouvements articulatoires en production.

Le désir d'améliorer l'étayage apporté aux apprenants, nos échanges fructueux et les réguliers retours vers la théorie (didactique, phonétique du français) ont débouché sur une prise de conscience croissante du rôle de la prosodie, non plus comme outil dont la manipulation des paramètres permet la correction d'éléments segmentaux, mais comme l'élément central d'une réelle amélioration des compétences orales.

Le rythme, particulièrement structurant tant en production qu'en compréhension, est notamment au cœur de notre pratique.

Tant que l'élève demeurera insensible au rythme parolier de la L2, tout autre travail phonétique est vain, notamment l'installation des segments

¹ <https://fonetix.org/>

vocaliques et consonantiques qui se fera sur un socle branlant sur lequel toute consolidation est pratiquement vouée à l'échec. (Billières, 2016).

Guberina (2003) insistait d'ailleurs sur l'importance d'une reproduction la plus fidèle possible du rythme et de l'intonation, plus facile selon lui à reproduire que les sons de parole car transmis par les fréquences basses, mais aussi condition d'une production correcte des sons de parole.

S'il jette les bases d'un travail spécifique sur la prosodie, incluant les mouvements corporels et les logatomes, éléments que nous avons repris à notre compte, c'est avant tout l'utilisation de la prosodie à des fins de correction segmentale qui reste la norme dans la pratique de la correction phonétique. Or, dans l'apprentissage des langues, le principe de crible s'étend aussi à la prosodie (Intravaia, 2005) : par exemple, l'accent de durée en fin de séquence rythmique propre au français est souvent mal perçu et reproduit par les apprenants (accent d'intensité, mauvaise position de la syllabe accentuée).

Nous avons ainsi imaginé et mis en place des activités permettant un travail de la prosodie, (Billières *et al.*, 2018) préalable à des activités de type MVT (travail de dialogues, jeux, etc.). Les retours tant des apprenants que des enseignants lors de classes d'observation nous poussent à croire en leur pertinence. Toutefois, quelle que soit la qualité des activités proposées, il manque un espace où l'apprenant puisse consolider ce qu'il a acquis en classe, répéter, comparer, valider, etc.

La création d'un « laboratoire de phonétique dématérialisé pour favoriser des apprentissages en autonomie » (Sauvage, 2019) nous semblait répondre à un besoin tant des apprenants que des enseignants et structures dispensant des cours de FLE. C'est donc en praticiens « éclairés » que nous avons tenté de relever le défi de la création d'une application web permettant aux apprenants de s'entraîner à la prononciation du français en dehors de la classe.

1.2. Technologie et pédagogie de la prononciation

Le déroulement d'une séance de remédiation phonétique selon la Méthode Verbo-Tonale peut être décrit comme une succession d'étapes circonscrivant les rôles respectifs de l'enseignant et de l'apprenant.

L'enseignant commence par proposer un modèle de production. L'apprenant répète, ce qui permet à l'enseignant d'identifier les erreurs segmentales et/ou suprasegmentales présentes dans la réalisation : voyelles déformées, contrastes consonantiques insuffisamment marqués, défauts d'accentuation, contour intonatif inadéquat, etc. À partir de ce repérage, l'enseignant établit un diagnostic et sélectionne les éléments à travailler en priorité. Il ajuste alors le retour en manipulant divers paramètres (amplitude, durée, rythme, intonation, gestes associés, contrastes perceptifs), afin d'optimiser les conditions de perception du son-cible. Une nouvelle phase de répétition permet d'évaluer les effets de cette manipulation du retour. L'enseignant vérifie ensuite si la production obtenue est satisfaisante. Si ce n'est pas le cas, le cycle d'ajustement et de répétition reprend, si la production atteint le niveau attendu, l'enseignant procède à la validation. Ce fonctionnement cyclique – modèle, répétition, diagnostic, manipulation du retour (*feedback*), nouvelle répétition, validation – constitue le cœur du travail verbo-tonal : un ajustement continu, guidé par l'écoute fine de l'enseignant et orienté vers une amélioration progressive de l'intelligibilité.

On conviendra aisément que demander à une solution numérique ne serait-ce que d'approcher une telle tâche relève de ce que l'on nomme aujourd'hui l'intelligence artificielle. Si nous avons eu vent d'études allant dans ce sens, force est de constater qu'à l'heure actuelle aucune solution n'a franchi le seuil des laboratoires de recherches. La tâche est, il est vrai, colossale.

De fait, si de nombreuses études ont permis de mieux cerner la manière d'améliorer l'entraînement perceptif des apprenants, la problématique centrale est à présent de pouvoir automatiquement évaluer - de manière formative - leur prononciation (en production), avec une architecture d'ingénierie pédagogique leur permettant de « progresser » (via le contenu, les modalités, les tâches et les interactions engagées) et de leur offrir, ainsi qu'aux enseignants, une ergonomie finement adaptée à leurs besoins. (Detey, 2020 : 31).

