

Tutoriel Prosogram

Prosogram est un outil d'analyse de la prosodie qui fournit la stylisation de la courbe de fréquence fondamentale d'un signal de parole. La spécificité de cette stylisation réside dans le fait qu'elle est vise à simuler la perception tonale.

Ce tutoriel aborde les points suivants : installation du logiciel, choix des paramètres (type de segmentation, seuils perceptifs) ...

1. Installation

1.1. Généralités

Prosogram est un script (un programme) qui suppose le logiciel Praat et qui produit des fichiers graphiques dans plusieurs formats, à savoir EPS et EMF, et également PDF sous Mac OSX. Les images EPS peuvent être affichées, imprimées et converties à l'aide de logiciels comme Ghostscript et GSView. Pour cette raison, quand on travaille sous Windows, il faut d'abord installer ces logiciels sur son ordinateur, si ce n'est pas encore le cas. Les fichiers EMF (Enhanced Metafile) sont un format propre à Microsoft et on ne les utilise généralement que sous Windows. Les utilisateurs de Mac n'ont pas besoin d'installer Ghostscript et GSview.

L'installation des logiciels suppose les manipulations générales suivantes.

- Afficher le contenu d'un répertoire sur le disque dur de votre ordinateur. (Pour les utilisateurs de Windows nous recommandons l'utilisation de l'application « Windows Explorer » qui fait partie de Windows.)
- Créer un nouveau répertoire sur votre ordinateur.
- Télécharger un fichier à partir d'un site internet. Cliquer avec le bouton droit sur le lien du fichier, choisir « Save link as... » ou « Save target as... », puis sélectionner le répertoire souhaité.
- Installer un logiciel sur votre ordinateur. Quand vous avez téléchargé un fichier exécutable sur votre ordinateur, il suffit de double-cliquer sur le nom du fichier pour lancer l'installation. (Attendre la fin de l'installation et lire les messages pour identifier les problèmes éventuels.)
- Ouvrir un fichier archive ZIP : double cliquer sur le nom de fichier. Pour en extraire des fichiers, choisir « Extract » et sélectionner le répertoire où on souhaite les placer.

Noms de fichier sous Windows

La plupart des noms de fichier comportent deux parties, séparées par un point : le nom de base et l'extension. L'extension indique normalement la nature et le format du fichier, par exemple « .doc » pour un document au format Word, « .gif » et « .jpg » pour des images aux formats GIF et JPEG, « .exe » pour certains programmes exécutables, et ainsi de suite.

Malheureusement, Windows n'affiche pas systématiquement l'extension. On peut alors à tort avoir l'impression qu'il y a plusieurs fichiers du même nom dans un répertoire ou que le fichier recherché n'y figure pas. Pour obtenir l'affichage des extensions sous Windows XP, aller dans « Settings | Control Panel | Folder Options | View » et désactiver « Hide file extensions for known types ».

1.2. Choix du répertoire des données

Il est recommandé de créer un répertoire dédié aux scripts de Prosogram et aux données. Sous Windows, évitez les chemins aux noms longs comme « My Documents » ou « Documents and Settings » et choisissez plutôt un chemin court du genre « C:\proso ». Ceci vous évitera de retaper à chaque fois le chemin long.

1.3. Installation de Praat

Sur le site <http://www.praat.org>, dans la rubrique "Download Praat", suivre le lien qui correspond à votre système de gestion (Windows, Macintosh, etc.). Télécharger le fichier en question, puis le lancer (double cliquer sur son nom) pour installer Praat.

Installer également les polices (fonts), comme indiqué sur la page de téléchargement de Praat.

1.4. Installation de Ghostscript et de GSview sous Windows

Installer Ghostscript. Sur le site <http://www.cs.wisc.edu/~ghost/>, choisir le lien « APFL Ghostscript ». Télécharger le fichier « gs853w32.exe » (taille 9.5 Mo) (ou une version plus récente) sur votre ordinateur. Lancer ce fichier pour installer Ghostscript sur votre ordinateur.

Installer GSview. Télécharger ensuite sur la même page le fichier « gsv47w32.exe » (taille 1.5 Mo) sur votre ordinateur et le lancer pour installer GSview sur votre machine.

Il est important d'installer Ghostscript avant GSview.

1.5. Installation de Ghostscript et de GSview sous Mac

Sous Mac OS X on peut utiliser le format graphique PDF (Postscript), qui s'ouvre dans l'application « Aperçu » (ou « Preview »). Dans ce cas, il n'est pas nécessaire d'installer Ghostscript et GSview.

Sous Mac OS X les fichiers EPS s'ouvrent dans l'application « Aperçu » (ou « Preview »). Cette application convertit d'abord le fichier EPS au format PDF avant d'afficher celui-ci.

En principe on n'aura pas besoin du format EMF sous Mac OS X. Pour afficher quand même les fichiers EMF on peut installer des outils tels que « NeoOffice », une version d'OpenOffice pour Mac.

Si vous souhaitez installer quand même GSview sous Mac (parce que l'affichage de GSview est supérieur à celui de Preview), il faut d'abord installer X-windows pour Mac, puis installer Ghostscript et GSview, par exemple à l'aide de Fink. Cependant, l'installation de X-windows sous Mac OS X varie d'une version à l'autre et est complexe, donc déconseillée au débutant.

1.6. Installation de Prosogram

Les scripts liés à Prosogram sont disponibles sur le site de « <http://bach.arts.kuleuven.be/pmertens/prosogram> ». (Il suffit de rechercher « prosogram » sur Google pour arriver sur cette adresse.) Suivre le lien « Download ». Télécharger le fichier « prosogram_vXX.zip » (où XX indique la version) dans le répertoire de données (cf. 1.2.).

1.7. Fichiers exemples

Télécharger de la même page « Download » le fichier « testdata.zip » dans le répertoire de données (cf. 1.2.). Ce fichier contient des exemples pour tester le fonctionnement de Prosogram.

