



**HAL**  
open science

# Méthodologie de la segmentation pour l'évaluation objective des compétences phonologiques avec **DIADOLAB3**

Anne Menin-Sicard, Etienne Sicard, Caroline Grenier, Tiphaine Bardelang,  
Sandrine Michel

► **To cite this version:**

Anne Menin-Sicard, Etienne Sicard, Caroline Grenier, Tiphaine Bardelang, Sandrine Michel.  
Méthodologie de la segmentation pour l'évaluation objective des compétences phonologiques avec  
DIADOLAB3. 2021. hal-03149560

**HAL Id: hal-03149560**

**<https://hal.science/hal-03149560>**

Preprint submitted on 23 Feb 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Méthodologie de la segmentation pour l'évaluation objective des compétences phonologiques avec DIADOLAB3

Anne Menin-Sicard<sup>1</sup>, Etienne Sicard<sup>1,2</sup>, Sandrine Michel<sup>1</sup>, Caroline Grenier<sup>1</sup>, Tiphaine Bardelang<sup>1</sup>

(1) Laboratoire LURCO/UNADREO, équipe de recherche ERU 46

(2) INSA-GEI, 135 Avenue de Ranguel, 31077 Toulouse

**Résumé :** Dans cet article, nous détaillons la procédure de segmentation de phrases courtes afin d'évaluer le score d'articulation, dans le cadre d'une méthodologie objective de l'évaluation de l'intelligibilité lors du bilan orthophonique pour des patients atteints de troubles de la parole. Nous évoquons les biais de l'évaluation subjective, justifions le choix du protocole et la méthodologie d'enregistrement, puis de segmentation assistée en vue de comparer les phonèmes produits avec ceux attendus. L'approche par processus de parole est explicitée. Au travers de cas de parole altérée, nous illustrons diverses notions telles que les imprécisions, confusions et contaminations. Nous évoquons aussi les biais de la segmentation et les conséquences sur la fiabilité des scores d'articulation.

**Mots-clés :** parole, segmentation, orthophonie, bilan, altération, pratiques probantes, intelligibilité, articulation, phonologie, DIADOLAB

### Introduction

Dans le cadre du bilan du langage oral, l'évaluation perceptive auditive faite par le clinicien est actuellement la plus couramment utilisée [Bouchet 2008] [Chevrie-Muller 1997] [Coquet 2007] [Khomsy 2001] [Thibault 2003]. Avec la généralisation des pratiques probantes [Maillart 2014] [Sicard 2017], force est de constater que l'évaluation subjective ne suffit plus et que notre profession doit se doter d'outils d'évaluation objective pour comparer des résultats à partir d'une ligne de base définie et stable. Les mesures instrumentales permettent de réduire les suivis inutiles qui ne relèvent pas du trouble mais de l'immaturation simple, de surseoir aux pressions scolaires et sociales et de se concentrer sur des cas avérés de trouble ce qui pourrait de ce fait réduire les listes d'attente et contribuer à aider les cliniciens à prioriser les suivis.

Dans cet article, nous évoquons les biais de l'évaluation subjective, notamment celle portant sur les logatomes. En support à une méthodologie objective de l'évaluation de l'intelligibilité avec le logiciel DIADOLAB [Menin-Sicard 2019], nous détaillons dans cet article, l'outil d'aide à la segmentation, la procédure avec la cotation associée, ainsi que le choix du protocole. Après avoir donné quelques rappels concernant les batteries d'évaluation du langage oral et justifié l'intérêt d'une analyse de la parole au niveau du phonème, nous rappelons quelques recommandations concernant l'enregistrement et le pré-traitement de la parole. Nous présentons les outils principaux de la segmentation, parmi lesquels le spectrogramme, l'allure temporelle ainsi que les marqueurs spécifiques. Les pistes concernant la segmentation assistée, voire la segmentation automatisée sont aussi évoquées.

Nous abordons aussi les notions de coarticulation et de contamination puis l'approche par processus hiérarchisés. Nous discutons enfin de l'apport de 3 phrases types étudiées sur un corpus de 80 enfants.

## Evaluation de la phonologie

Traditionnellement, l'évaluation de la phonologie se réalise à partir d'une épreuve de répétition de logatomes ou de pseudo-mots ou encore de mots usuels. Un tableau recensant les batteries d'évaluation du langage oral les plus utilisées par les orthophonistes est proposé dans [Menin Sicard 2019]. La majorité des batteries d'évaluation du langage oral recommande d'enregistrer les épreuves de phonologie afin que l'orthophoniste puisse écouter à posteriori les enregistrements et analyser la production des patients dans diverses situations. En pratique, la réalité terrain est tout autre.

### Batteries d'évaluation du langage oral les plus utilisées

Sur 50 orthophonistes interrogées via un questionnaire pré-formation [AMS] en 2019, 50% d'entre elles utilisent la batterie Exalang [Thibault 2003] pour évaluer la phonologie, contre 44% qui utilisent la batterie EVALO [Coquet 2007]. Cette dernière comporte une liste de logatomes et une épreuve de répétition de mots usuels. Sur les 50 orthophonistes interrogées, seulement 28 % d'entre elles enregistrent et réécoutent les patients contre 54 % qui passent outre et procèdent à la cotation au fur et à mesure du déroulé de la passation par manque de temps. L'épreuve de répétition de logatomes est une épreuve encore aujourd'hui considérée comme incontournable.

Un certain nombre de questions se posent par rapport à l'utilisation des listes de logatomes.

- Biais de mémoire : pour les logatomes de plus de 3 syllabes, n'y a-t-il pas **deux paramètres évalués** en même temps ? en l'occurrence la phonologie et la mémoire auditivo-verbale. Ne devrait-on pas limiter les items des listes à des logatomes ou non-mots de deux ou trois syllabes maximum pour éviter ce biais d'empan auditif ?
- **Logatomes complexes** comparables à des virelangues comme par exemple /sinzanchujon/. Ces logatomes incluant 4 constrictives non maîtrisées à 5 ans peuvent-ils être raisonnablement proposés à des enfants de 4 ans ?
- Ces logatomes, sélectionnés dans les épreuves de phonologie, obéissent-ils aux règles phonotactiques de la langue française ?
- Le risque de **lexicalisation** doit-il être écarté ? si oui comment ? (ex /sozisson/ =>saucisson /panbi/=>Bambi).
- Des logatomes longs et comportant **des risques d'erreurs de décodage** grapho-phonémique comme /faviker/ comme c'est le cas lorsqu'ils sont générés informatiquement et de façon aléatoire peuvent-ils être conservés dans ces listes ?

La retranscription des logatomes à l'oreille seule ou après écoutes attentives de l'enregistrement sont aussi sujettes à discussion. Les altérations audibles avec un fort consensus à l'évaluation subjective concernent le plus souvent des altérations grossières de type inversion ou omissions. Le taux de consensus dans ce cas est certes assez élevé. En revanche, certaines substitutions, ébauches, imprécisions, et autres altérations subtiles sont souvent peu ou pas repérées à l'oreille seules, notamment celles qui concernent les voyelles (consensus faible). Cela signifierait que beaucoup d'altérations pourraient passer inaperçues et ne seraient pas comptabilisées dans le score final des évaluations phonologiques traditionnelles. C'est d'autant plus dommageable que cette « indulgence » liée à notre suppléance mentale concerne une liste de logatomes et donc se répète au cours de l'épreuve.

Le risque d'erreur d'évaluation est d'autant plus significatif que le nombre d'item administré est important. De ce fait, l'évaluation subjective avec ou sans enregistrement des productions d'un sujet est-elle fiable ?

D'autres questions viennent aussi à l'esprit. Dans quelle mesure des facteurs propres à l'évaluateur vont-ils influencer la perception auditive ? Nous pouvons citer des facteurs tels que :

- Sa perception auditive propre (courbe d'audition binaurale)
- Son éducation linguistique (langue maternelle et d'apprentissage, pauvre ou riche phonétiquement)
- Son exposition précoce à d'autres langues (patois ou dialectes)
- La participation de la lecture labiale et linguale lors de l'épreuve
- La présence ou non du sujet en face de l'évaluateur
- Ses contraintes temporelles
- Son état de fatigue et de concentration
- Le comportement du sujet évalué
- Son degré de concentration
- Sa fatigabilité au cours de l'épreuve
- La présence ou non des parents
- La familiarité/attachement à l'enfant

Beaucoup de biais difficiles à maîtriser contribuent à augmenter **significativement la subjectivité** et le phénomène appelé **attraction de familiarité**. Pour échapper autant que possible à ces multiples biais, il est nécessaire de travailler à minima sur enregistrements de bonne qualité. Ce cadre méthodologique n'exclue pas le biais de la perception de l'évaluateur certes mais réduit significativement la subjectivité et donc l'imprécision de l'évaluation. La question de l'expertise de l'orthophoniste mérite d'être posée. La segmentation varie-t-elle en fonction de l'expérience du clinicien ?

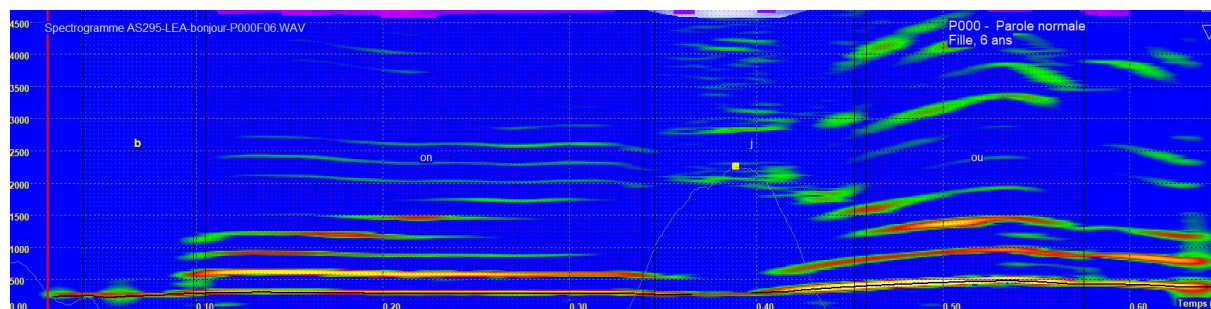
### Intérêt de la segmentation

La segmentation de la parole enregistrée permet d'isoler chaque phonème prononcé, préciser le phonème perçu, et par comparaison avec la liste des phonèmes attendus, dresser la liste des altérations. Cette approche permet d'avoir une vision précise des compétences phonologiques du patient. Souvent réservée aux cliniciens chercheurs, nous défendons l'idée que la segmentation de la parole se généralise à la pratique clinique, afin de réaliser des bilans objectifs de la parole altérée et pathologique de meilleure qualité.