Développer une telle technologie dépasserait nos moyens techniques et financiers. Sans nier l'énorme intérêt de ces recherches, nous avons abordé le problème dans un sens inverse : comment tirer le meilleur parti de technologies simples et déjà existantes pour proposer aux apprenants et

aux enseignants un contenu dont l'élaboration constitue le cœur de notre métier. C'est donc bien conscients de notre incapacité à « recréer » l'équivalent d'une interaction avec un enseignant et afin de créer un espace d'entraînement en complément des cours que nous avons organisé le contenu, conçu les tâches et les interactions sur le laboratoire de langue. Deux précisions sont nécessaires pour expliciter l'approche retenue.

- Guberina souligne que l'on peut influencer la perception soit au moment de l'émission, soit au moment de la transmission de l'énoncé :

Pour prononcer correctement une émission donnée (un son, un mot, un groupe rythmique ou une phrase), il faut l'avoir correctement entendue. Pour obtenir une réception (audition) correcte, on peut, d'après le Système Verbo-tonal, intervenir dans l'émission même (différentes possibilités de l'articulation nuancée, choix des mots, choix de différents types d'intonation, etc.) ou bien au cours de la transmission (filtres acoustiques) » (Guberina, 2003 : 181).

Les différentes générations de Suvag agissaient sur la transmission grâce à des filtres acoustiques. À l'inverse, notre recherche de l'optimale correctrice passe par la création d'énoncés choisis pour leurs qualités phonétiques (position des sons dans la chaîne parlée, intonation, présence/absence de phénomènes tels qu'enchaînements, liaison, etc.), donc au niveau de l'émission.

- Cette orientation implique que le contexte d'énonciation — central chez Guberina — n'est pas pris en compte dans l'état actuel de développement de la plateforme : nous travaillons sur des énoncés isolés, choisis pour leurs caractéristiques acoustiques et prosodiques, permettant aux apprenants d'écouter, répéter et manipuler des unités calibrées pour l'entraînement perceptif.

2. Concevoir des optimales correctives pour la formation asynchrone

2.1. Produire un étayage perceptif

Comme dit précédemment, à l'heure actuelle et à notre connaissance, aucun système informatique n'est capable de se substituer à un enseignant

pour évaluer la prononciation d'un apprenant et proposer un retour adéquat. La création d'exercices réalisables en autonomie nécessite donc de proposer aux apprenants des conditions optimales, qui leur permettent de produire des énoncés « corrects » dans la langue cible.

Afin de proposer de tels exercices, nous avons listé les « erreurs » attendues et les avons réparties en deux catégories : prosodie et sons de parole. Pour chacune de ces « erreurs », nous avons créé des exercices de répétition, permettant aux apprenants d'enregistrer, écouter et comparer leur production au modèle.

La validation (ou non) de la production de l'apprenant et donc de sa maîtrise des difficultés traitées se fait en deux étapes. Nous lui proposons dans un premier temps de s'autoévaluer en comparant sa production au modèle, autant de fois que souhaité. Ici se pose évidemment la question de la surdité phonologique (Troubetzkoy, 1939) et accentuelle (Dupoux, Peperkamp, 2002) de l'apprenant. Comment un apprenant pourrait-il évaluer sa production d'un enchaînement vocalique si ce phénomène n'existe pas dans sa langue maternelle et qu'il ne le perçoit pas ? Comment comparer sa production d'un [e] si on le perçoit comme un /i/ ?

Une étape de validation perceptive intervient alors sous la forme d'un questionnaire visant à déterminer si l'apprenant perçoit effectivement le phénomène travaillé. Si l'apprenant parvient à percevoir la différence entre un énoncé contenant un enchaînement consonantique et un énoncé n'en contenant pas, nous pensons qu'il sera plus facilement en capacité de percevoir cette différence entre sa production et le modèle présenté. S'il ne la perçoit pas, il constatera que sa réponse au questionnaire est fausse et disposera d'une série supplémentaire pour mettre le doigt sur ce qui est travaillé. Si après trois séries le phénomène travaillé n'est toujours pas maîtrisé, l'apprenant passe à la suite et ce point sera retravaillé plus tard. Ce travail par séries nous permet de gérer l'hétérogénéité entre les apprenants. Trois types de questionnaire permettent de couvrir l'ensemble des difficultés sur lesquelles nous souhaitons apporter un étayage :

- Questionnaires à choix multiples (QCM)

Les QCM portent sur de l'audio. L'apprenant entend un énoncé et doit choisir la bonne réponse à la question qui est écrite. La réponse peut être

de type textuel ou numérique (nombre de syllabes, correct/incorrect) ou audio (logatome, battement, audios à comparer au modèle).