Les fichiers avec l'extension « zip » sont des archives compressées utilisées couramment sous Windows. Pour les décompresser, on utilise un logiciel comme WinZip ou 7zip (<http://www.7-zip.org/>). Si vous n'avez pas ce type de logiciel sur votre ordinateur ou si vous

travaillez sous Mac OS X (et si celui-ci ne permet pas de l'ouvrir), vous pouvez télécharger les fichiers séparés depuis la page « Download ».

1.8. Vérification du fonctionnement de GSview (Windows)

Parmi les fichiers exemples dans « testdata.zip » (que vous avez téléchargé) dans le répertoire de données (cf. 1.2.), il y a « fg00150wr.eps ». Lorsqu'on double-clique sur le nom de ce fichier, une fenêtre intitulée « GSview » devrait apparaître avec l'image en question. Essayez les boutons « + » et « - » pour changer la taille de l'image.

2. Utilisation du prosogramme

Le logiciel Prosogram vise à réaliser une transcription de la hauteur perçue d'un enregistrement de parole, et plus particulièrement d'un corpus oral.

2.1. Nature de la transcription

Il existe plusieurs types de transcription prosodique, allant de l'analyse acoustique de la fréquence fondamentale (F_0) à une notation abstraite (de nature phonologique) à l'aide de symboles. Dans le prosogramme nous optons pour une représentation de la *mélodie perçue*, afin d'obtenir une notation objective, indépendante de telle ou telle théorie linguistique. En effet, ces théories abondent et il importe donc de disposer de données objectives et reproductibles pour en évaluer la validité.

Comme les recherches linguistiques sur la langue parlée font intervenir plusieurs types d'annotation (transcription textuelle orthographique, phonétique, prosodique, syntaxique, sémantique, etc.), le prosogramme permet d'afficher ces annotations en même temps que la transcription de la prosodie perçue.

2.2. Choix du type de segmentation

Le prosogramme peut être réalisé de plusieurs façons, avec ou sans alignement phonétique.

- La *segmentation automatique*.
Elle permet d'obtenir des transcriptions prosodiques d'un échantillon important de parole *dans un temps minimal*. La segmentation est basée sur les propriétés acoustiques du signal de parole, en l'occurrence sur l'intensité du signal filtré. L'inconvénient de ce type de segmentation réside dans les erreurs de segmentation éventuelles, quand le nombre d'éléments ne correspond pas exactement au nombre de syllabes.
- La segmentation basée sur un *alignement phonétique*.
L'avantage de cette approche est qu'il n'y aura pas d'erreurs en termes du nombre de syllabes. Quand l'alignement phonétique est réalisée à la main, il demande un investissement important. Mais dans certains cas l'alignement phonétique est déjà disponible ou peut être obtenu de façon semi-automatique ou automatique.
- La segmentation automatique corrigée manuellement.
Dans ce cas, on obtient d'abord la segmentation automatique, que l'on corrige à la main et qu'on sauvegarde pour l'utiliser ensuite comme entrée de la segmentation.

2.3. Préparation d'un fichier d'annotation (TextGrid)

S'il est possible d'obtenir le prosogramme d'un signal de parole sans aucune annotation, le plus souvent on voudra ajouter une transcription orthographique ou phonétique afin d'interpréter la prosodie en relation avec la chaîne linéaire (les mots, les constituants, les phrases).

En l'absence d'un fichier d'annotation, celui-ci peut être créé à l'aide de Praat, selon la procédure habituelle. Dans Praat, un fichier d'annotation adopte un format appelé TextGrid.

- Lancer Praat.
- Lire l'enregistrement sonore pour lequel on veut créer un fichier TextGrid.
Sur le menu principal, utiliser « Open », puis « Read from file... » pour un signal bref (de moins de 30 s). Pour un signal long, utiliser plutôt « Open long sound file... ». Après la lecture du signal, il sera sélectionné dans la liste des objets.
- Une liste de commandes apparaît dans la partie droite à côté de l'objet sélectionné (en l'occurrence le signal de parole). Choisir « Annotate », puis « To TextGrid... ». Il faut alors remplir deux champs. Le premier, « Tier names », indiquera le nom des couches d'annotation souhaitées. Remplacer le contenu du champ par « text » pour indiquer qu'on aura une seule couche d'annotation (« tire ») comportant la transcription textuelle du signal. Laisser vide le champ « Point tiers ». Appuyer sur « OK ».
- Un nouvel objet apparaît dans la liste des objets et il est sélectionné. Pour le sauvegarder, cliquer sur l'option « Save » du menu principal, puis sur « Save as text file... », puis chercher le répertoire souhaité et indiquer le nom souhaité. Ce fichier devra être placé dans le même répertoire que le fichier son dont il est dérivé.

Le fichier TextGrid peut être modifié interactivement dans Praat. A cet effet, suivre la procédure décrite dans l'introduction, disponible sous le menu « Help », puis « Praat Intro... », puis « Intro 7. Annotation ». Alternativement on peut consulter un des textes introductifs disponibles sur le site www.praat.org.

2.4. Lancer le script Prosogram

Pour illustrer l'utilisation du prosogramme, on a prévu un fichier son, le fichier textgrid correspondant, ainsi que quelques fichiers graphiques réalisés avec Prosogram. Consultez la section 1.7. « Fichiers exemples » pour obtenir ces fichiers.

Chargement du script

D'abord il faut charger le script dans Praat selon la procédure habituelle.

- Lancer Praat.
- Sur le menu principal, cliquer sur « Praat », puis sur « Open Praat script... ». Dans le répertoire de données, sélectionner le fichier « prosogram.praat ».
- Une fenêtre apparaît intitulée « Script prosogram.praat » avec son propre menu, où apparaît la commande « Run ».

Quand on clique sur « Run » ou quand on tape Ctrl-R, le formulaire suivant apparaît.