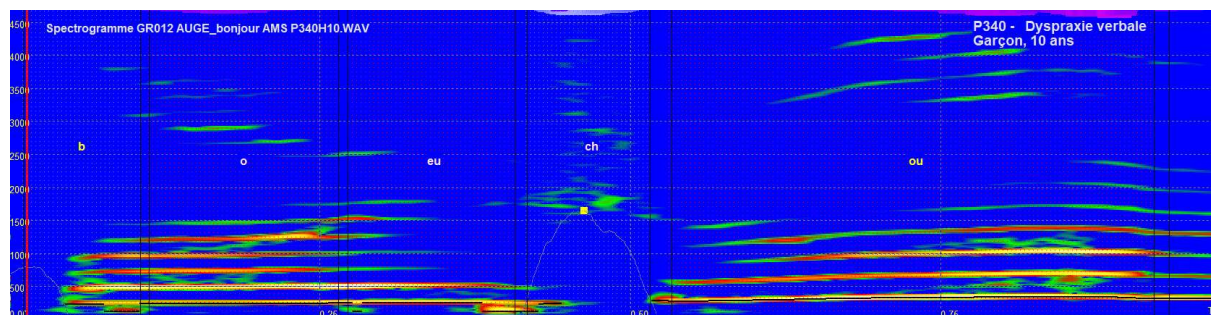
L'adhésion du plus grand nombre à la segmentation de la parole passe cependant par un certain nombre de prérequis :

- La phrase à segmenter doit être efficace, compacte, et identique d'un patient à un autre
- Le clinicien doit disposer d'outils dédiés, fiables et pratiques pour faciliter la segmentation : spectrogramme, allure temporel, marqueurs spécifiques, pré-alignement, etc.
- La liste des altérations doit être produite automatiquement, ainsi que le calcul d'un score d'articulation, pour éviter les calculs fastidieux et limiter les risques d'erreurs
- Développer les compétences en phonétiques acoustiques des cliniciens selon la méthodologie décrite plus bas.

Avec l'aide du spectrogramme et de l'allure temporelle, nous pouvons préciser les frontières de chaque phonème, et sur la base de l'écoute sélective répétée en cas de doute, lui associer le phonème perçu. Nous limitons l'échantillon de parole au strict minimum, en y intégrant un maximum de processus. Le choix de la phrase est justifié à la section 3 de cet article.



Séquence /Bonjour/ parole normale, fille de 6 ans



Séquence /Bonjour/ parole altérée, garçon de 10 ans

Figure 1 : segmentation du mot /Bonjour/, parole normale (en haut) et altérée (en bas)

Si la parole est peu altérée, cette étape est simple et rapide. Elle devient plus complexe et chronophage dans le cas de la parole très altérée ou pathologique. Deux exemples de segmentation du mot /Bonjour/ sont donnés figure 1. Celui du haut correspond à la parole normale, celui du bas à une parole altérée.

## Recommandations d'enregistrement et de mise en forme

### Matériel et configuration d'enregistrement

Il existe une offre considérable en termes de microphones et de cartes sons. Nous recommandons l'usage d'un microphone directionnel qui isole bien la voix du patient, associé à une carte son externe pour limiter les bruits parasites. Parmi les microphones directifs (super cardioïdes ou unidirectionnels) les plus utilisés se trouvent le T.Bone MB85 BETA ou Shure SM58. Concernant la connexion micro-ordinateur, les câbles XLR > USB tels que Alesis MicLink, T.Bone 1X soient très simples de mise en œuvre et bon marché, mais ils souffrent d'un niveau de bruit un peu élevé, perceptible lors de l'écoute. Le microphone T.bone SC 440 USB n'isole pas véritablement la voix du patient mais a une excellente dynamique. Nous utilisons en formation une carte son externe de type Asus Xonar X5.

Les options de traitement du signal telles que le contrôle automatique du gain ou la suppression du bruit doivent être désactivés. Dans ces conditions, les fichiers de parole sont exploitables et de bonne qualité. Le logiciel DIADOLAB [Sicard 2020] utilise une fréquence d'échantillonnage de 22050 Hz, 16 Bit, mono, une configuration très largement suffisante pour analyser les composantes principales de la parole.

Au niveau protocole, là encore, on observe une grande variabilité d'approches : le thérapeute ne donne pas toujours la même consigne, ne met pas toujours le patient en situation identique, ce qui peut altérer la pertinence des analyses et des comparaisons avant/après prise en charge.

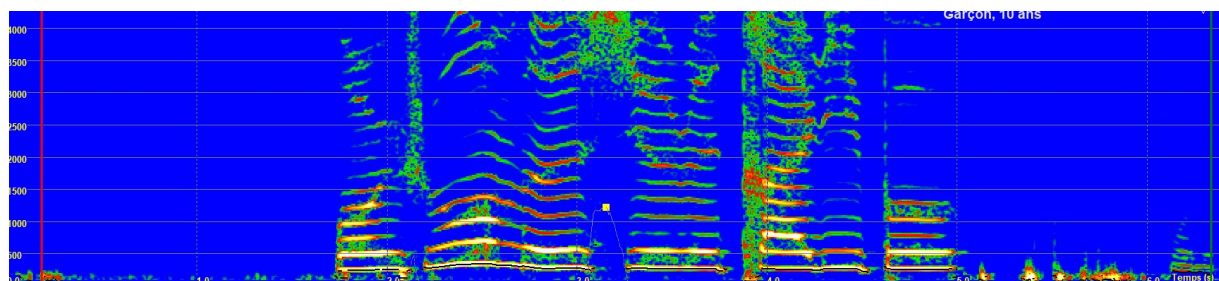
Nous recommandons de :

- Contrôler la distance bouche-micro à l'aide d'une règle solidement fixée sur le micro.
- Donner strictement la même consigne, avec le modèle adapté
- Faire un 1<sup>er</sup> essai d'enregistrement pour s'assurer que le patient a bien intégré la consigne et valider le niveau d'enregistrement.

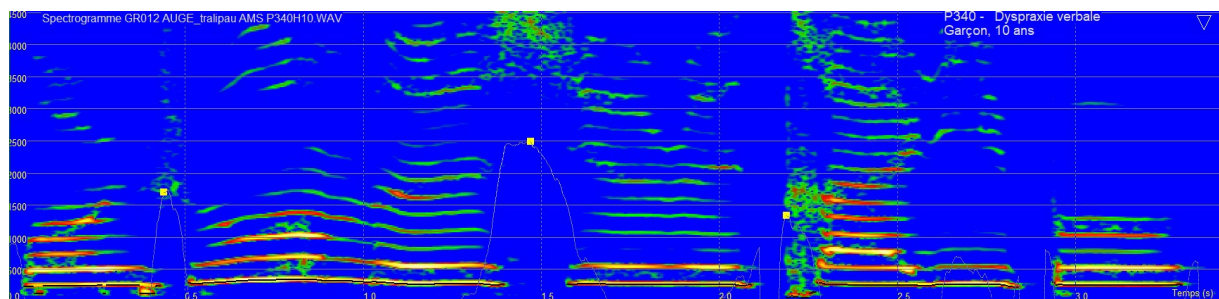
Nous proposons dans les phases d'enregistrement de la parole d'afficher à l'écran le protocole détaillé ainsi que les exemples de réalisations homme, femme, enfant [Menin-Sicard 2019]. L'interface d'enregistrement de la phrase /Bonjour Monsieur Tralipau/ est reportée figure 2.



Figure 2 : Protocole d'enregistrement de la phrase, avec accès aux enregistrements de phrases modèles Homme, Femme & Enfant dans DIADOLAB



Enregistrement initial



Suppression des portions inutiles

Figure 3 : Sélection de la partie utile de l'enregistrement

## Sélection de la partie utile

La partie à segmenter doit être préalablement débarrassée des parties silencieuses, inutiles ou encore comportant des bruits, consignes, commentaires inutiles afin de réaliser la segmentation dans les meilleures conditions possibles. L'exemple de la figure 3 concerne un enregistrement où la parole du patient n'apparaît qu'après 2 secondes environ. Le clinicien n'a arrêté l'enregistrement que 2 secondes après la fin de la phrase. Il convient ici de ne sélectionner que la partie utile de la phrase.

## Choix de la phrase à segmenter

### Approche par processus hiérarchisés

Les processus de parole tels que nous les envisageons sont hiérarchisés en fonction de leur impact sur l'intelligibilité. Les 4 premiers blocs en haut de la figure 4 (constriction, occlusion, voisement et nasalité) correspondent aux processus de base que le sujet doit pouvoir maîtriser afin de pouvoir les combiner aisément. Les blocs apparaissant dans une couleur foncée signifient qu'ils ont un impact important sur l'intelligibilité. Ce sont les extrêmes consonantiques et vocaliques et ceux qui ont une forte occurrence dans la langue française tels que le /s/ et le /r/ [Wioland 1991].

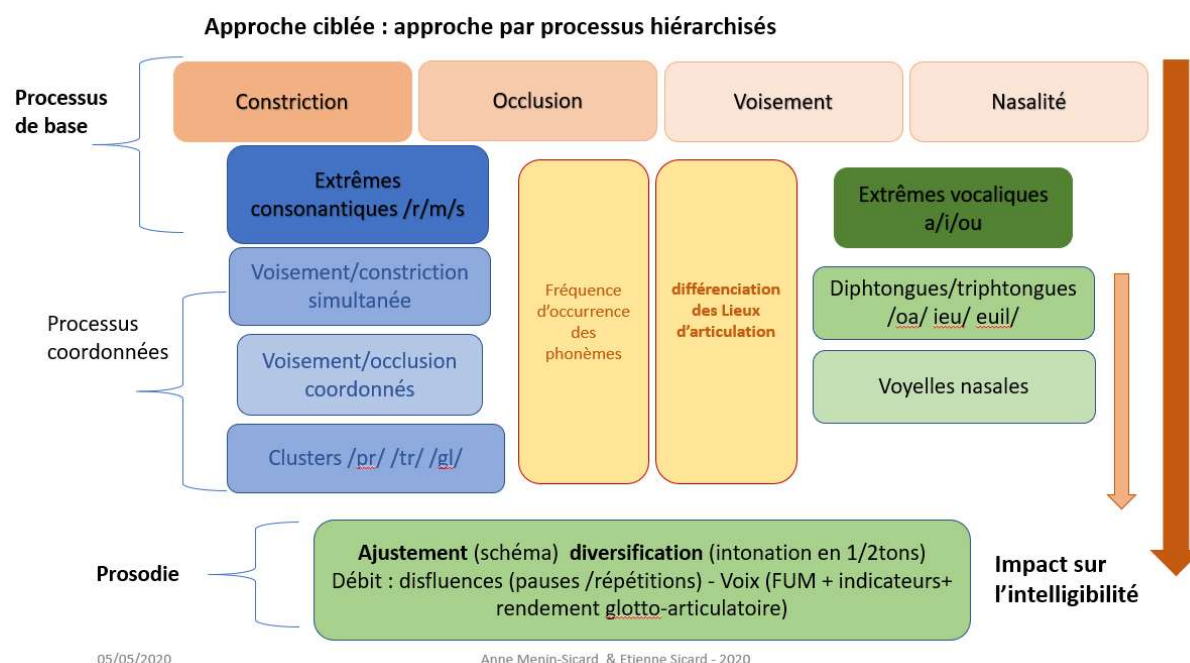


Figure 4 : Illustration de l'approche par processus de parole hiérarchisés

Les phonèmes légèrement plus longs (constrictives, diphtongues) ont un impact plus important sur l'intelligibilité que les phonèmes très courts sujets à contamination tels que le /l/ par exemple. La différenciation des lieux d'articulation est prise en compte mais n'est pas prioritaire. Les phonèmes postérieurs donc non visibles par la lecture labiale et linguale seront traités et travaillés en priorité par rapport à ceux qui sont visibles tels que le /p/ et le /b/.