- Localiser le son

L'apprenant entend un énoncé et les syllabes correspondant à cet énoncé apparaissent à l'écran. On demande à l'apprenant de cliquer sur la/les syllabe(s) dans la/lesquelle(s) se trouve le son/la syllabe inscrit(e) en haut de page. Pour ce type d'exercice, nous avons recours à l'API (les symboles API sont cliquables et renvoient au son correspondant).

- Identifier le son

L'apprenant entend un énoncé et les syllabes correspondant à cet énoncé apparaissent à l'écran. On demande à l'apprenant de choisir entre plusieurs sons celui qu'il entend. La syllabe sur laquelle porte la question est mise en évidence. Pour ce type d'exercice, nous avons recours à l'API.

La validation finale revient à l'enseignant accompagnant l'apprenant-utilisateur : les audios produits par l'apprenant sont consultables par l'enseignant référent pour validation.

Afin de favoriser l'automatisation en oral spontané, des productions semi-guidées sont proposées après une phase d'entraînement sur un point donné. Il s'agit pour l'apprenant, à partir d'une question et d'un stimulus visuel, de produire un énoncé. Ainsi, pour travailler l'enchaînement vocalique, l'apprenant est invité à répondre à une question du type « Il/elle va où ? », associée à une image représentant un lieu (par exemple une banque ou un marché). La production attendue est alors une phrase telle que « Il/elle va à la banque ». L'apprenant enregistre sa production et peut la comparer avec un modèle de contrôle. Ces données étant stockées pour étude, cette tâche pourra également nous permettre de mesurer l'efficacité de l'entraînement reçu par l'apprenant sur sa production spontanée. Nous espérons en effet que ce travail favorisera un transfert vers l'oral spontané, un effet déjà documenté pour les voyelles dans une étude de Lengeris (2018).

Le travail sur la prosodie est un préalable à tout travail sur les sons de parole. L'enjeu premier est de faire percevoir certains « traits saillants » de

la prosodie du français, notamment la relative régularité en termes de durée et d'intensité des syllabes non accentuées, la durée supérieure de la syllabe accentuée (généralement concomitante avec une variation de hauteur en « glissando »), la syllabation prioritairement ouverte, les phénomènes d'enchaînements consonantiques et vocaliques. La maîtrise de ces éléments entraîne une amélioration de l'intelligibilité et de la compréhensibilité des apprenants (Hahn, 2004, Kang *et al.*, 2010), de même qu'elle a un impact sur leur compréhension de l'oral.

Sans possibilité d'interaction avec l'apprenant, nous avons imaginé des stratégies lui permettant d'acquérir progressivement le système prosodique du français, impliquant un choix quant à l'ordre dans lequel les éléments prosodiques sont travaillés. L'apprenant commence par l'appréhension de la syllabe et ce quel que soit son niveau général de français ou sa langue maternelle. Nous proposons à la répétition des items audios permettant d'orienter favorablement la perception des apprenants. Nous nous sommes pour cela inspirés de notre expérience d'enseignement dans les ateliers évoqués plus haut. Le principe général est le suivant : il ne s'agit pas de présenter d'emblée des énoncés produits avec une prosodie « naturelle », mais d'opérer des modifications sur la prosodie visant à faire percevoir un élément donné en modifiant le débit, en ajoutant des pauses, en exagérant la durée d'une syllabe accentuée, la hauteur d'une intonation... avant de revenir à un modèle plus proche d'une prosodie naturelle. Par ailleurs, là encore inspirés par notre pratique, nous proposons pour les exercices visant l'acquisition de la syllabation des audios accompagnés d'un « *tapping* » régulier que l'apprenant doit reproduire, lui permettant ainsi d'intégrer les notions de syllabe et d'enchaînements (Astésano, 2022). On retrouve ici la recherche d'optimale perceptive, appliquée à la prosodie au niveau de l'émission, avec recours possible aux mouvements du corps (*tapping*).

L'intégration de gestes dans les activités prosodiques trouve également un appui dans les travaux récents sur l'entraînement incarné de la prosodie. Baills *et al.* (2022) montrent qu'un entraînement fondé sur l'imitation de logatomes accompagnés de gestes reproduisant les mouvements de hauteur améliore significativement l'accent et la précision suprasegmentale, là où la répétition seule produit des effets moindres. De manière complémentaire, Li *et al.* (2023) mettent en évidence que des

gestes soulignant les proéminences rythmiques et mélodiques favorisent l'amélioration de la précision segmentale. Ces résultats convergent avec les principes verbo-tonaux en soulignant l'intérêt du geste comme support perceptif et moteur dans l'apprentissage phonétique.