2.5. Le formulaire du script Prosogram

Prosogram 2.7 (June 18, 2010)

Task: Prosogram and prosodic profile

Input files:
C:/corpus/*.wav

Analysis parameters:

Time range (s): 0.0 0.0 (=all)

F0 detection range (Hz): 60 (=default) 450 (=default)

Parameter calculation: Full (saved in file)

Frame period (s): 0.005

Segmentation method: Automatic: acoustic syllables

Thresholds: G=0.32/T², DG=20, dmin=0.035

Save intermediate data

Plotting options:

View: 4: Wide rich

Time interval per strip (s): 3.0

Tiers to show (*convert to IPA): *1, 2, 3

Pitch range (ST): 0 (=autorang) 100

Output mode: Fill page with strips

Output format: EPS (Encapsulated Postscript)

Output path and filename for EPS/EMF files (number and extension added automatically):
<input_directory>/<basename>_

Standards Cancel Apply OK

Ce formulaire contient les champs nécessaires pour réaliser des prosogrammes ainsi que les options qui permettent de modifier les paramètres de la stylisation et les modalités de l'affichage graphique. Tous ces aspects sont décrits dans le guide d'utilisation disponible sur la page Web dédiée au prosogramme. Cependant, il suffit de remplir deux champs pour obtenir des résultats satisfaisants dans la plupart des cas.

Les champs essentiels

1. Dans le champ « Input files » on fournit le nom du fichier son à traiter. Par exemple :

C:\proso\fg00150.wav (Windows)
C:/proso/fg00150.wav (accepté également sous Windows)
/Users/mertens/corpus/fg00150.wav (Mac OS)

Sous Mac OS X, le répertoire personnel de l'utilisateur apparaît généralement comme un sous-répertoire du répertoire « /Users/ ».

Si on laisse vide le champ « Input files », Prosogramme ouvrira une fenêtre qui permettra de sélectionner un fichier. Dans ce cas, il faut d'abord choisir les autres paramètres.

2. Dans le champ « Segmentation method » on choisit la méthode de segmentation à utiliser.

- « Automatic: acoustic syllables » sélectionne la segmentation automatique en syllabes. Ce mode ne suppose pas de fichier d'annotation. Cependant, si ce fichier est disponible (et quel que soit son contenu), il sera affiché dans le prosogramme, selon les modalités choisies par l'utilisateur (cf. infra).
- « Vocalic nuclei : local peak of vowels in tier "phon*" or tier 1 » sélectionne la segmentation basée sur l'alignement phonétique présent dans le fichier d'annotation.

Elle suppose donc un fichier d'annotation (TextGrid) avec une tire donnant l'alignement phonétique. Cette tire est repérée grâce à son nom, qui commencera par "phon", ou, en l'absence d'une tire répondant à ce critère, la tire 1.

La méthode utilisée par défaut est la segmentation automatique.

3. Les champs « Time range » indiquent la partie (l'intervalle temporel) du signal à analyser, à savoir les temps du début et de la fin. Les valeurs par défaut correspondent à l'ensemble du signal.

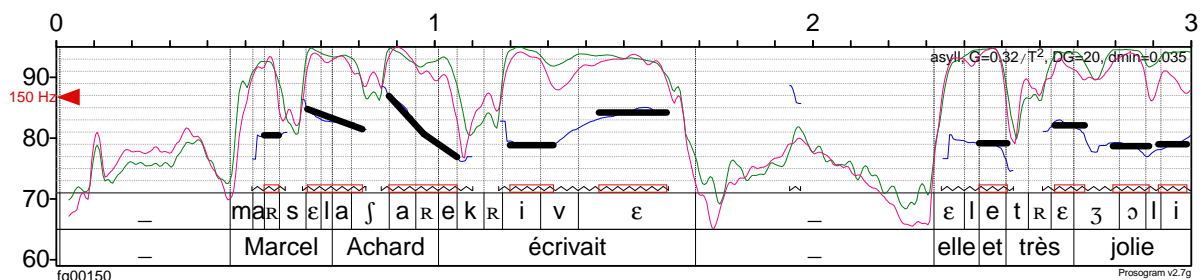
Calcul du prosogramme

Essayons l'analyse à segmentation automatique pour le fichier son « C:\proso\fg00150.wav », pour l'intervalle temporel compris entre 0 et 3 s (on indique 3 dans le champ droit de « Time Range »). On clique sur le bouton « Apply » en bas du formulaire d'entrée.

Après quelques instants on voit apparaître des courbes dans la fenêtre « Praat Picture ». En même temps un fichier graphique sera sauvegardé dans le répertoire de données, avec le nom « fg00150_001.eps ». (On revient plus loin sur les noms de fichier.) Affichez le prosogramme avec GView selon les indications données plus haut. Vous obtenez le résultat suivant.

2.6. Choix du format d'affichage

Les formats d'affichage diffèrent selon la taille de l'illustration et la quantité de paramètres affichés. Le format « large riche » donne le plus d'informations.



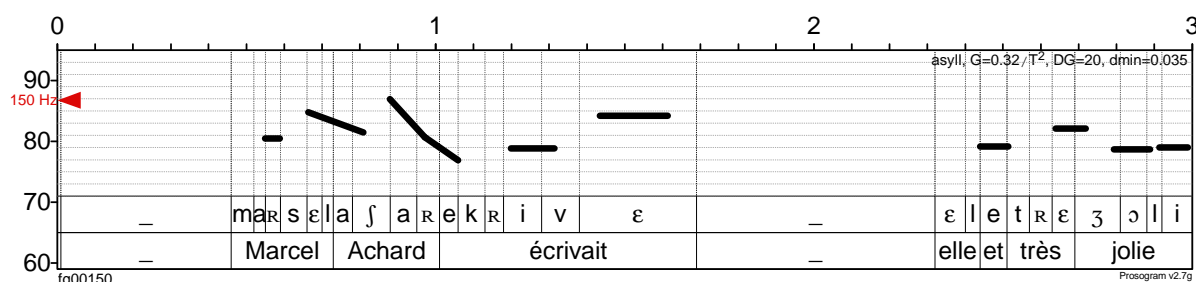
Le trait noir épais donne une approximation de la hauteur perçue. Il s'agit de hauteur stylisée selon un modèle de la perception tonale chez l'auditeur humain et appliqué aux intervalles segmentés.