Les paramètres prosodiques sont pris en considération : on recherchera l'amplitude et la diversification intonative, un bon rendement glotto-articulatoire qui soulagera le travail purement glottique ainsi qu'un débit en accord avec les valeurs référentielles de l'âge et du genre.

### Phrase de référence

Le choix de la phrase du protocole a donné lieu à une démarche d'analyse minutieuse selon des critères décrits dans [Sicard 2020], étayé par deux mémoires d'orthophonie, l'un de C. Grenier [Grenier 2020], l'autre de T. Bardelang [Bardelang 2020]. Nous avons retenu les critères permettant de considérer qu'une phrase est efficace dans le cadre du bilan orthophonique de la parole, c'est-à-dire susceptible de fournir le maximum d'information sur les capacités phonologiques du patient en un minimum de temps d'analyse :

- Elle est courte
- Elle comprend les extrêmes vocaliques
- Elle comprend les extrêmes consonantiques
- Elle couvre au mieux les processus de parole (Figure 5)
- Elle comprend les phonèmes qui ont le plus fort taux d'occurrence
- Elle est peu redondante
- Elle comporte des accents et contours prosodiques variés
- Sa phonotactique est proche de la langue française.

Plus la phrase est courte, plus les efforts d'analyse et d'interprétation seront réduits. Si l'on utilise des outils tels que PRAAT [Boersma 2019] ou DIADOLAB [Menin-Sicard 2019] pour la segmentation des mots, syllabes ou phonèmes, l'effort d'identification des frontières, de comparaison entre phonème perçu et phonème attendu ainsi que le codage de la confusion éventuelle pour chaque phonème sera d'autant moins coûteux en temps que le nombre d'items à analyser sera réduit. De plus, une phrase courte sollicitera peu les capacités de mémoire immédiate ou les compétences morphosyntaxiques du sujet ce qui élimine ces biais susceptibles de perturber la répétabilité de l'évaluation du score d'articulation [Sicard 2020].

PHONEME	PROCESSUS	COMMENTAIRE
<b>B</b>	Voisement puis occlusion	
<b>ON</b>	Voyelle nasale	Peu sensible à la régionalisation
<b>J</b>	Constriction et voisement	Processus simultanés
<b>OU</b>	Extrême vocalique postérieure	Accent tonique, prolongation, Extrême du triangle vocalique
<b>R</b>	Voisement constriction	Extrême consonantique
<b>M</b>	Voisement, nasalisation, occlusion antérieure extrême	Energie concentrée vers les basses fréquences
<b>EU</b>	Phonème neutre, central	
<b>S</b>	Constriction extrême	Energie concentrée vers les hautes fréquences
<b>IEU</b>	Diphthongue	
<b>TR</b>	Cluster	Souvent substitué en KR
<b>A</b>	Extrême vocalique bouche ouverte	Extrême du triangle vocalique
<b>L</b>	Voisement puis occlusion	Souvent substitué « ieu »
<b>I</b>	Extrême vocalique antérieure langue	Extrême du triangle vocalique
<b>P</b>	Occlusion, explosion extrême	Visible en lecture labiale
<b>AU</b>	Voyelle	Grande fréquence d'occurrence

Figure 5 : Caractéristiques de la phrase de référence /Bonjour Monsieur Tralipau/



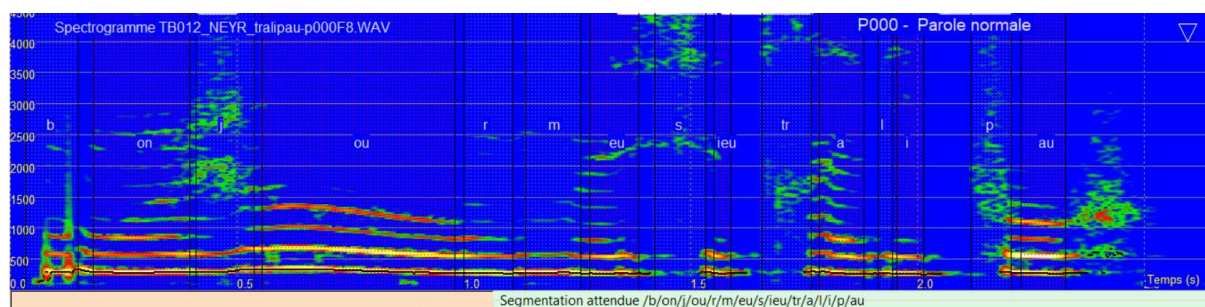


Figure 6 : Spectrogramme de la phrase /Bonjour Monsieur Tralipau/, score maximum 30, fille de 8 ans parole normale

La segmentation de la phrase /Bonjour Monsieur Tralipau/ est relativement simple et peu sujette à la régionalisation ni à la contamination. Elle a été construite dans le but de balayer l'essentiel des processus de parole tels que le voisement, la constriction, l'occlusion, le triangle vocalique, la nasalité, les diphtongues et les clusters. Cette phrase se révèle d'une grande efficacité : non biaisée par la mémoire verbale car l'effet de récence permet d'aider le sujet à répéter correctement le logatome. Un exemple de segmentation conforme à la cible, permettant d'obtenir un score maximum 30/30 est donnée figure 6 et correspond à une fille témoin de 8 ans [Bardelang 2020].

### Phrases complémentaires

Nous avons ajouté deux phrases dans le but de balayer presque tous les phonèmes de la langue française complémentaire de /Bonjour Monsieur Tralipau/ (Figure 7).

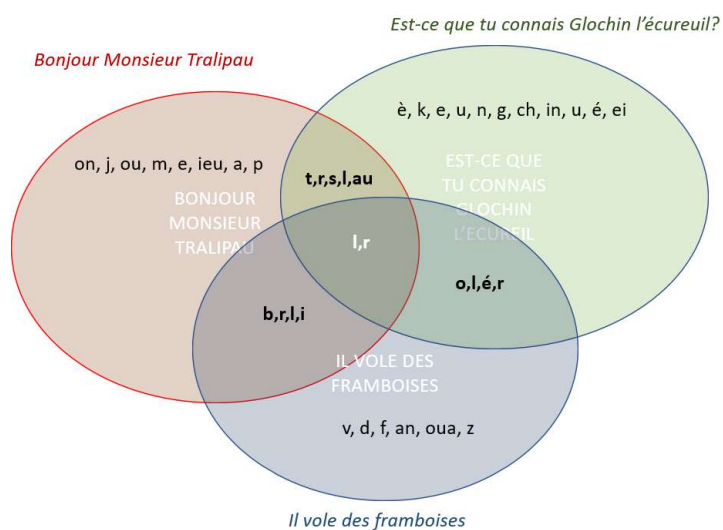


Figure 7 : illustration de la complémentarité entre les phrases proposées

La phrase /Est-ce que tu connais Glochin l'écureuil ?/ [Bardelang 2020] remplit le cahier des charges suivant :

- Elle complète une partie des phonèmes non présents dans /Bonjour Monsieur Tralipau/
- Elle est interrogative avec une formulation très fréquente et usuelle : /Est-ce que ... /, et propose ainsi un contour intonatif différent et varié.
- Elle présente une structure syntaxique ainsi qu'un lexique accessible à de jeunes enfants même en situation de bilinguisme, ne favorisant ainsi aucun milieu socio-éducatif en particulier
- Elle n'est pas trop longue pour ne pas surcharger la mémoire de travail et la segmentation (Figure 8)

- Elle est cohérente au niveau du sens global et facilement compréhensible hors contexte
- Elle contient un logatome qui n'est pas sensible à la lexicalisation : « Glochin ». Cela permet d'évaluer la capacité de l'enfant à traiter et acquérir de nouvelles formes sonores.
- Elle contient un mot usuel de trois syllabes : « écureuil »
- Elle contient les clusters /sk/ /tr/ et /gl/. Le cluster /sk/ est particulièrement intéressant car il s'agit d'une locution très utilisée par les jeunes enfants. De plus, il combine deux lieux articulatoires éloignés (alvéolaire puis vélaire), en opposition avec le cluster /gl/ (vélaire puis alvéolaire).