Les exercices de correction des sons de parole s'inspirent quant à eux des principes de MVT mis au point par Guberina, repris par la suite et développés par les verbo-tonalistes à savoir :

- repérage des erreurs ;
- diagnostic ;
- manipulation du retour (*feedback*) ;
- validation.

Prenons par exemple une erreur fréquente : un /y/ prononcé [u]. Le diagnostic que nous en faisons est que l'apprenant valorise les fréquences basses, on parlera en MVT de timbre sombre (Renard, 1979). L'énoncé est donc modifié de manière à proposer une optimale corrective (Intravaia, 2005).

Pour ce faire, la MVT offre plusieurs possibilités (Billières *et al.*, 2013) :

- placer le son en sommet intonatif (par exemple « Tu as lu ? ») ;
- le mettre au contact d'une consonne éclaircissante (par exemple le [s]) ;
- avoir recours à la prononciation nuancée : produire un allophone de /y/ tendant vers [i].

Nous n'avons pas retenu l'option de la prononciation nuancée pour la correction des sons, dans la mesure où le degré de modification requis dépend étroitement de l'interaction en temps réel entre l'enseignant et l'apprenant. Dans un contexte asynchrone, ce paramètre nous a paru trop difficile à contrôler de manière fiable. Ainsi, pour chaque type d'erreur nous avons créé des exercices composés de phrases discrètes dans lesquelles le son est en position optimale par rapport à une perception erronée donnée. Afin que l'apprenant puisse progresser, les sons sont dans un premier temps placés en situation très favorable, puis de moins en moins favorable et enfin défavorable.

2.2. L'interface comme support multimodal à la perception

Dans les écrits de Guberina, la perception est présentée comme multicanal « la langue parlée est surtout perçue par l'œil et par l'oreille » (Guberina, 2003 : 143) : dans la conception du matériel pédagogique SGAV, l'image fournit (au moins dans un premier temps) une indication de contexte du dialogue entendu et soutient grandement l'accès au sens. En l'état actuel de son développement, la plateforme fonetix.org ne propose que de manière assez exceptionnelle des déclencheurs visuels liés au sens (images). En revanche, nous avons fait en sorte que l'interface facilite la perception et la compréhension des phénomènes travaillés. Dans ce but, nous nous éloignons parfois de la doxa verbo-tonaliste, mais l'expérience de terrain, soutenue par les auteurs que nous citons par la suite, nous encourage à tester des solutions alternatives dans le but de fournir le meilleur étayage possible dans un contexte asynchrone. Nous estimons que le travail de la prononciation doit être accessible à tout apprenant de FLE, même (et surtout !) aux débutants. Nous avons donc souhaité que l'interface de notre plateforme soit sobre, simple d'utilisation et intuitive. Nous avons également essayé de limiter le recours à l'écrit.

Les apprenants effectuent les différentes actions nécessaires au moyen de gros boutons contenant des icônes : un haut-parleur pour écouter, un microphone pour répéter et enregistrer sa production, un phylactère pour lire les transcriptions de l'énoncé (orthographique et en Alphabet Phonétique International-API), un œil pour faire apparaître un document déclencheur. Si l'apprenant ne saisit pas la consigne, il peut cliquer sur un point d'interrogation donnant accès à une vidéo explicative précisant l'action attendue dans l'exercice.

Afin de limiter les effets négatifs de l'orthographe sur la prononciation des apprenants (Santiago, 2018), la transcription n'est accessible qu'après un premier enregistrement réalisé uniquement sur la base de la perception de l'énoncé. Par la suite, les utilisateurs ne peuvent enregistrer leur production en présence de la transcription : ils doivent la fermer pour accéder à la fonction « enregistrer ». Cette contrainte répond au principe verbo-tonal selon lequel l'écrit, en mobilisant des correspondances graphie-phonie souvent inadéquates, peut perturber la perception et contrarier l'étayage perceptif nécessaire à l'installation des formes sonores.

Pour le travail sur les sons de parole, nous avons choisi d'avoir recours à l'API. En effet, il nous fallait pouvoir faire correspondre de manière biunivoque un symbole écrit et un son, pour les questionnaires notamment. Les liens graphie-phonie du français étant particulièrement irréguliers, l'utilisation du seul code écrit ne nous semblait pas pertinente. Nous avons pu constater, dans le cadre de formations universitaires de FLE où la maîtrise de l'API constitue un objectif explicite, que l'usage de cette notation pouvait favoriser l'appropriation du système phonologique dès lors que l'apprenant commence à percevoir et à produire le son de manière plus fiable. Les travaux de Hamada (2018) montrent d'ailleurs que le *IPA-shadowing*² peut avoir un effet mesurable sur l'amélioration des segments, les apprenants du groupe API ayant significativement progressé sur les traits segmentaux au terme de l'entraînement.