D'autres informations apparaissent également dans cette analyse :

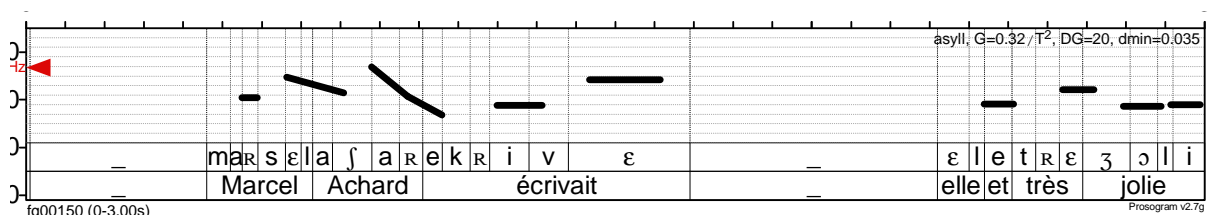
- la courbe bleue donne la fréquence fondamentale mesurée (F0), sur une échelle en demi-tons (relative à une fréquence de référence de 1 Hz) ; dans l'exemple en question elle est presque entièrement recouverte par le trait noir épais, ce qui permet de vérifier le fonctionnement de la stylisation ;
- la courbe verte donne l'intensité (en dB) ;
- la courbe en cyan donne l'intensité du signal filtré en passe-bande (utilisée pour la segmentation automatique) ;
- la ligne en zigzag indique les parties voisées ;
- les carrés rouges indiquent les intervalles utilisés pour la stylisation, qui correspondent aux noyaux syllabiques ;
- les traits pointillés horizontaux indiquent des repères (calibration) espacés de 2 demi-tons ;

- en haut sur l'axe horizontal apparaissent des repères indiquant le temps en secondes et en dixièmes de seconde.

Ci-dessous le format « large simple », qui ne donne que la stylisation.



Voici le format « compact simple ».



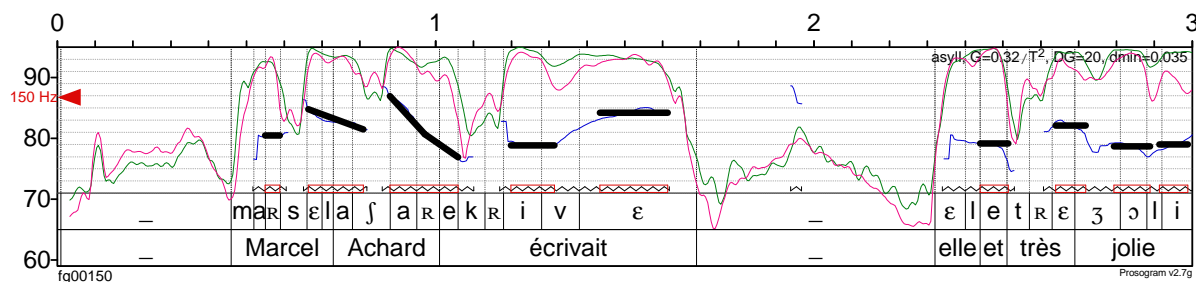
Dans tous les formats le nom (de base) du fichier son est indiqué dans le coin gauche inférieur.

Les variantes larges donnent la calibration de l'axe du temps (en seconde) et de l'axe de hauteur en demi-tons (par rapport à une valeur de référence de 1 Hz). Les axes ne sont pas calibrés dans le format compact. Cependant le temps est affiché après le nom du fichier analysé. La valeur de 150 Hz est indiquée par une flèche rouge dans tous les formats.

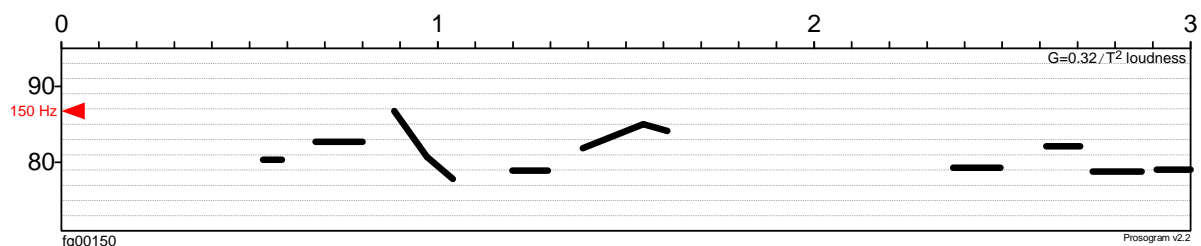
2.7. Sélectionner les tires d'annotation à afficher

Les exemples donnés plus haut comportent deux couches d'annotation, soit le nombre de couches affichées par défaut, si le fichier TextGrid en comporte deux ou plus.

Le champ « Tiers to shown (*convert to IPA) » permet d'indiquer les tires à afficher dans le prosogramme, en précisant leur nom ou leur numéro, par exemple « 1, 2 » ou « phon, syll ». Ça suppose évidemment que ces tires soient présentes dans le TextGrid. On peut préciser l'ordre dans lequel les tires sélectionnées apparaîtront dans le prosogramme : « syll, phon » ou « 3, 1, 2 ». Enfin, l'astérisque devant le nom ou numéro d'une tire indique que cette tire est au format SAMPA et doit être convertie en symboles phonétique de l'IPA : « *1, 3 ».



S'il n'y a pas de fichier d'annotation, ou son nom n'est pas conforme aux conventions adoptées par Prosogram, l'annotation n'est pas affichée. Cette possibilité n'existe que pour la segmentation automatique. Dans l'autre cas, un message d'erreur s'affiche.



2.8. Fichiers graphiques de sortie : formats et noms de fichiers

Prosogram fournit deux formats graphiques sous Windows : EPS (Encapsulated Postscript) et EMF (Windows Enhanced Metafile) et deux formats sous Max : EPS et PDF (Postscript). On peut sélectionner les formats souhaités dans le champ « Output Format » du formulaire du script.