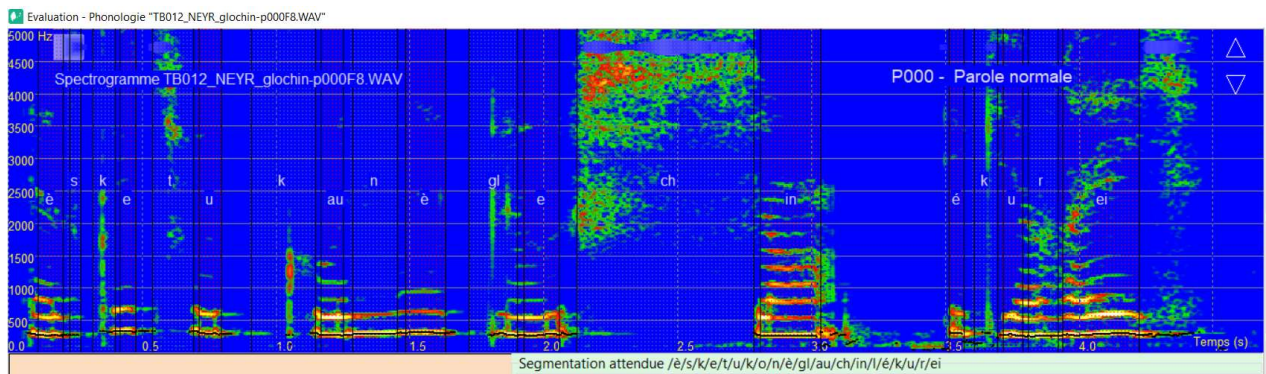


Figure 8 : Spectrogramme et segmentation de la phrase « Est-ce que tu connais Glochin l'écureuil ? », fille de 8 ans parole normale

Enfin, une 3<sup>ème</sup> phrase, couplée à /Bonjour Monsieur Tralipau/ et à /Est-ce que tu connais Glochin l'écureuil ?/ complète la couverture des phonèmes de la langue française. La phrase /Il vole des framboises/ remplit le cahier des charges suivant [Bardelang 2020]:

- Elle complète une partie les phonèmes non présents dans /Bonjour Monsieur Tralipau/ et de /Est-ce que tu connais Glochin l'écureuil ?/
- Elle présente une structure syntaxique ainsi qu'un lexique accessible à de jeunes enfants même en situation de bilinguisme, ne favorisant ainsi pas un milieu socio-éducatif en particulier
- Elle n'est pas trop longue pour ne pas surcharger la mémoire de travail.
- Sa longueur et sa simplicité syntaxique se rapprochant de celle de Tralipau font qu'elles semblent toutes deux aussi faciles à produire (Figure 9)
- Elle est cohérente au niveau du sens global et facilement compréhensible hors contexte
- Elle contient le cluster /fr/particulièrement intéressant car il combine deux lieux articulatoires extrêmes (labio-dentale et uvulaire)
- Elle contient le mot usuel « framboise » qui est complexe au niveau de ses cibles articulatoires.

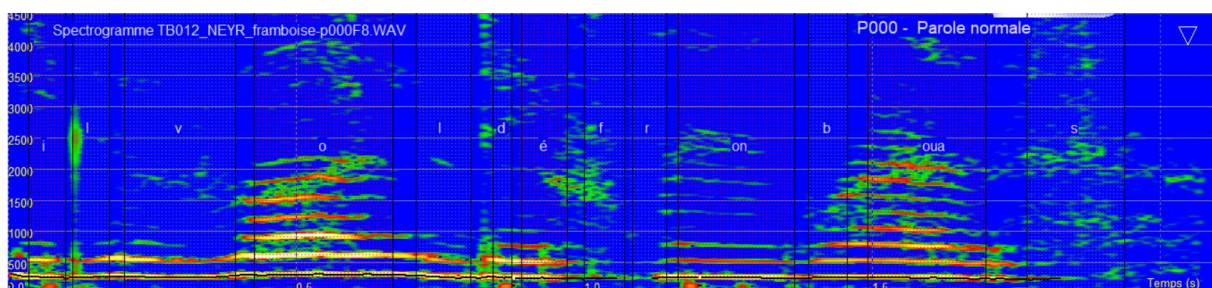


Figure 9 : Spectrogramme et segmentation de la phrase « il vole des framboises ».

Dans le cadre d'un mémoire d'orthophonie, une étude a été menée afin de savoir si évaluer la phonologie avec la phrase de référence /Bonjour Monsieur Trapliau/ était aussi fiable et aussi efficace que d'évaluer la phonologie avec les trois phrases citées plus haut, qui balaient ensemble la quasi-totalité des phonèmes de la langue française. Cette étude a été menée auprès de 92 enfants entre 5 et 8 ans par [Bardelang 2020]. Les résultats ont montré que la phrase Tralipau ne requiert l'évaluation des deux autres phrases que dans les cas où le résultat obtenu est limite (c'est-à-dire, proche du 10<sup>e</sup> percentile). Dans tous les autres cas, la phrase Tralipau, qui serait encore plus fine si elle comportait davantage de voyelles nasales, est alors fiable et suffisante pour évaluer la phonologie.

### Coarticulation et contamination :

La coarticulation est un phénomène naturel de modifications phonémique pouvant être accentuée par un débit élevé. La coarticulation est particulièrement visible dans /ieu/ de /Monsieur/ (Figure 10) : la frontière entre /i/ et /eu/ est très difficile à situer, dans la mesure où l'évolution de la structure formantique du /i/ vers le /eu/ est continue. C'est pour cette raison que nous considérons /ieu/ comme une diphtongue et non comme deux voyelles séparées.

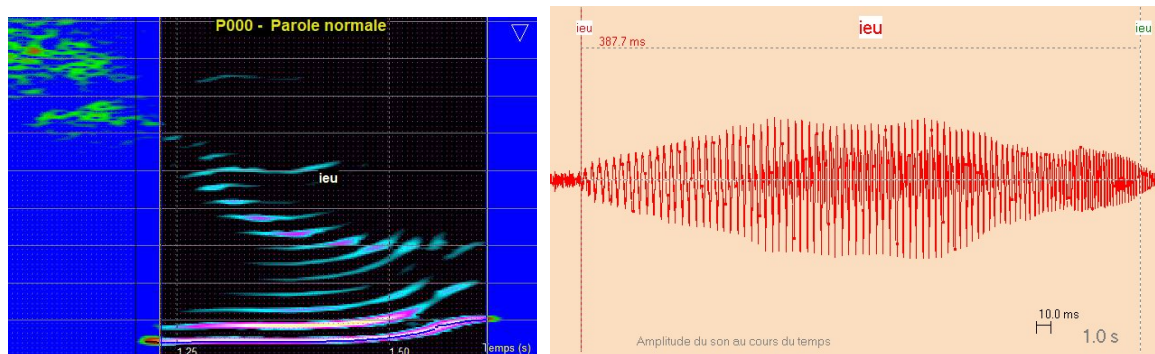


Figure 10 : coarticulation du /i/ et /eu/ dans /Monsieur/, spectrogramme et allure temporelle

La contamination est une coarticulation trop élevée où un phonème est structurellement modifié dans sa portion stable sous l'influence du précédent ou de celui qui suit. Concernant le cas SM010 (Garçon de 4 ans, trouble de la parole), le mot /Tralipau/ est perçu comme /t/e/i/p/au/ (Figure 11), avec des voyelles /e-i/ très peu différenciées, au moment du bilan initial. Après rééducation, même si le mot cible n'est pas encore bien prononcé, il n'y a plus de contamination entre les voyelles, l'espace fréquentiel des voyelles /a,i,au/ est beaucoup plus large, avec une différenciation temporelle nette (Figure 12).

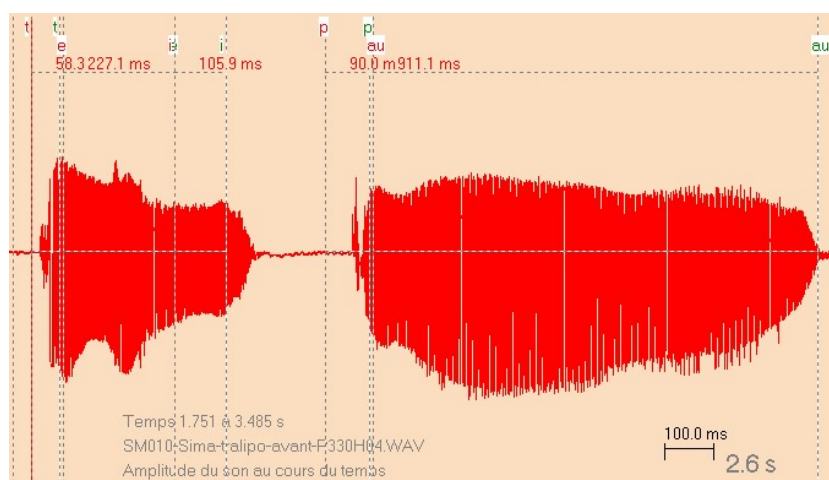


Figure 11 : Allure temporelle de /tralipau/ avec contrastes faibles et une contamination du /e-i/ (SM010 avant)

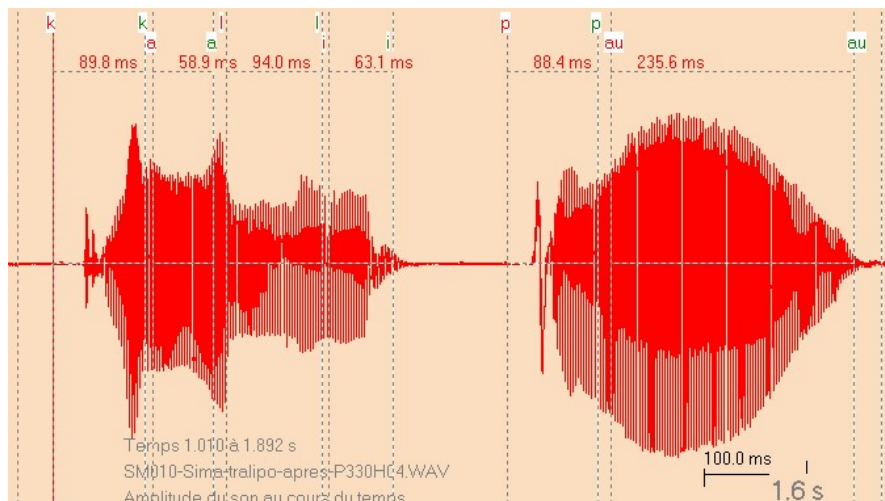


Figure 12 : Allure temporelle du mot /Tralipau/ avec contrastes plus nets, et une bonne séparation /a-i- au/ (SM010 après)

## Outils d'aide à la segmentation spectrale

Nous décrivons ci-après les principaux d'outils d'aide à la segmentation : le spectrogramme, l'allure temporelle, ainsi que certains marqueurs spécifiques.

### Spectrogramme

Le spectrogramme représente l'évolution dans le temps des composantes fréquentielles de la parole. Il peut être considéré comme l'un des piliers de l'analyse de la parole, et un outil très précieux pour la segmentation. Très informatif, impartial, d'une précision bien supérieure à l'oreille humaine, le spectrogramme donne des images très contrastées des voyelles et des consonnes, permettant de positionner les frontières de manière assez efficace. Une description détaillée de l'utilisation du spectrogramme en clinique orthophonique est proposée dans [Sicard 2021]. Un exemple de positionnement des frontières du mot /Tralipau/ est reporté figure 13.