Si notre dispositif ne recourt pas au *IPA-shadowing*, l'API est néanmoins fournie immédiatement après la phase de *shadowing* (Vale Narciso, 2022) afin d'offrir un appui visuel susceptible de renforcer la stabilisation des contrastes perçus. Par ailleurs, la notation en API demeure largement présente dans les dictionnaires et ressources en ligne ; lorsqu'elle est maîtrisée, elle peut constituer pour les apprenants un outil complémentaire de compréhension et de vérification phonologique.

Afin que les utilisateurs puissent faire le lien entre un symbole API et un son de parole, deux tableaux (voyelles et consonnes) sont accessibles dans la partie ressources. En cliquant sur un symbole, l'apprenant entend le son isolé et en contexte.

Enfin, la notation en API permet de rendre compte de la syllabation du français, ce que ne permet pas vraiment la présence de « lettres muettes » dans une transcription orthographique : « tu-par-les a-ve-cu-ne a-mie » nous semble bien moins compréhensible que [ty-paʁ-la-vɛ-ky-na-mi]. Ainsi, à côté de la transcription orthographique, nous proposons ce type de notation API avec découpage syllabique.

La conscience du nombre de syllabes étant un élément central en français (Wioland, 2005), nous avons choisi de matérialiser visuellement ces

² IPA-shadowing (Répétition immédiate effectuée à partir d'une transcription en alphabet phonétique international).

unités afin d'en faciliter le repérage et le comptage. Nous nous sommes ici inspirés des travaux montrant l'intérêt pédagogique des visualisations prosodiques, notamment en intonation (Hardison, 2007 ; Liu, Tseng, 2019). Dans cette perspective, nous avons intégré dans la plateforme un ensemble d'indices visuels complémentaires : segmentation syllabique, mise en évidence des syllabes accentuées — représentées par une barre plus longue, conformément au rôle prépondérant de la durée dans l'accentuation en français — ainsi que l'affichage de l'onde sonore et de la courbe d'intonation.

L'affichage de l'onde sonore et de la courbe d'intonation répond à des objectifs complémentaires. L'onde sonore constitue d'abord un appui visuel utile pour le travail segmental : elle permet d'observer la présence ou l'absence de vibration des cordes vocales dans les consonnes voisées et non voisées, ainsi que la continuité ou la rupture du signal lors de la réalisation des enchaînements vocaliques. La courbe d'intonation offre pour sa part un accès direct aux variations mélodiques : elle rend visibles les montées et les descentes de la voix, et permet d'identifier, sur les syllabes accentuées, les *glissandi* caractéristiques du mouvement intonatif. Ces *glissandi*, associés à un allongement vocalique, peuvent parfois être interprétés à tort comme un accent d'intensité ; la représentation graphique contribue alors à clarifier cette distinction. L'ensemble de ces indices fournit un appui perceptif supplémentaire pour comprendre et stabiliser les phénomènes prosodiques travaillés.

Pour déterminer quelles syllabes devaient être signalées comme accentuées et quels groupes portaient l'accent, nous nous sommes référés au mot phonétique (Wioland, 2005) ou mot prosodique (Astésano, Bertrand, 2016), niveau intermédiaire entre la syllabe et la séquence prosodique (Martin, 2024) particulièrement pertinent en didactique de la prononciation. La délimitation de ces unités dépend toutefois de divers paramètres (nombre de syllabes, débit, style énonciatif). Conscients de cette variabilité, nous avons retenu une solution pragmatique : la syllabe finale d'un mot phonétique consensuel entre concepteurs est indiquée, privilégiant ainsi l'étayage visuel lorsque l'apprenant produit l'énoncé.

Cette multimodalité et la taille réduite des énoncés sont cohérents avec les résultats récents montrant que les visualisations acoustiques (onde, F0)

présentées dans des phrases discrètes peuvent favoriser l'acquisition de propriétés prosodiques en L2 (Liu, Tseng 2019). Elle vise à fournir aux apprenants un environnement riche permettant d'appuyer simultanément la perception des phénomènes segmentaux et suprasegmentaux.

Suite à une première phase de test de notre plateforme auprès d'apprenants étrangers, nous avons inclus des capsules présentant explicitement les points abordés au moyen d'animations, de manière que les utilisateurs comprennent le but recherché par la série d'exercices qu'ils s'apprêtent à réaliser.

Outre les retours des utilisateurs, nous avons été encouragés à déroger au principe MVT d'acquisition implicite du fait des effets bénéfiques que peut avoir un enseignement explicite des paramètres suprasegmentaux sur la compréhensibilité et sur les compétences prosodiques des apprenants, quel que soit leur niveau (Saito et Saito, 2017), notamment lorsque leurs habitudes sont « fossilisées » (Derwing *et al.*, 2014).