- Le format EPS convient pour l’affichage à l’aide de GSview, pour les documents Postscript, Latex, etc.
- Le format EMF convient pour l’insertion dans des documents Word ou dans des présentations Powerpoint.
- Le format PDF convient pour l’utilisation sous Mac. Selon les informations sur www.praat.org, il peut être inséré dans des documents Word sauvegardés au format docx.

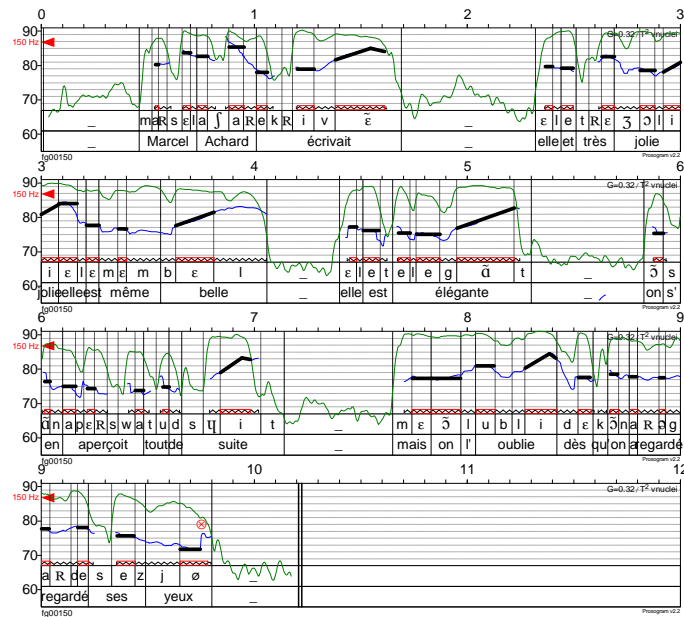
Ces fichiers seront placés dans un répertoire précis que l’on peut indiquer dans le champ « Output path and filename for EPS/EMF/PDF files (number and extension added automatically) » du formulaire.

- La valeur par défaut « <input_directory>/<basename>_ » utilise le répertoire et le nom du fichier son analysé.
- On peut remplacer « <input_directory> » par un *répertoire existant* ; les fichiers graphiques seront alors placés dans ce répertoire.

Un fichier son peut résulter en plusieurs fichiers graphiques successifs. Ils seront numérotés. Par exemple, pour un fichier « abc.wav » (où le nom de base correspond à « abc »), les fichiers graphiques seront nommés « abc_001.eps », « abc_002.eps », « abc_003.eps », etc.

Enfin, le champ « Output mode » permet de préciser le nombre de bandes (prosogrammes) par fichier graphique. Normalement, le script met autant de bandes que possible sur une page A4. Mais si on indique pour le champ « Output mode » la valeur « one strip per file » chaque fichier ne comportera qu’un seul prosogramme.

Ci-dessous un exemple d’un fichier comportant plusieurs bandes successives.



2.9. Calcul des paramètres

2.9.1. Sauvergarde des paramètres

La stylisation repose sur certains paramètres acoustiques comme la fréquence fondamentale (parfois appelée « Pitch ») et l'intensité. La segmentation automatique suppose l'intensité du signal filtré. Le script Prosogram tente de trouver les fichiers correspondants, mais s'ils ne se trouvent pas à l'endroit (répertoire) attendu, ces paramètres seront calculés et sauvegardés dans des fichiers (dans la plupart des cas). C'est la raison pour laquelle la stylisation se fait plus rapidement dès la deuxième analyse d'un même fichier son : les paramètres seront alors disponibles dans des fichiers.

2.9.2. Calcul partiel

Lorsqu'on travaille avec de gros fichiers son, d'une durée de 10 à 30 minutes ou même plus, le calcul des paramètres devient extrêmement long. Il arrive qu'on s'intéresse à un passage seulement d'un corpus très long, et que le calcul des paramètres du corpus complet est superflu. On a prévu la possibilité de réaliser des prosogrammes pour des passages précis et de limiter le calcul des paramètres aux passages en question. Pour obtenir ce fonctionnement, choisir pour le champ « Parameter calculation » la valeur « Partial (not saved in file) » et indiquer les temps du début et de fin souhaités.

2.9.3. Calcul préalable

Pour les gros fichiers son qui seront analysés dans leur totalité, il est plus efficace de faire le calcul des paramètres avant le calcul des prosogrammes. A cet effet, on choisit la tâche (« Task ») « Recalculate pitch for entire sound », ce qui permet en outre de modifier les valeurs minimale et maximale de la plage de fréquence fondamentale acceptée (dans « F0 detection range »). Alternativement on peut utiliser le script indépendant « prosoprep.praat » qui calcule les paramètres nécessaires et les met dans des fichiers.

2.9.4. Calcul de la fréquence fondamentale

Par défaut, Prosogram utilise la détection du fondamental fournie par Praat, qui correspond à la méthode d'autocorrélation modifiée. L'utilisateur peut entrer les valeurs minimale et maximale de la plage de fréquence acceptée dans les champs « F0 detection range ». Ces valeurs sont choisies en fonction de la tessiture du locuteur.

S'il est nécessaire de changer ces valeurs, il faut recalculer la fréquence fondamentale : on choisit la tâche (« Task ») « Recalculate pitch for entire sound ».

Il est possible d'utiliser des données de fréquence fondamentale obtenus avec d'autres logiciels, à condition que elles soient disponibles dans le format (« *.Pitch ») de Praat.

2.10. Traiter un corpus comportant plusieurs fichiers sons

Avec les gros corpus deux cas de figure se présentent : soit le corpus comporte plusieurs fichiers son plus courts, soit il ne comporte qu'un seul fichier long.