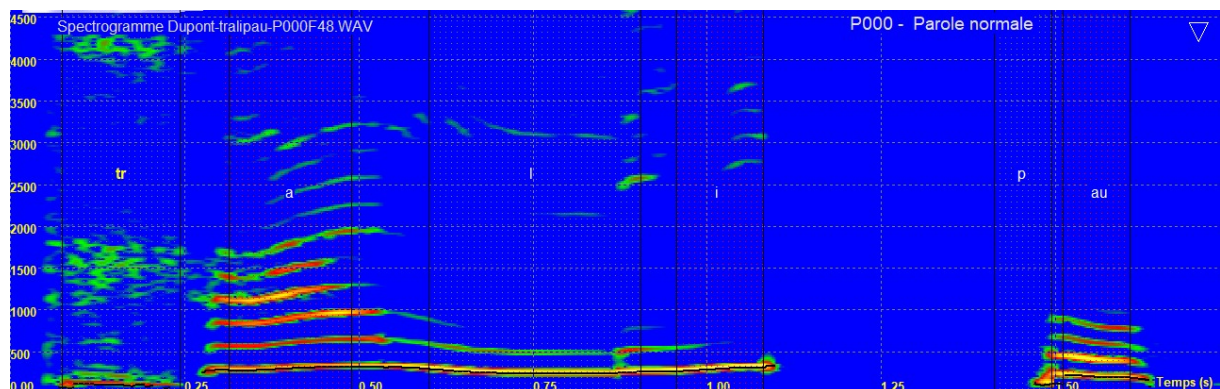


Figure 13 : Définition des segments de phonèmes (dont un cluster /tr/) sur la base du spectrogramme

On note ici une segmentation « au cœur du phonème » en évitant de superposer la fin d'un phonème et le début d'un autre, ce qui évite d'inclure dans les segments les transitions, silences et zones de coarticulation.

### Allure temporelle

L'allure temporelle permet un ajustement très précis du début et de la fin de la production sonore d'un phonème.

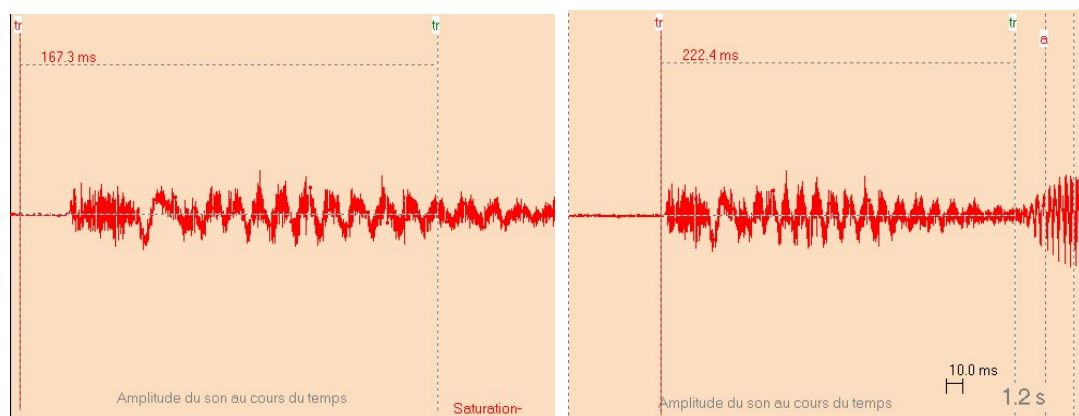


Figure 14 : Ajustement des frontières (ici le segment /tr/) à l'aide l'allure temporelle

Dans la figure 14, les frontières initialement posées sur la base du spectrogramme sont réajustées de manière à faire coïncider précisément le début et fin du cluster /tr/ avec la zone de forte constriction. Cet ajustement n'a pas d'incidence précise sur l'exploitation ultérieure des données et la fiabilité des résultats cliniques, mais permet de faire porter des analyses statistiques telles que la durée typique de chaque phonème [Sicard 2019].

### Ecoute sélective

L'écoute sélective (en général consonne-voyelle ou voyelle-consonne) permet d'isoler le phonème de son contexte et supprimer de ce fait toute béquille ou repère pouvant influencer la décision du clinicien sur le phonème perçu. L'incapacité à identifier un phonème, comme dans le cas de « bouillies articulaires », peut se traduire par la cotation de type « phonème indéterminé » représenté par le symbole /?/. La réécoute partielle et phonémique, sélective aléatoire ou dans le désordre aide l'opérateur à résister à l'attraction du modèle.

### Spectre moyen

La moyenne du spectre dans l'intervalle de production du phonème permet d'obtenir le profil de l'énergie en fonction de la fréquence. Associé à la détection des formants et par comparaison aux profils type, ou en superposant deux phonèmes comme ici le /a/ et le /au/ (Figure 15). Ces représentations aident le clinicien à évaluer la distance entre le modèle et la production moyenne, ou encore les contrastes entre deux phonèmes prononcés par le patient. Le spectre moyen constitue alors une aide à la décision pour le codage du segment.

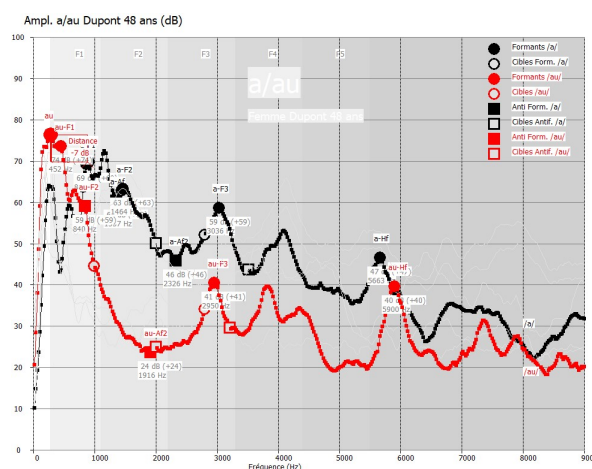


Figure 15 : Profils comparés du /au/ et du /a/ : énergie en fonction de la fréquence

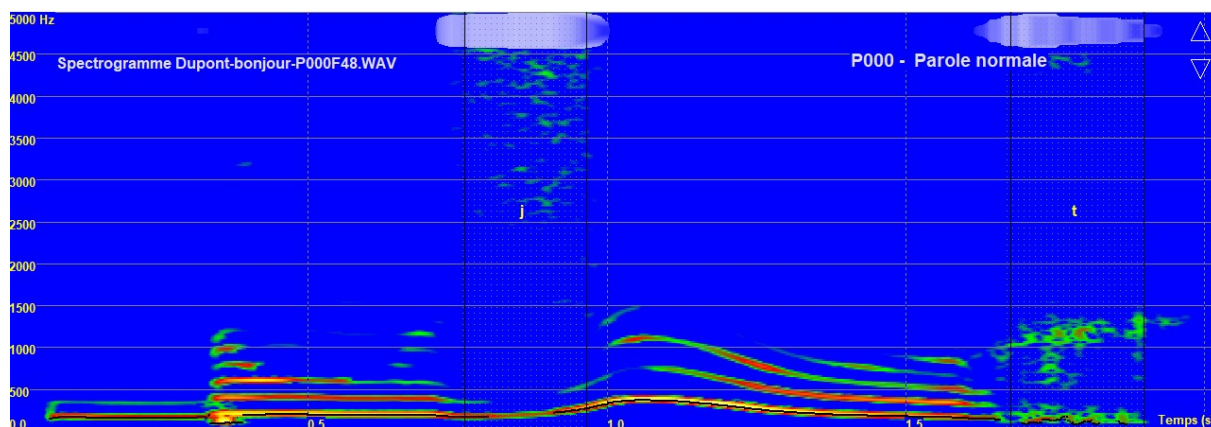


Figure 16 : Indice de constriction superposé au spectrogramme pour faciliter la segmentation du /j/ et du /r/ dans /Bonjour/

### Marqueurs spécifiques

Nous avons ajouté au spectrogramme une information complémentaire à l'énergie fonction du temps et de la fréquence : le degré de constriction. Cette information apparaît en haut du spectrogramme sous forme d'une palette de blanc. Plus il y a de « nuages » blancs, plus la constriction est élevée. La quantification de la constriction affichée permet donc d'aider à repérer les zones de constriction. Dans le mot /Bonjour/, ce marqueur facilite le positionnement du segment /j/ et /r/, comme le montre la figure 16.

### Segmentation assistée

L'intérêt d'une segmentation assistée est de raccourcir le temps de segmentation en se reposant sur les capacités du logiciel à positionner correctement les frontières, puis en faisant une réécoute sélective pour affiner les altérations. Dans la mesure où nous connaissons bien la durée typique des phonèmes dans le contexte des phrases types décrites au chapitre précédent, nous pouvons tenter de superposer la segmentation attendue à la production sonore. Il faut cependant prendre un certain nombre de précautions :

- Bien identifier le début et la fin du son
- Avoir une réalisation de la phrase proche du modèle typique
- Disposer de quelques indices pour aligner les phonèmes à l'intérieur de la phrase : constriction, explosions, voyelles.

Un premier niveau de segmentation assistée consiste à placer tous les phonèmes sur la production, quel que soit le niveau d'altération, sous forme d'un alignement forcé. Le résultat est plutôt bon pour la parole normale (Figure 17) ou faiblement altérée, mais se révèle vite inexploitable pour la parole pathologique. Par comparaison avec les modèles existants, il est possible de demander au logiciel de donner un avis dans un choix restreint d'alternatives, en lien avec les confusions les plus fréquentes. Par exemple, l'outil de segmentation assistée considère que le /j/ de bonjour est plus proche du /ch/, ce que le clinicien peut confirmer ou infirmer par écoute sélective.

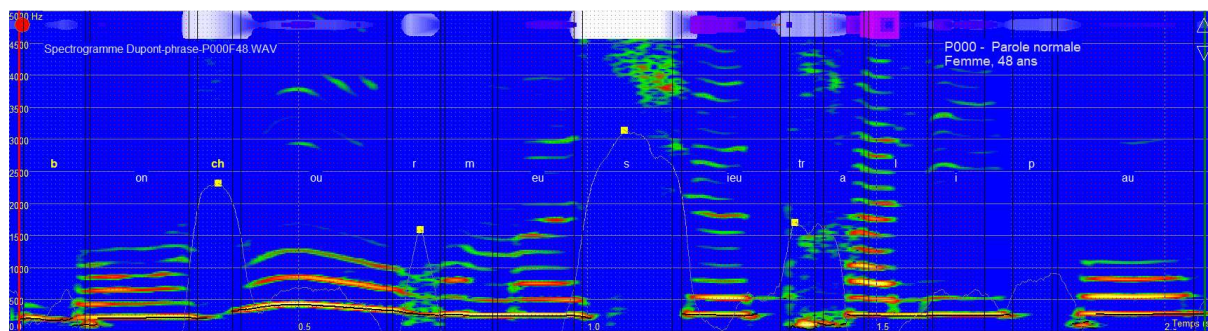


Figure 17 : Segmentation assistée de la phrase /Bonjour Monsieur Tralipau/

### Segmentation automatisée

La mise en place de la segmentation automatisée suppose que l'outil de reconnaissance des phonèmes est significativement plus fiable que l'oreille du clinicien, et qu'il a été validé sur un très grand nombre de cas, le tout avec une qualité d'analyse très élevée.

