

UNIVERSITÉ DE GENÈVE
FACULTÉ DES LETTRES
DÉPARTEMENT DE LINGUISTIQUE
(Informatique pour les sciences humaines) /
ÉCOLE DE LANGUE ET DE CIVILISATION FRANÇAISES

Rapport sur le projet de NTIC

***" Système interactif de correction de la prosodie dans
l'apprentissage d'une langue seconde "***

Chargé d'enseignement: M.°Luka°Nerima

Chargée d'enseignement: Mme Isabelle Racine

Chargé d'enseignement: M. Jean-Philippe Goldman

Etudiant de Master en FLE :

Bolychev Dmitry

22, ch. François-Lehmann

1218 Le Grand Saconnex

E-Mail : bolyche0@etu.unige.ch

No. d'immatriculation : 09-328-337

Sommaire

- 1. Introduction**
- 2. La prosodie et les paramètres choisis (théorie)**
- 3. Présentation des outils NTIC utilisés et justification du choix**
- 4. Une description des tests effectués, des résultats obtenus (la progression du travail)**
- 5. L'interface utilisateur**
- 6. Les conclusions**
- 7. Bibliographie**
- 8. Annexes**

1. Introduction

Dans le cadre du projet de Nouvelles technologies de l'information et de la communication (NTIC), on a commencé à développer un *Système interactif de correction de la prosodie dans l'apprentissage d'une langue seconde*.

Une fois réalisé, on présuppose que ce système devra permettre à corriger la prosodie dans l'apprentissage d'une langue seconde. Le principe de fonctionnement de ce programme est le suivant^o: le système doit contenir un certain corpus des modèles prosodiques jugées d'être standards, ce qui servira de base pour l'évaluation de la prosodie de l'apprenant. Un modèle prosodique est un enregistrement audio d'un locuteur natif. Le programme doit pouvoir visualiser la prosodie de ces sons grâce aux telles technologies WEB comme le SVG, le PHP, le JavaScript et le (x)HTML dont le choix sera expliqué plus tard dans un chapitre relatif. Une autre fonctionnalité du système devra enregistrer en direct l'apprenant, notamment sa lecture/prononciation de mêmes textes qui ont été utilisés pour les modèles. Ensuite, la machine compare l'enregistrement de l'apprenant et montre où est le problème, si, bien sûr, il y en a un.

Le projet de NTIC prévoit la création d'un prototype interactif de ce programme qui sera capable de coller le texte à une courbe intonative à partir d'un document *"txt"* qui contient l'analyse d'un enregistrement fait grâce au logiciel *Praat*¹. Le prototype doit aussi réguler d'autres paramètres prosodiques dont la description sera donnée dans le chapitre suivant.

L'idée de la création d'un tel système automatique de "curvilignisation" des textes est venue d'un article de Rude Markus (2012). Dans cet article l'auteur décrit le concept de "*l'écriture prosodique*" (l'EP) (la traduction de l'auteur de l'anglais^o: "*prosodic writing*") (cf. Annexe 1) qui exprime les paramètres prosodique par les distorsions définies d'un texte écrit (Rude, 2012). Selon les conclusions faites dans

¹ "*Praat* est un logiciel libre pour l'analyse, la manipulation et l'annotation de sons." (Goldman Jean-Philippe, Décembre 2006)

cet article, toutes ces distorsions de texte dessinées à la main ont influencé positivement l'acquisition d'une bonne intonation par les apprenants.

2. La prosodie et les paramètres choisis (théorie)

Pour comprendre comment fonctionne le système, nous sommes obligés de passer par une base théorique, notamment, en expliquant le terme de la prosodie et les paramètres qui ont été choisis pour le projet.

"*La prosodie* (aspects supra-segmentaux) étudie une dimension qui vient s'ajouter par-dessus les sons du langage" (Racine, 2010-2011).

En français, la prosodie peut être représentée par les paramètres suivants:

- a. "*Fréquence fondamentale* (désigne l'estimation du son laryngien à partir du signal acoustique à un instant donné),
- b. de *durée* (mesure d'un intervalle de temps nécessaire pour émettre le signal de parole ; elle concerne donc l'organisation temporelle du message et comprend
 - le *débit de parole* (nombres de syllabes par seconde),
 - *le tempo* (accélération ou ralentissement du débit à l'intérieur d'un groupe prosodique)
 - et *les pauses*)
- c. et *d'intensité* (est relative à l'énergie contenue dans le signal de parole) " (Lacheret-Dujour et Beaugendre, 1999).

Dans le cadre de notre projet nous s'intéresseront à la fréquence fondamentale (f0) pour dessiner la courbe intonative, à la durée pour ajuster les dimensions des lettres dans des syllabes et, au lieu de l'intensité, nous prendrons la *proéminence* des syllabes pour marquer celles qui ont été considérées comme proéminentes de point de vue de la prosodie, ce qu'on définit en se basant sur 4 paramètres prosodiques^o: la durée relative, la f0 relative, le mouvement de la f0 et la durée de la pause (Goldman, Avanzi, Auchlin, Simon, 2012). Une étude réussie sur la détection automatique de la proéminence menée par M. Goldman, M. Avanzi,

M. Auchlin et Mme Simon (2012) a permis de distinguer 5 degrés de proéminence (de 0 (une syllabe non-proéminente) à 4 (une syllabe extrêmement proéminente)). Ces degrés sont donc utilisées dans notre système pour évaluer la proéminence des syllabes.