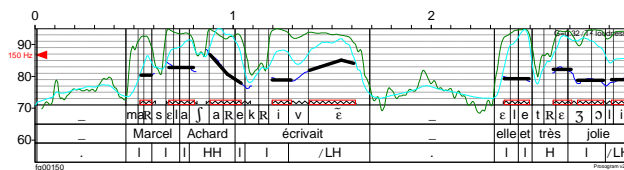
Lorsque le corpus comporte plusieurs fichiers son, les noms de fichier seront normalement dans l'ordre alphanumérique : a001.wav, a002.wav, a003.wav, etc. Pour réaliser les prosogrammes de tout le corpus, il suffit alors d'utiliser des wildcard dans la spécification des fichiers d'entrée : « C:\proso*.wav ». Les fichiers seront analysés les uns après les autres et les prosogrammes seront placés dans des fichiers sortie numérotés : 001.eps, 002.eps, 003.eps, etc.

Lorsque le corpus ne comporte qu'un seul fichier son, il suffit de préciser ce nom. Mais il sera généralement efficace de procéder au calcul préalable des paramètres (cf. supra).

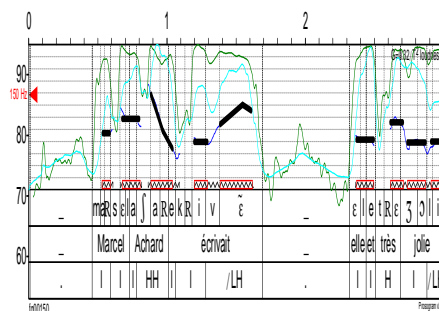
2.11. Utilisation des prosogrammes dans Word ou dans Powerpoint

Il suffit d'obtenir le prosogramme au format EMF et de l'insérer dans le document à l'aide de « Insert | Picture | From file... ».

Pour changer la taille de l'illustration, glisser (« drag ») un des coins pour garder les proportions du rectangle (« aspect ratio »).



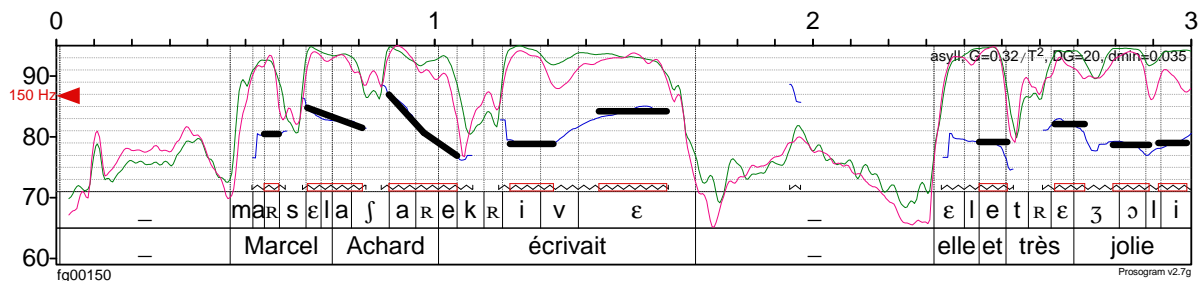
Au contraire, quand on déplace un des bords, la proportion du rectangle sera modifiée.



2.12. Conversion des prosogrammes vers d'autres formats graphiques

Les fichiers EPS peuvent être convertis vers d'autres formats graphiques tels que JPG ou TIFF à l'aide de GSview ou Ghostscript. Le script « eps_conv.praat » (disponible sur le site de Prosogram) permet de faciliter cette conversion. Il faut modifier les premières lignes de ce script pour préciser la place de ces logiciels sur votre machine.

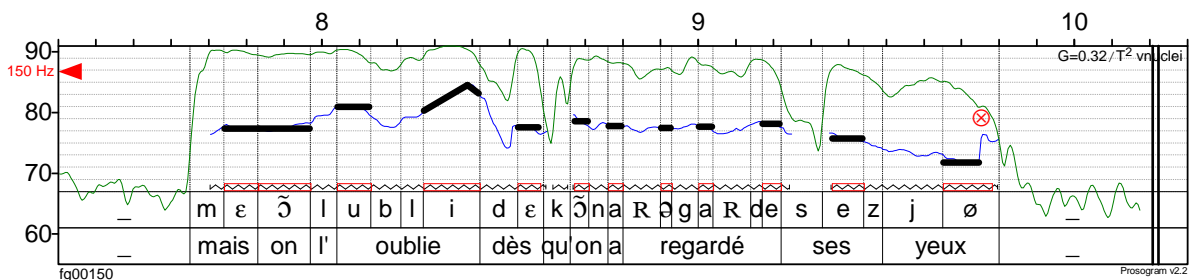
3. Interprétation des prosogrammes



3.1. Informations sur les paramètres d'analyse et d'affichage

- La légende dans le coin haut à droite indique
 - o le seuil de glissando utilisé pour la stylisation ($G=0.32/T^2$),
 - o le seuil de glissando différentiel utilise ($DG=20$),
 - o la durée minimale d'un segment tonal ($dmin=0.035$)
 - o le type de segmentation
 - vnuclei = voiced nuclei (noyaux vocaliques, base sur l'alignement phonétique dans le fichier d'annotation TextGrid) ;
 - asyll = automatic segmentation into syllables and syllabic nuclei, based on several acoustic parameters including intensity of filtered speech and intensity of full band signal.
- Dans le coin gauche en bas, apparaît le nom de base du fichier son analysé.
- Les pointillés horizontaux indiquent la calibration de l'axe de hauteur mélodique. La distance entre deux lignes successives équivaut à 2 demi-tons (2 ST).

3.2. Paramètres acoustiques affichés, annotation



- La ligne verticale double indique la fin du fichier son (vers 10.2 s dans l'exemple ci-dessus).
- Les traits pointillés verticaux indiquent les frontières des intervalles dans les couches d'annotation.
- Le temps apparaît sur l'axe horizontal, divisé en pas de 100ms.

Les données suivantes sont affichées dans le format riche uniquement.