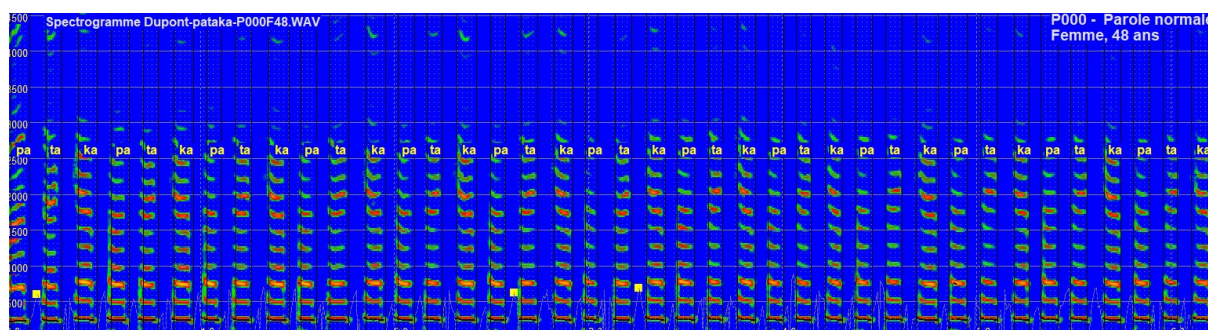


Figure 18 : Segmentation automatisée des séries de /PaTaKa/ disponible dans DIADOLAB

La segmentation automatisée a pu être mise au point dans le cas spécifique de la diadococinésie sur des répétitions de /PaTaKa/ [Sicard 2020], comme le montre la figure 18. Elle a été testée avec succès sur des corpus de grande taille tels que ceux de dysarthrie de J. Rusz [Rusz 2015] et du projet Voice4PD [Das 2019], évitant ainsi une segmentation manuelle fastidieuse et très répétitive de segments identiques. Elle fonctionne cependant mal en situation de réverbération, les frontières entre syllabes devenant floues et les limites devenant imprécises.

La segmentation automatisée de phrases telles que celles décrites au chapitre précédent, pour tous types d'altération, avec une haute qualité d'analyse, que ce soit pour homme, femme ou enfant, reste encore à ce jour au stade du prototype. Nous avons entamé dans le cadre de l'ERU 46 du LURCO la construction d'un corpus de référence pour lequel la qualité de la segmentation automatique va être évaluée par étapes successives, en espérant mettre disposition des cliniciens un outil de pré-segmentation et évaluation automatisée des confusions, qui devrait ensuite être vérifié et éventuellement ajusté par l'orthophoniste.

### Cotation des altérations

Comment coter justement les altérations sans pénaliser des phénomènes considérés comme normaux ? Comment prendre en compte les différents paramètres sans pénaliser des sujets plus que d'autres ? Les discussions entre orthophonistes, cliniciens ou chercheurs nous ont amené à réaliser un tableau de classification précis des altérations en tentant d'être au plus près de la réalité terrain et en s'affranchissant des classifications du domaine phonétique traditionnellement utilisées, théoriques, peu adaptées aux pathologies de la parole.

La cotation des altérations que nous proposons obéit à une terminologie simple non ambiguë, factuelle, liée à une réalité avant tout acoustique et physiologique. Le choix du terme ne devant pas représenter une difficulté surajoutée à la segmentation.

### Typologie des erreurs

Nous décrivons ci-dessous les différentes altérations repérables par DIADOLAB.

**Substitution** : Celle-ci est la plus fréquente. Nous faisons une différence entre une substitution proche (un seul trait omis) ou une substitution éloignée (deux traits omis), entre les voyelles et les consonnes. Toute substitution d'un trait donne lieu au retrait d'un point ou encore lorsque le phonème est peu marqué. Pour les voyelles : lorsque la substitution concerne un phonème proche (ex : ou/o ; a/è), DIADOLAB compte -1 mais sanctionne plus durement lorsque les voyelles sont très éloignées (ex a/i, u/è) -2.

**Imprécision** : L'imprécision concerne les consonnes et les voyelles. Lorsque la substitution n'est pas claire, l'orthophoniste a un doute et après réécoutes, ne parvient pas à identifier clairement l'un ou l'autre des oppositions : ex : le phonème /ch/ qui se rapproche d'un /s/ ou un /a/ qui se rapproche d'un /è/, le segmentateur cote -1 tout comme la substitution. Dans le cas où le sujet a un trouble d'articulation (sigmatisme, schlintement, rhotacisme), cela rentre dans le cadre des imprécisions liées à une malposition linguale. Cela pourra être noté cliniquement par ailleurs dans le dossier du patient.

**Omission de clusters, syllabes, mots, logatomes, locutions** : Les omissions de phonèmes donne lieu à une sanction de -2 points par phonème omis. Les cas des omissions est plus complexe qu'il n'y paraît à première vue. Plusieurs cas de figures peuvent se présenter. Dans le cas de l'omission d'une syllabe entière : il peut y avoir des syllabes de deux ou trois phonèmes. L'omission d'une syllabe comportant 3 phonèmes ex : /tra/, le /tr/ étant considéré comme un cluster, entité unique, on ne sanctionnera que deux points par « entité » donc moins -4 ici. En cas d'omission d'un mot : la question du statut du logatomes par rapport au mot usuel s'est posée. Nous proposons d'enlever 4 points lorsqu'il s'agit du mot « Monsieur » et 6 points lorsqu'il s'agit du logatome Tralipau. Dans ce cas précis, un deuxième essai permet au sujet d'être évalué sur le plan phonologique et non pas mnésique.

**Ajout de syllabes ou de phonèmes** : l'ajout incongru de phonèmes qu'ils soient vocaliques ou consonantiques est sanctionnée à -2. Cette altération est considérée comme une altération plus rare et atypique. Il peut faire l'objet d'une segmentation sauf dans le cas du « e » à la fin d'un mot considéré comme une régionalisation. L'ajout d'une syllabe est plus délicat car considéré comme une disfluence. Il peut être segmenté mais il ne sera pas sanctionné en phonologie. Cette erreur sera prise en compte dans le score de fluence sauf si la syllabe ajoutée est incongrue. Ex : Bonjour Monsieur taralipau : le cluster tr est simplifié (-1) un /a / est ajouté (-2). Si l'enfant dit : « Bonjour Mr tra...tralipau », le premier /tra/ peut être segmenté mais comptera comme une répétition et sanctionnée uniquement en disfluence.

**Répétition de phonèmes ou de syllabes** : la répétition de phonème n'est pas sanctionnée dans le score de phonologie mais dans le score de fluence.

**Ebauche de phonèmes ou encore phonèmes « fantômes »**. Le segmentateur peut percevoir le phonème lors de l'écoute sélective mais il ne le voit pas ou ne l'identifie pas clairement sur le spectre. Cela peut être le cas lorsque le phonème est seulement ébauché ou de très faible amplitude (atonie). Cela suffit parfois pour créer l'illusion du phonème (simple mouvement de recul de langue pouvant donner l'impression auditive que le sujet a prononcé le phonème /r/ par exemple). Si le phonème perçu est « intermédiaire », le segmentateur va être plus naturellement orienté vers le phonème attendu plutôt que vers le phonème opposé. Certains phonèmes se prêtent plus que d'autres à cet



effet. Le /l/ par exemple, phonème dit liquide se traduit sur le spectre par une sorte de petite « pelote » à basse fréquence. Il est parfois difficile à repérer. Le phonème /r/ en position finale peut être concerné aussi. Le sujet peut donner l'illusion qu'il le réalise en reculant légèrement la langue après le phonème/ou/ mais la constriction n'a pas lieu, n'est pas visible sur le spectre. Soit l'orthophoniste le considère comme une ébauche et le cote -1, soit comme une omission et elle le cote -2.

**L'effet de masquage de la vibration par la constriction** peut être aussi rencontré. Par exemple pour le /j/ du mot « Bonjour ». Une vibration nette correspondant au trait de voisement est repérée sur le spectre mais à l'écoute sélective, il est assez nettement perçu comme un phonème dévoisé comme /ch/. Dans la mesure où le sujet a réalisé la vibration, on pourrait le gratifier, sauf que la sensation auditive est bien celle du phonème sourd et non du sonore. L'hypothèse serait que la constriction marquée (hautes fréquences) masque un voisement ébauché donnant ainsi la sensation du phonème sourd. C'est aussi le cas lorsque la vibration est ébauchée, ou lorsqu'il s'agit simplement d'une rémanence de la vibration du phonème précédent ou de la préparation du phonème suivant. La coarticulation est considérée comme normale si elle reste limitée aux bords des phonèmes concernés. Dans ce cas, c'est la sensation auditive qui prime, et n'est sanctionné que de -1 si elle ne correspond pas à la cible.

**Phonèmes impossibles à identifier** : le segmentateur peut aussi se retrouver devant un phonème impossible à catégoriser même en écoutant à plusieurs reprises le phonème isolé ou dans une sélection plus large, comme si le phonème n'existait pas dans notre langue ou ne ressemblait à rien de connu. Ces altérations se retrouvent dans des cas de dysarthrie, bilinguisme ou d'insuffisance vélaire ou encore de chirurgies réparatrices : coups de glotte, bruits de souffle nasale, clicks étranges ou autres bruits. Dans ce cas, il est noté /?/ et cela sera sanctionné -2 car erreur considérée comme atypique et rare sur ces phrases.

**Régionalisation** : Certains phonèmes vocaliques sont prononcés différemment par les locuteurs en fonction des régions. C'est le cas des phonèmes /è/ pouvant être prononcés /é/, /e/>/eu/, /o/>/au/ ou bien l'ajout d'un /e/ en fin de mot, fréquent dans le sud. Ces modifications minimales ne sont pas pénalisées car considérées comme des régionalisations.

**Dysphonie** : Dans notre base de données, plusieurs sujets ont été repérés comme étant dysphoniques. Le spectre apparaît bruité, le FUM (Fondamental usuel moyen décalé vers les graves) et certains phonèmes vocaliques sont partiellement voire complètement désonorisés. Les consonnes voisées peuvent être aussi partiellement ou totalement dévoisées en fonction de leur longueur moyenne. Lorsque la dysphonie est légère et qu'elle ne se traduit que par un spectre un peu bruité, elle impacte peu l'intelligibilité et elle n'est pas sanctionnable sur le plan phonologique. Elle sera simplement notée en observation clinique par ailleurs. Par contre lorsque cela se traduit par des désonorisations partielles ou complètes de voyelles ou de consonnes, ce sera pris en compte en substitution (-1). Cela affectera dans ce cas, le score de phonologie donc l'intelligibilité.