3. Créer des parcours adaptés aux cribles phonologiques et prosodiques des apprenants

3.1. Vers des optimales correctives différenciées selon la langue première

Même s'il existe des variations interpersonnelles, les productions orales des locuteurs d'une communauté linguistique donnée vont présenter les mêmes écarts par rapport à la norme du français, tant au niveau segmental que suprasegmental. Ce « système d'erreurs » (Billières, 2016), est ce qui permet souvent à un natif d'identifier l'origine linguistique d'un non-natif. Ainsi, les anglophones auront tendance à réaliser une syllabation fermée contrairement aux francophones qui privilégient une syllabation ouverte, /e/ sera souvent perçu et produit [i] par un arabophone mais [ɛ] par un polonais, etc.

L'établissement exhaustif de ces « erreurs » repose à la fois sur une analyse contrastive entre la langue cible — en l'occurrence le français — et la langue de l'apprenant, sur notre expérience de terrain, ainsi que sur les

apports de la littérature scientifique consacrée aux difficultés phonétiques propres aux différentes langues premières.

Cette démarche est à l'origine de la création de parcours pour hispanophones, lusophones, sinophones... La notion de parcours est fondamentale à deux titres : elle renvoie d'une part à un cheminement propre à chaque groupe linguistique, et d'autre part à une idée de progression construite à partir des difficultés anticipées pour chaque langue première. Dans cette logique, l'organisation interne de chaque parcours répond aux principes déjà évoqués : la priorité est donnée aux éléments suprasegmentaux, afin de créer un contexte favorable à la perception puis à la production des segments. Les premières activités sont donc consacrées à la prosodie, avant d'aborder le travail sur les sons proprement dits : les voyelles dans un premier temps, puis les consonnes.

Nous avons à l'heure actuelle implémenté des parcours pour les locuteurs de quatre langues : anglais, arabe, chinois (mandarin), espagnol. Sont en préparation des parcours à destination des locuteurs du japonais, de l'italien, du portugais, du russe et du vietnamien. Si nous comptons compléter ce catalogue, il serait illusoire de proposer des parcours pour l'ensemble des langues du monde. Par ailleurs, il nous a semblé pertinent de proposer une solution pour des structures accueillant un public international et désireuses de proposer un soutien numérique dans l'acquisition de la prononciation du français, et ce quelle que soit la langue maternelle de l'apprenant.

3.2. Une boîte à outils au service du diagnostic enseignant

Les parcours créés en fonction de la LM de l'apprenant reposent sur un diagnostic *a priori* des écarts d'un locuteur de cette langue par rapport à la norme du français. Nous avons insisté sur le fait qu'un tel diagnostic ne peut actuellement pas être réalisé automatiquement par notre plateforme sur la base des productions d'un utilisateur. Pour des classes composées d'apprenants locuteurs de langues pour lesquelles nous n'avons pas créé de parcours, c'est donc à l'enseignant accompagnateur d'évaluer la production de chaque apprenant et d'imaginer le parcours apportant le meilleur étayage possible, en complément d'activités menées en classe. Nous avons donc conçu une « boîte à outils », proposant des exercices que l'enseignant

attribue à un apprenant donné, créant ainsi un parcours individuel via notre interface. Sur la base de notre expérience de classe, nous avons listé l'ensemble des difficultés envisageables tant sur les éléments segmentaux que supra-segmentaux et conçu des exercices suivant les principes décrits précédemment. Ici encore, le travail sur la prosodie est obligatoirement un préalable au travail sur les voyelles et les consonnes.

Pour les sons de parole, l'interface est conçue de la manière suivante : l'ensemble des sons du français sont présentés en notation API, l'enseignant clique sur un son réalisé de manière erronée par un apprenant. Apparaissent alors différentes notations API décrivant les manières dont ce son peut être produit par un apprenant ne le maîtrisant pas. Par exemple /y/ peut être produit ou tendre vers [u], [i] ou [ju]. En choisissant la réalisation du son qu'il a perçue chez un apprenant, l'enseignant attribue une série d'exercices dans laquelle le son est en position optimale pour cette difficulté précise.

Afin de soutenir l'enseignant dans la construction de parcours individualisés et de favoriser une appropriation progressive des principes mobilisés, nous avons développé un ensemble de ressources complémentaires. Ces matériaux comprennent des fiches pédagogiques détaillant les principaux phénomènes prosodiques et segmentaux ainsi que les erreurs typiques observées selon les langues premières, des capsules vidéo présentant les caractéristiques prosodiques et les sons du français, et des tutoriels explicitant l'usage de la boîte à outils. L'objectif est de fournir un étayage méthodologique destiné aux enseignants, parallèlement à celui proposé aux apprenants : ces ressources permettent ainsi de consolider les compétences diagnostiques des enseignants, qu'ils soient déjà formés ou non à la remédiation phonétique, et s'inscrivent dans une démarche collective fondée sur la mutualisation des pratiques au sein de la communauté Fonetix.