- En bleu le fondamental mesuré (F0).
- En vert foncé, le tracé d'intensité (en dB).
- En magenta, le tracé d'intensité du signal filtré en bande passante (en dB).

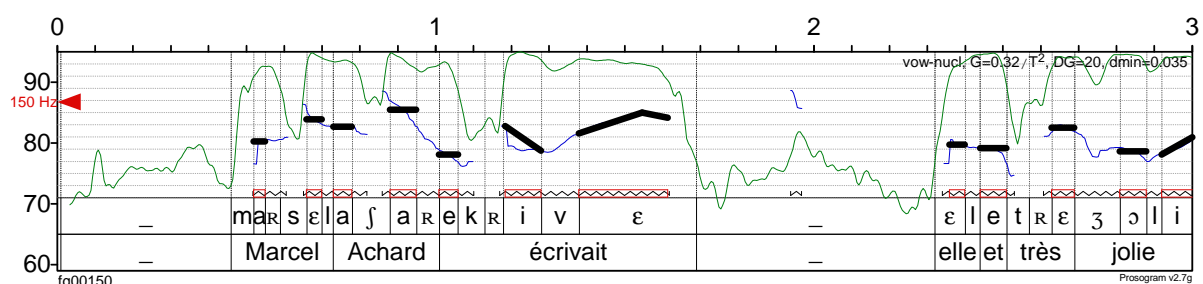
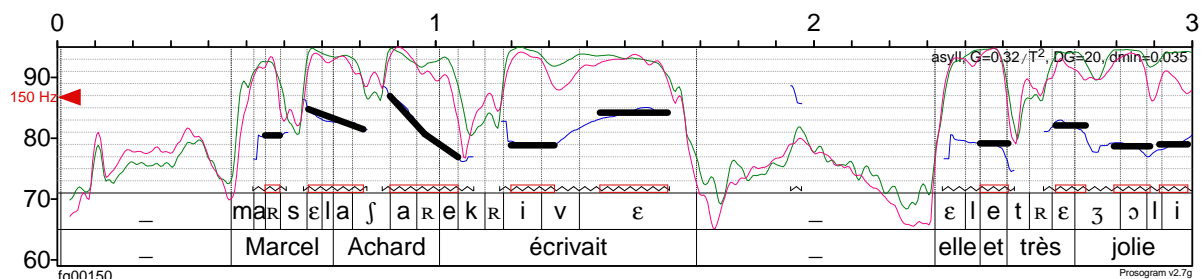
- En cyan, la sonie (intensité perçue, loudness), si elle est disponible.
- La ligne noire en zigzag indique une région voisée.
- Les rectangles en rouge indiquent les régions stylisées.

3.3. Interpréter la stylisation

- Le trait noir épais donne une approximation de la hauteur perçue. Il s'agit de hauteur stylisée selon un modèle de la perception tonale chez l'auditeur humain et appliqué aux intervalles segmentés.
- Le signe en forme de x encadré (dans la syllabe « yeux » dans l'exemple plus haut) indique (un instant temporel présentant) une discontinuité dans les données de fréquence fondamentale (trait fin en bleu). Comme ce type de discontinuité résulte le plus souvent d'une erreur de détection du fondamental, comme le saut d'octave, qui entraînerait des erreurs dans la stylisation résultante, l'intervalle temporel stylisé est limité à droite à l'instant de la discontinuité.

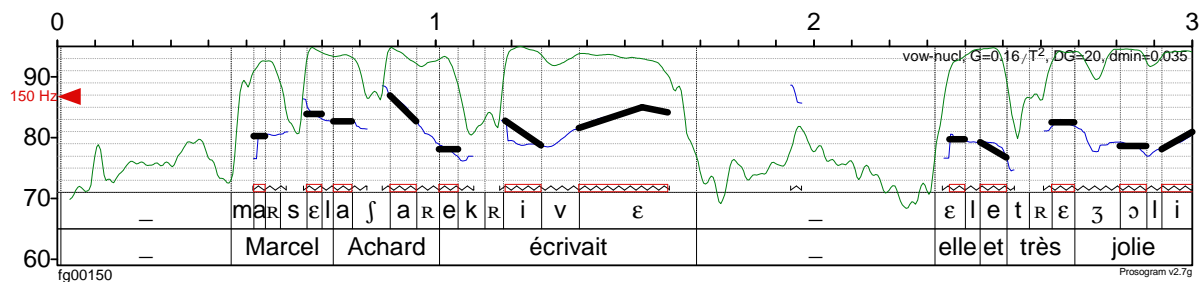
3.4. Impact du type de segmentation sur la stylisation

Il est illustré par les deux prosogrammes suivants. Le premier (A) utilise la segmentation automatique, le deuxième (B) les noyaux vocaliques. La différence est évidente pour la région de 0.850 s à 1.050 s. Dans A il n'y a qu'une seule unité de segmentation (qui correspond à la chaîne [aRe]) ; dans B il y en a deux ([a] et [e]). Les unités de segmentation sont données par les rectangles rouges. C'est à elles qu'est appliquée la stylisation. Dans le cas de [aRe] cette unité a une durée nettement plus longue que dans [a], ceci qui affecte évidemment la stylisation, le seuil de glissando étant lié au temps. C'est la raison pour laquelle il est crucial d'incorporer l'annotation textuelle et l'alignement phonétique si disponible.



3.5. Impact du seuil de glissando sur la stylisation

Il est illustré par le prosogramme suivant (C) qu'il faut mettre en regard avec B. Le seuil utilisé dans B est de $0.32 \text{ ST}/T^2$, dans C il est plus bas $0.16 \text{ ST}/T^2$, ce qui se manifeste par la présence d'un glissando dans la deuxième voyelle [a] du mot « Achard ». Le choix du seuil est discuté dans les publications sur le prosogramme.