## Cotation des altérations

Nous présentons ci-dessous le tableau de cotation des altérations.

TYPE D'ALTERATIONS	PERCEPTION AUDITIVE	SPECTROGRAMME	COTATION	EXEMPLES
OMISSION - <b>DEUX POINTS PAR PHONEMES</b>	Segment auditivement non perceptible	Segment invisible spectralement	-2 par phonèmes	onjour monsieur Tralipau
PHONEME INIDENTIFIABLE	Perceptible mais non catégorisable ni identifiable	Visible et segmentable noté ?	-2	<i>Click, bruit singulier coup de glotte ou autre</i>
SUBSTITUTION (2 TRAIT) - <b>CONSONNES</b> - <b>VOYELLES TRES ELOIGNEES</b> - <b>CLUSTERS</b> - <b>DIPHONGUES</b>	Aisément perceptible en réécoute sélective	Bien visible et segmentable sur le spectre. Substitution identifiable	-2	<i>bonsour kalipau tralipi i/au glochu u/in i/o a/ou</i>
AJOUT DE PHONEMES - <b>VOYELLES (HORS REGIONALISATION)</b> - <b>CONSONNES</b>	Aisément perceptible en réécoute sélective	Visible et segmentable sur le spectre	-2	<i>Golochin Tralkipau  Taralipau</i>
SUBSTITUTION 1 TRAIT - <b>ORALISATION</b> - <b>ASSOURDISSEMENT</b> - <b>NASALISATION</b>	Moins facile à percevoir  Plusieurs ré-écoutes nécessaires Doutes possibles	A rapprocher des profils types Standard pour valider + allure temporelle	-1	bojour bonchour tralimau bojour
IMPRECISION - <b>VOYELLES</b> - <b>CONSONNES</b> - <b>SIGMATISME</b> - <b>SCHLINTMENT</b> - <b>DIPHONGUE</b>	Moins facile à percevoir  Plusieurs ré-écoutes nécessaires Doutes possibles	A rapprocher des profils types Standard pour valider + allure temporelle	-1	baéjour
INVERSION	Facile à percevoir	Bien visible et Segmentable sur le spectre	-1	tarlipau trapilau
EBAUCHE - <b>VOYELLES</b> - <b>CONSONNES</b>	A peine perceptible	Portion de segment visible sur le spectre	-1	
REPETITION - <b>PHONEMES</b> - <b>SYLLABES</b> - <b>MOTS</b>	Facile à percevoir auditivement	Bout de segment visible sur le spectre. À segmenter pour que Diadolab puisse le traiter	<b>0</b> <b>Sanctionnée Dans la fluence uniquement</b>	

Tableau 1 : Classification et cotation des altérations

CRITERE	PERCEPTION AUDITIVE	SPECTROGRAMME	COTATION	EXEMPLES
REGIONALISATION	Peu perceptible	A peine visible	<b>0</b>	é/è o/au eu/e in/un ajout d'un /e/ final
COARTICULATION NORMALE	Déformation visible chez la grande majorité des tout-venants	Visible sur le spectre + réécoute sélective	<b>0</b>	Le /t/ de /tu/ fricativé Le /o/ de glochin substitué en /e/
DYSPHONIE	Substitution /désonorisations Peu perceptibles	Visibles sur l'ensemble du spectre. Analyse spectrale du A tenu à faire en complément	<b>0</b> <i>Non sanctionnée dans le score d'articulation mais dans le module prosodie</i>	

Tableau 2 : Typologie des altérations et cotation de DIADOLAB

## Discussion

Dans cette discussion, nous notons

- « Tralipau » : /Bonjour Monsieur Tralipau/
- « Glochin » /Est-ce que tu connais Glochin l'écureuil ? »
- « Framboise » : /Il vole des framboises/

### Approche progressive de la segmentation

Pour apprendre à segmenter, nous suggérons de commencer par des exemples de parole de sujets sans difficulté particulière, pour aller progressivement vers de la parole altérée puis pathologique. A cette intention, nous avons séparé les 3 mots de la phrase « Tralipau » dans le cas d'étude « Mes patients\Dupont femme », afin de faire la segmentation de chaque mot isolé et faire ainsi le lien entre les aspects temporels et fréquentiels sur des cas simples. La segmentation est un travail d'analyse minutieux qui demande des compétences en phonétique et acoustique. Dans la mesure où la phrase étudiée est toujours la même et les motifs attendus aisément identifiables, l'expertise en segmentation augmente rapidement avec la pratique.

### Apports de la phrase « Glochin »

La remontée terrain des cliniciennes concernant l'enregistrement des phrases dans le cadre du bilan de la parole a mis en évidence les points suivants :

- « Glochin » suscite des rires et/ou de la surprise. L'ensemble /Glochin l'écureuil / est parfois mal compris : certains enfants la prononcent /le chant de l'écureuil/.
- La phrase « Glochin » est plus longue et plus complexe à segmenter que « Tralipau »

La phrase « Glochin » présente de nombreux phonèmes coarticulés pouvant être tolérés et elle est plus sujette à coarticulation que Tralipau. Cela concerne 21 sujets/74 sujets tout-venants.

Dans, « Glochin », le /o/ est fréquemment substitué en /e/, le /t/ de « tu » souvent fricativé, à cause de la position très antérieure de la langue qui ne laisse pas la place suffisante pour réaliser une véritable occlusion. Cela concerne 39 sujets/74 sujets tout-venants ce qui est considérable. Cette phrase est plus sensible à la régionalisation de par la présence des phonèmes : é/è, o/au/in et l'ajout de e à la fin des mots.

- La phrase est construite en trois unités signifiantes : /est-ce que/ /tu connais/ / Glochin l'écureuil/. Le logatome est ainsi au milieu, ce qui prive le sujet de l'effet de récence et de primauté.
- Elle est très intéressante au niveau de la prosodie car elle est plus riche que « Tralipau » et peut ainsi donner des informations complémentaires.
- La formulation « est-ce que » peut être difficile à produire par les plus petits (en-dessous de 5 ans) du fait de sa complexité articulatoire.
- La réponse à la question au lieu d'une répétition, peut évoquer des difficultés de type pragmatique (difficultés à adopter le point de vue de l'autre).

### Apports de la phrase « Framboise »

Concernant « Framboise », la remontée terrain des cliniciennes a mis en évidence les points suivants :

- « Vole » est parfois compris « veulent »
- La séquence « vole des » est difficile à segmenter car elle est naturellement coarticulée
- Elle est très sensible aux désonorisations : v/f, d/t, p/b, s/z. Le /z/ de « framboises » est fréquemment dévoisé pour 51 cas sur 72 tout-venants, ce qui est beaucoup car positionné en toute fin d'énoncé.
- Elle est sensible aux confusions vocaliques : on/an

Dans une étude réalisée par [Grenier 2020] dans le cadre de son mémoire d'orthophonie, les scores d'articulation obtenus à ces trois phrases ont été comparés. Sur la base de 30 patients, nous obtenons un coefficient de corrélation de l'ordre de 70% entre la phrase « Tralipau » et « Glochin » et entre la phrase « Tralipau » et « Framboise ». Cette corrélation élevée tend à montrer une similitude des résultats obtenus sur les aspects articulatoires et phonologiques étudiés, notamment en termes de processus de parole.

### Analyse comparée des scores

Certains enfants même tout-venant peuvent avoir de meilleurs scores à la phrase « Glochin » qu'à la phrase « Tralipau », ce qui semble à première vue surprenant puisque « Glochin » est plus longue et plus complexe. Il est noté que dans l'étude réalisée par [Bardelang 2020], 10 sujets sur 80 étaient dans ce cas de figure.

Lorsque l'on compare les scores d'articulation entre plusieurs segmentateurs de niveau d'expérience différents, ces derniers ne varient que de l'ordre de 4% par rapport au score maximal. Les différences inter-expérimentateurs apparaissent donc faibles [Grenier 2019]. Cette valeur est à rapprocher de variations qui pourraient être observées lors de transcriptions phonétiques manuelles par des cliniciens de niveau d'expertise différent.

### Conditionnement phonémique

Le protocole Diadolab proposé dans le logiciel doit être respecté afin que les données, l'extraction des paramètres soit tangible et la comparaison à des normes pertinente. Il est en effet difficile de s'extraire du conditionnement phonémique lié à notre apprentissage linguistique et à la familiarité de la tâche. La tendance à « coller » à la structure phonémique idéale, à celle du modèle, à celle que l'on voudrait entendre est très forte.

L'évaluation objective contribue à structurer la perception auditive et développe nos aptitudes à percevoir les altérations phonologiques fines. Cependant l'oreille perçoit la parole comme un tout et ne peut dissocier les différents paramètres surtout lorsqu'ils sont très brefs ou qu'une seule composante du phonème est altérée. Des questionnements ont émergé des phases de segmentation en fonction des différents cas cliniques rencontrés. Plus de 600 segmentations ont été conduites grâce à la base de données de parole créées dans le cadre de l'ERU46 (Laboratoire LURCO) et avec l'aide précieuse d'une trentaine de cliniciennes.

## Conclusion

Dans cet article, nous avons défendu l'intérêt de l'évaluation objective de la parole au travers de l'enregistrement et de la segmentation d'une phrase type, une approche susceptible de fiabiliser les scores de phonologie et de limiter l'impact des biais de l'évaluation à l'oreille. Après avoir décrit différents biais auxquels sont soumis les cliniciens dans leur pratique, nous avons comparé l'efficacité des différentes phrases et les résultats en termes de score d'articulation obtenus pour chacune d'elle. Nous proposons une méthodologie précise de segmentation spectrale en s'aidant non seulement du spectrogramme mais aussi d'autres outils tels que la segmentation assistée, l'allure temporelle, la réécoute phonémique sélective et l'aide des profils-type. Nous avons dressé un tableau exhaustif de la typologie des altérations avec leur cotation associée puis décrit l'approche par processus hiérarchisés qui permet au clinicien de disposer d'un guide méthodologique clair et efficace pour le suivi des patients atteints de trouble de l'intelligibilité et de la fluence.