Conclusion

La réalisation de cette plateforme est le fruit d'un travail collectif mené sur plusieurs années et nourri par une réflexion continue autour de la Méthode verbo-tonale et de la notion d'optimale corrective. L'outil est en

ligne depuis 2023 et son adoption croissante par des structures universitaires, des organismes de formation et des enseignants de français langue étrangère ou de français langue d'intégration (FLE/FLI) atteste de l'intérêt porté à une approche articulant entraînement autonome, étayage perceptif et accompagnement pédagogique. Conscients du caractère perfectible de notre proposition, nous avons conçu une plateforme volontairement agile : les contenus, l'organisation des parcours et l'interface peuvent être ajustés, enrichis ou transformés à la lumière des retours émanant tant des apprenants que des enseignants. C'est d'ailleurs dans ce cadre que nous avons récemment intégré la visualisation de la courbe intonative, désormais appliquée au modèle et à la production de l'apprenant.

Si notre démarche est avant tout celle de praticiens, elle s'inscrit dans un dialogue constant avec la recherche. Les données massives générées par la plateforme — productions audio, trajectoires d'apprentissage, résultats aux questionnaires perceptifs — constituent une ressource inédite pour étudier l'efficacité des différents paramètres de l'optimale corrective (position prosodique, entourage segmental, statut accentuel). Une première collaboration avec le laboratoire de neuropsycholinguistique de l'université Toulouse 2 vise à analyser ces données et à déterminer dans quelle mesure certaines positions optimales (prosodie, entourage, position dans l'énoncé) favorisent l'amélioration des productions segmentales.

Par ailleurs, nous avons amorcé l'intégration de modules d'intelligence artificielle — notamment de reconnaissance vocale — ouvrant la voie à des développements futurs portant sur la visualisation de la prosodie ou l'évaluation automatique de certains traits segmentaux et phénomènes suprasegmentaux. À terme, il serait souhaitable de concevoir des tâches d'entraînement contextualisées permettant de mieux prendre en compte le rôle du discours et du dialogue dans l'acquisition phonétique et son transfert en oral spontané.

Enfin, si la plateforme se veut un soutien à l'autonomie des apprenants, elle n'a nullement vocation à se substituer à l'enseignant. Nous souhaitons au contraire qu'elle devienne un outil à sa disposition, valorisant son expertise diagnostique et facilitant la mise en place d'activités phonétiques en classe. Dans cette perspective, nous espérons que ce dispositif

contribuera également à renforcer la formation des enseignants de FLE/FLI à la correction phonétique, domaine encore largement marginalisé dans les cursus.

Bibliographie

Astésano, C. 2022. « De la supramodalité du rythme : implications pour la description prosodique, la remédiation linguistique et l'apprentissage des langues ». *XXXIV^e Journées d'Études sur la Parole*, p. 1-14. [En ligne] : doi.org/10.21437/JEP.2022-1 [consulté le 30 novembre 2025].

Astésano, C., Bertrand, R. 2016. « Accentuation et niveaux de constituance en français : enjeux phonologiques et psycholinguistiques ». *Langue française*, 191 (3), p. 11-30. [En ligne] : doi.org/10.3917/lf.191.0011 [consulté le 30 novembre 2025]

Billières, M. 2016. « Prédire les erreurs phonétiques des apprenants en FLE ». *Au son du FLE*. [En ligne] : <https://www.verbotonale-phonetique.com/predire-erreurs-phonetiques-apprenants-fle/> [consulté le 30 novembre 2025].

Billières, M., Alazard, C., Astésano, C., Nocaudie, O. 2013. « La correction des éléments segmentaux : voyelles et consonnes ». In : *Phonétique corrective en FLE – Méthode verbo- tonale*. Toulouse : MVT- UOH. [En ligne] : <https://mvt-uoh.univ-tlse2.fr/seq05P0101.html> [consulté le 30 novembre 2025].

Billières, M. *et al.* 2018. *Pratiques de l'enseignement de la prononciation en FLE*. FUN- MOOC. [En ligne] : <https://www.fun-mooc.fr/fr/cours/pratiques-de-lenseignement-de-la-prononciation-en-fle/> [consulté le 30 novembre 2025].

Derwing, T.M., Munro, M.J., Foote, J.A., Waugh, E., Fleming, J. 2014. « Opening the window on comprehensible pronunciation after 19 years: a workplace training study ». *Language Learning*, 64 (3), p. 526-548. doi.org/10.1111/lang.12053 [consulté le 30 novembre 2025].

Detey, S. 2020. « Que signifie « corriger la prononciation du français » aujourd'hui ? De 80 ans de crible phonologique aux problématiques contemporaines ». *Le langage et l'homme*, 55 (2), p. 15-42.