4. Resynthèse

- Dans le formulaire d'entrée du script prosogramme, cocher l'option « Save intermediate data » pour sauvegarder le contour stylisé dans un fichier du type PitchTier. Après le calcul du prosogramme ce fichier est sauvegardé un fichier avec le nom « <basename>_styl.PitchTier », où <basename> représente le nom de base du fichier son analysé. Attention : la stylisation se limite à l'intervalle temporel indiqué dans le formulaire d'entrée.
- Lire le fichier son dans Praat.
- Créer un objet du type « Manipulation » à partir de cet objet son (« Sound »), comme suit. Sélectionner l'objet son et cliquer sur « To Manipulation... ». Choisir le « Time Step » identique à celui de la stylisation, c'est-à-dire à la valeur dans le champ « Frame Period » et fournir des valeurs adaptées pour « Maximum pitch » et « Minimum pitch ». L'objet manipulation apparaît dans la liste d'objets de Praat.
- Lire le fichier de la stylisation (styl.PitchTier).
- Sélectionner à la fois l'objet Manipulation et l'objet PitchTier. Cliquer ensuite sur « Replace Pitch tier ». Maintenant, l'objet Manipulation contient le pitch de la stylisation.
- Editer éventuellement l'objet Manipulation pour écouter le résultat.
- Sélectionner l'objet Manipulation et cliquer le bouton « Get resynthesis (PSOLA) ». Un nouvel objet son apparaît dans la liste des objets. Le sauvegarder de la façon habituelle.

5. Petites astuces

Dans certains cas, afin d'éviter de retaper les noms de fichiers ou les valeurs des paramètres, on peut les entrer directement dans le script « prosogram.praat » et changer les valeurs par défaut.

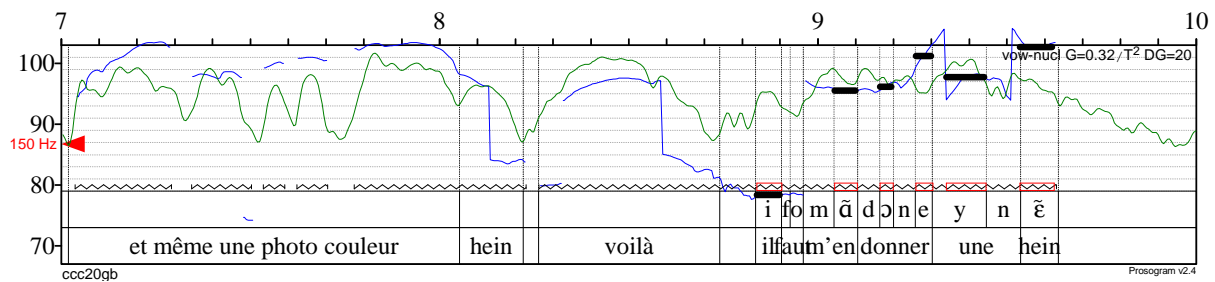
Quand on ouvre le script dans Praat, le texte source s'affiche dans une fenêtre et on peut le modifier. Après les commentaires du début (lignes commençant par #), il y a une partie intitulée « form Prosogram » où l'on retrouve les informations du formulaire Praat qui apparaît à l'écran quand on lance le script (avec « Run »). Il suffit de modifier les valeurs par défaut et de sauvegarder le fichier pour les retrouver la prochaine fois qu'on ouvre le script.

6. Problèmes

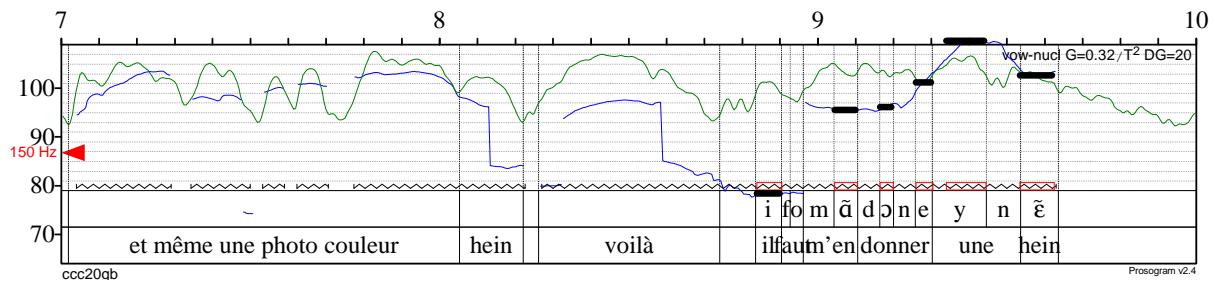
Erreurs de détection du fondamental

L'algorithme de détection de la fréquence fondamentale suppose que le fondamental se situe dans une plage de valeurs de fréquence, choisie en fonction de la tessiture du locuteur (homme, femme, enfant). Les valeurs de défaut de Praat sont de 75 à 500 Hz, celles du prosogramme de 65 à 450 Hz. Si pour une partie du signal analysé le fondamental effectif

dépasse la plage de fréquences sélectionnée, les valeurs affichées seront erronées. Dans la figure suivante, la plage de fréquence est de 65 à 450 Hz. Pour la syllabe à 9.4 s, le F0 dépasse la valeur supérieure, ce qui se manifeste par des discontinuités à 9.33 s et à 9.51 s.



Dans ce cas il sera nécessaire d'ajuster la plage de fréquence du F0 en fonction de la voix du locuteur. Ci-dessous le résultat après ajustement de la plage de fréquence pour la détection du fondamental.



Pour obtenir le résultat souhaité, d'abord recalculer le F0 (Task = « Recalculate Pitch for entire sound file ») avec les valeurs minimale et maximale nouvelles (« F0 detection range »), puis recalculer le prosogramme (Task = « Prosogram and Prosodic profile »).

Parties inutilisables aux bouts du signal de parole

La segmentation est basée sur la courbe d'intensité. Comme celle-ci est calculée sur une fenêtre temporelle d'une certaine durée, elle est non définie aux deux bouts (environ 35 ms) du signal de parole. Ces parties ne seront pas donc disponibles pour l'analyse. Dans les cas où on aimerait les inclure, il suffit d'insérer au début et à la fin du signal un silence d'environ 50ms.