## Remerciements

Nous tenons à remercier les orthophonistes contributrices à l'ERU 46 du laboratoire LURCO ainsi que les orthophonistes qui ont accepté d'enregistrer des témoins et des patients ou de co-encadrer des mémoires en lien avec la parole et ses pathologies. Nous saluons aussi les membres des projets C2SI et Voice4PD qui nous ont permis, par le biais de corpus conséquents, d'affiner nos outils d'analyse notamment en lien avec la segmentation, les spectres moyens et l'extraction des formants.

Nous saluons aussi les efforts de mise en archive ouvertes d'un très grand nombre de résultats scientifiques, en particulier dans le domaine de la parole, ce qui permet une diffusion très large, simple et sans droit d'accès, autorisant de ce fait des analyses comparatives et méta-analyses qui accélèrent ainsi leur mise en application clinique. Enfin, nous remercions l'ensemble des orthophonistes, étudiants et chercheurs qui nous ont accordé leur confiance en utilisant nos méthodologies, approches et outils dans le cadre de leur pratique clinique ou de leurs travaux scientifiques.

## A propos des auteurs



**Etienne SICARD** est professeur à l'INSA Toulouse en électronique, informatique et traitement du signal. Il est directeur de recherches de l'équipe ERU 46 sur la parole pathologique au Laboratoire LURCO et chercheur associé à l'IRIT dans le cadre du projet ANR Voice4PD sur le diagnostic différentiel des

Syndromes Parkinsoniens. Etienne SICARD est co-auteur des logiciels VOCALAB et DIADOLAB pour le bilan et la rééducation orthophonique de la voix et la parole, ainsi que MICROWIND et IC-EMC en lien avec les circuits intégrés. Il a co-écrit une vingtaine d'ouvrages et plus de 200 publications scientifiques dans le domaine de l'analyse de la voix et parole, des circuits intégrés, et de la compatibilité



**Anne MENIN-SICARD** est orthophoniste et titulaire d'un master recherche en Sciences du langage de l'université de Grenoble. Elle a exercé en libéral à Toulouse de 1994 à 2015. Anne MENIN-SICARD est co-auteur des logiciels VOCALAB et DIADOLAB. Dans le cadre de AMS Formation,

elle organise et anime des formations aux orthophonistes sur le bilan et la prise en charge de la voix et parole pathologique. Elle a fondé en 2019 Opop, pour la formation en ligne personnalisée à destination des orthophonistes. Chercheuse associée à l'ERU 46 du LURCO, elle développe des outils de mesures objectives et de rééducation de la parole en lien avec le logiciel DIADOLAB. Elle est auteur d'un

électromagnétique. Il a été nommé en 2006 *Distinguished Lecturer* de la société savante IEEE.

Site : [www.etienne-sicard.fr](http://www.etienne-sicard.fr)

Email : [etienne.sicard@insa-toulouse.fr](mailto:etienne.sicard@insa-toulouse.fr)



Tiphaine BARDELANG a obtenu son diplôme d'orthophonie du Centre de formation en Orthophonie de l'Université Claude Bernard de Lyon en Juin 2020. Elle a été membre étudiante du LURCO/UNADREO dans le cadre de son mémoire qui portait sur l'étude comparative de deux protocoles d'évaluation acoustique de la parole, dans le cadre du logiciel DIADOLAB3. Son mémoire était co-encadré par Anne MENIN-SICARD et Océane BARBERA. Elle est clinicienne libérale à Lyon depuis Janvier 2021.

Email : [tiphaine.bardelang@gmail.com](mailto:tiphaine.bardelang@gmail.com)

#### **Sandrine MICHEL**

Sandrine MICHEL est orthophoniste en cabinet libéral à Castelginest. Elle est aussi rattachée à l'Unité Régionale d'Evaluation des TSA et autres troubles du neurodéveloppement, Hôpital La Grave à Toulouse. Elle est aussi chargée de cours au Centre de Formation des Orthophonistes de l'Université de Toulouse, dans les domaines tels que la rééducation vocale.

Dans le cadre de l'équipe ERU 46 du LURCO/UNADREO, elle conduit des recherches sur la parole pathologique, notamment en lien avec les troubles d'articulation, dyspraxies, surdités, mais aussi les dysarthries. Elle a co-encadré différents mémoires, notamment dans le domaine de l'autisme, ainsi que celui de C. Grenier en 2020 sur l'étude comparative de tests phonologiques existants et de protocoles d'analyse objective de la parole.

Email : [s.michel81@orange.fr](mailto:s.michel81@orange.fr)

ouvrage sur l'évaluation objective de la voix (DeBoeck, 2017) ainsi que de nombreuses publications dans le domaine de la voix et de la parole pathologique.

Site : <https://www.formationsvoixparole.fr>

Email : [anne.sicard2@orange.fr](mailto:anne.sicard2@orange.fr)



Caroline GRENIER a obtenu un diplôme d'ingénieur en informatique en 2001 à l'INP-ENSEEIH. Après avoir passé 15 années dans des sociétés de service pour le compte de groupes industriels aéronautiques, elle se réoriente vers une carrière d'orthophoniste. Elle obtient son diplôme d'orthophoniste en Juin 2020 à l'Université de Toulouse. Elle a été membre étudiante du LURCO/UNADREO dans le cadre de son mémoire qui évaluait l'intérêt de la mesure objective de l'intelligibilité et de la fluence à partir du logiciel DIADOLAB comme aide au diagnostic dans le cadre de difficultés de parole chez des enfants âgés de 3 à 7 ans. Son mémoire était co-encadré par Sandrine MICHEL et Etienne SICARD. Elle est clinicienne libérale à Tournefeuille depuis Septembre 2020.

Email : [carolinegrenier@yahoo.fr](mailto:carolinegrenier@yahoo.fr)

## Références

[AMS] <https://www.formationsvoixparole.fr>

[Boersma 2019] Boersma, Paul & Weenink, David (2019). Praat: doing phonetics by computer [Computer program]. Version 6.1.06, retrieved 8 November 2019 from <http://www.praat.org/>

[Bouchet 2008] Bouchet, M., & Boutard, C. (2008). Protocole d'évaluation de l'expression syntaxique (PEES). Isbergues: Ortho-édition.

[Bardelang 2020] Bardelang T. (2020), L'apport d'une phrase supplémentaire au protocole de DIADOLAB3. Mémoire d'orthophonie. Université de Lyon. Encadré par A. Menin-Sicard.

[Diadolab] <https://www.gerip.com/fr/home/73-diadolab-evaluation.html>

[Grenier 2020] Grenier C. (2020), Etude comparative de tests phonologiques existants et du protocole d'analyse objective de la parole avec Diadolab : intérêt et limites d'un test court et des mesures objectives d'intelligibilité et de fluence. Mémoire de Master d'orthophonie. Université de Toulouse. Co-encadrement : S. Michel, E. Sicard.

[Chevrie-Muller 1997] Chevrie-Muller, C., Simon, A. M., Fournier, S., & Brochet, M. O. (1997). L2MA: batterie, langage oral, langage écrit, mémoire, attention. Editions du Centre de psychologie appliquée.

[Coquet 2007] Coquet, F., Roustit, J., & Jeunier, B. (2007). Batterie EVALO 2-6 Évaluation du développement du langage oral et des comportements non verbaux du jeune enfant. *Rééducation orthophonique*, 45(231), 203-225.

[Das 2019] Das, B., Daoudi, K., Klempir, J., & Rusz, J. (2019, May). Towards disease-specific speech markers for differential diagnosis in Parkinsonism. In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 5846-5850). IEEE.

[Khomsy 2001] Khomsy, A. (2001). ELO: évaluation du langage oral. ECPA, les Éd. du Centre de psychologie appliquée.

[Maillart 2014] Maillart, C., & Durieux, N. (2014). L'évidence-based practice à portée des orthophonistes: intérêt des recommandations pour la pratique clinique. *Rééducation orthophonique*, 257, 71-82.

[Menin-Sicard 2019] Menin-Sicard, A., Sicard, E. (2019). Méthodologie d'évaluation objective de la phonologie, de la fluence et de la prosodie - Vers un bilan rapide à destination des orthophonistes. Journées de Phonétique Clinique JPC 2019, May 2019, Mons, Belgique. <https://hal.archives-ouvertes.fr/hal-02127039>

[Menin-Sicard 2019b] Menin-Sicard, A, Sicard, E. (2019) DIADOLAB 3 - Logiciel d'évaluation et de rééducation de la parole - Manuel d'utilisation. Archives ouvertes du CNRS <https://hal.archives-ouvertes.fr/hal-02073314>

[Menin-Sicard 2020] Menin-Sicard, A., Sicard, E. (2020). Outils de mesures et d'évaluation objective de la phonologie, de la fluence et de la prosodie avec DIADOLAB3, *Rééducation Orthophonique*, N° 281, Le bilan Orthophonique. pp 161-168.

[Rusz 2015] Rusz, J. (2015). Speech disorders reflect differing pathophysiology in Parkinson's disease, progressive supranuclear palsy and multiple system atrophy, *Journal of Neurology*. 2015; 262(4): 992

[Sicard 2017] Sicard, E., Menin-Sicard, A., Daumet, M., Perriere, S., & Barbera, O. (2017). Evaluation de la voix avant/après suivi orthophonique-à propos de la méthodologie d'évaluation et de la fiabilité des mesures chiffrées.

[Sicard 2019] Sicard, E., Menin-Sicard, A. (2019) Analyse de la qualité des phrases pour un bilan objectif de la parole. [Rapport de recherche] INSA Toulouse. 2019. <https://hal.archives-ouvertes.fr/hal-02389764>

[Sicard 2020] Sicard, E., & Menin-Sicard, A. (2020). Analyse d'une phrase type dans le cadre du bilan orthophonique de la parole. Rapport de recherches en archives ouvertes. INSA Toulouse, LURCO/UNADREO. <https://hal.archives-ouvertes.fr/hal-02568392>

[Sicard 2020b] Sicard, E., & Menin-Sicard, A. (2020). La diadococinésie et son application en orthophonie clinique. Rapport de recherches en archives ouvertes. INSA Toulouse, LURCO/UNADREO, <https://hal.archives-ouvertes.fr/hal-02512856/>

[Sicard 2021] Sicard, E., Menin-Sicard, A. (2021) Le spectrogramme et son application en clinique orthophonique. INSA Toulouse, LURCO/UNADREO, <https://hal.archives-ouvertes.fr/hal-03107434>

[Thibault 2003] Thibault, M. P., Helloin, M. C., & Croteau, B. (2003). Exalang–5/8. Une batterie d'examen du langage oral et écrit chez l'enfant de 5 à 8 ans. *Revue Tranel (Travaux neuchâtelois de linguistique)*, 38, 129-152.

[Wioland 1991] WIOLAND, F., (1991). *Prononcer les mots du français*. Hachette