Dupoux, E., Peperkamp, S. 2002. Fossil markers of language development: phonological « deafnesses » in adult speech processing. In: Durand, J., Laks, B., éd., *Phonetics, phonology and cognition*, p. 168-190. Oxford: Oxford University Press.

Guberina, P. 2003. *Rétrospection*. Roberge, C. (Rédacteur en chef et éditeur), Zagreb: Artresor Naklada.

Hahn, L.D. 2004. « Primary stress and intelligibility: research to motivate the teaching of suprasegmentals ». *TESOL Quarterly*, 38 (2), p. 201-223. [En ligne] : doi.org/10.2307/3588378 [consulté le 30 novembre 2025].

Hamada, Y. 2018. « Shadowing for pronunciation development: haptic-shadowing and IPA-shadowing ». *Journal of AsiaTEFL*, 15 (1), p. 167-183. [En ligne] : doi.org/10.18823/asiatefl.2018.15.1.11.167 [consulté le 30 novembre 2025].

- Hardison, D.M. 2007. The visual element in phonological perception and learning. In: Pennington, M.C., éd., *Phonology in context*, Londres: Palgrave Macmillan, p. 135-158.
- Intravaia, P. 2005. « La formation verbo-tonale des professeurs de langue ». In : *Apprentissage d'une langue étrangère / seconde*, vol. 4, p. 139-169. Bruxelles : De Boeck Supérieur. [En ligne] : doi.org/10.3917/dbu.barbe.2005.01.0139 [consulté le 30 novembre 2025].
- Kang, O., Rubin, D., Pickering, L. 2010. « Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English ». *The Modern Language Journal*, 94 (4), p. 554-566. [En ligne] : doi.org/10.1111/j.1540-4781.2010.01091.x [consulté le 30 novembre 2025].
- Lengeris, A. 2018. « Computer-based auditory training improves second-language vowel production in spontaneous speech ». *Journal of the Acoustical Society of America*, 144 (3), p. EL165-EL171. [En ligne] : doi.org/10.1121/1.5052201 [consulté le 30 novembre 2025].
- Liu, Y.-T., Tseng, W.-T. 2019. « Optimal implementation settings for computerized visualization cues in assisting L2 intonation production ». *System*, 87, article 102145. [En ligne] : doi.org/10.1016/j.system.2019.102145 [consulté le 30 novembre 2025].
- Martin, P. 2024. *Une nouvelle grammaire de l'intonation de la phrase française*. [En ligne] : https://inria.hal.science/hal-04623104 [consulté le 30 novembre 2025].
- Munro, M.J., Derwing, T.M. 1995. « Foreign accent, comprehensibility and intelligibility in the speech of second-language learners ». *Language Learning*, 45 (1), p. 73-97. [En ligne] : doi.org/10.1111/j.1467-1770.1995.tb00963.x [consulté le 30 novembre 2025].
- Renard, R. 1979. *La méthode verbo-tonale de correction phonétique*. 3^e éd. Paris : Didier.
- Roberge, C. 2003. « Préface ». In : Petar Guberina. *Rétrospection*. Zagreb : Artresor Naklada.
- Saito, Y., Saito, K. 2017. « Differential effects of instruction on the development of second-language comprehensibility, word stress, rhythm and intonation ». *Language Teaching Research*, 21(5), p.589-608. [En ligne] : doi.org/10.1177/1362168816643111 [consulté le 30 novembre 2025].
- Santiago, F. 2018. « Effets de l'orthographe dans la prononciation du français L2 ». *Journées d'Études sur la Parole*. [En ligne] : https://hal.science/hal-01768620 [consulté le 30 novembre 2025].
- Sauvage, J. 2019. « Humanité numérique et formation innovante en didactique du FLE ». *Recherches en didactique des langues et cultures*, 16 (2). [En ligne] : doi.org/10.4000/rdlc.7051 [consulté le 30 novembre 2025].
- Troubetzkoy, N.S. 1939. *Principes de phonologie*. Paris : Klincksieck.
- Valle Narciso, J. M. 2022. « L'application de la variation de la vitesse dans la technique de la répétition décalée du discours pour l'amélioration de l'intonation et de la fluidité en

français ». *Synergies Espagne*, n° 15, p. 41- 57. [En ligne] : <https://gerflint.fr/Base/Espagne15/valle.pdf> [consulté le 30 novembre 2025].

Wioland, F. 2005. *La vie sociale des sons du français*. Paris : L'Harmattan.



© *Synergies Europe*, Revue du GERFLINT, n° 20, Année 2025.
ARK : <https://catalogue.bnf.fr/ark:/12148/cb42696564q>
Bibliothèque nationale de France - Décembre 2025.

Éléments sous droits d'auteur.
www.gerflint.fr
<https://gerflint.fr/synergies-europe>