Comme le résultat, on a créé un fichier TXT (cf. Annexe 2) qui est une analyse de l'enregistrement de la lecture des fragments tirés de *Le petit Prince d'Antoine de Saint-Exupéry*. Ce fichier contient six colonnes suivantes°:

1-ère – les syllabes/les étiquettes (écritures orthographiques), qui on été écrites avec son système spécifique de notation: si l'étiquette est un trait c'est une pause; si elle finit par un tiret (ex: con-), cela indique qu'elle devra être jointe avec la syllabe suivante; si elle ne finit pas avec un tiret, il faut rajouter un espace; s'il y a un espace a l'intérieur de l'étiquette, il faut le laisser ce qui montre qu'il y a eu l'enchaînement ou la liaison (comme dans "chez les hommes" → chez – les –z – hommes);

2-ème et 3-ème – le début et la fin (en secondes) respectivement;

4-ème – la fréquence fondamentale (f0 ou en anglais "pitch") en Hz;

5-ème – l'intensité en Db;

6-ème – la proéminence (de 0 (une syllabe non-proéminente) à 4 (une syllabe extrêmement proéminente)).

3. Présentation des outils NTIC utilisés et justification du choix

Comme on a déjà remarqué avant, ce programme est écrit majoritairement en PHP qui génère le SVG dans un encadrement de (x)HTML et un ornement de CSS et utilise un peu le JavaScript.

Le SVG (Scalable Vector Graphics) joue le rôle principal dans le système parce que, grâce à ce langage, on fait les distorsions, manipulations visuelles avec

le texte. Etant un langage XML, le SVG était créé pour faire les graphiques vectoriels adaptables de deux dimensions. Les graphiques *vectoriels* se distinguent de les graphiques *matriciels* par sa capacité d'être extensible à l'infini, d'où son nom "adaptable". Les créateurs des cartes géographiques ont apprivoisé cette technologie car la qualité des cartes en SVG est infiniment mieux que de celles du format *PNG*, par exemple. Une telle résolution est possible dans une image SVG, grâce à sa composition. L'image matricielle est liée aux pixels tandis que l'image SVG se construit à partir de différentes données et paramètres qu'on donne dans le code SVG. Cependant, supporté complètement par le Chrome, moins par le Firefox, le Safari, and l'Opera, le SVG n'est pas la bienvenue sur Internet Explorer. Pour visualiser nos résultats, on utilise donc le Google Chrome.

Le langage PHP a été choisi pour ces capacités. Le PHP (*Hypertext Preprocessor*) est un langage de coté serveur que l'utilisateur ne voit pas mais il peut percevoir seulement le résultat de ce langage. Cependant, ce dernier permet l'interaction avec l'utilisateur, par exemple, à travers de renvoie des requêtes à travers les formulaires qui ont apparues déjà dans la deuxième version de PHP² en 1995, un an après son invention par Rasmus Lerdorf³. Après chaque telle interaction le PHP génère le nouveau page WEB, le PHP serve donc à créer des pages WEB dynamiques. Une autre capacité de PHP qui nous est intéressante est d'interpréter des fichiers TXT. Nous profitons de cette fonctionnalité pour lire le fichier contenant l'analyse d'un son à fin d'entrer les données nécessaires pour leur visualisation prochaine. Enfin, le PHP se conjugue bien avec le (x)HTML et permet même de générer le (x)HTML mais aussi le SVG. Comme c'est un langage de coté serveur, il faut avoir un serveur sur lequel les fichiers PHP doit être posés.

Enfin, le CSS fait une mise en page de l'interface pour qu'il soit plus au moins esthétiquement confortable de travailler avec la page WEB.

² Le 14 mars 2013, la version 5.4.13 de PHP est sorti (<http://php.net>)

³ Rasmus Lerdorf – est un programmeur danois et canadien de 44 ans (http://fr.wikipedia.org/wiki/Rasmus_Lerdorf)

4. Une description des tests effectués, des résultats obtenus (la progression du travail)

Notre travail a commencé à partir d'étude approfondie du langage SVG et, notamment, de son traitement de texte qui mit en disposition une large palette des possibilités du changement visuel du texte. Ci-dessous, nous expliqueront quelques principes de base que l'on a utilisés pour notre projet.

En gros, le code SVG doit être entre les balises <svg> dans laquelle on définit les dimensions du cadre de l'affichage, sa position, la possibilité de redimensionner l'image ou pas, etc. La deuxième balise importante est <defs> qui est responsable de toutes les définitions, dans notre cas, de la définition du *path*, du chemin par lequel le texte passera. Enfin, le <g> signifie le groupe sur lequel toutes les définitions seront exécutées et, plus précisément, sur notre texte. Ce dernier doit être imbriqué dans la balise <text> et après dans <textPath> et, ensuite, il peut être déchiré par <tspan> qui contient principalement des attributs de style.

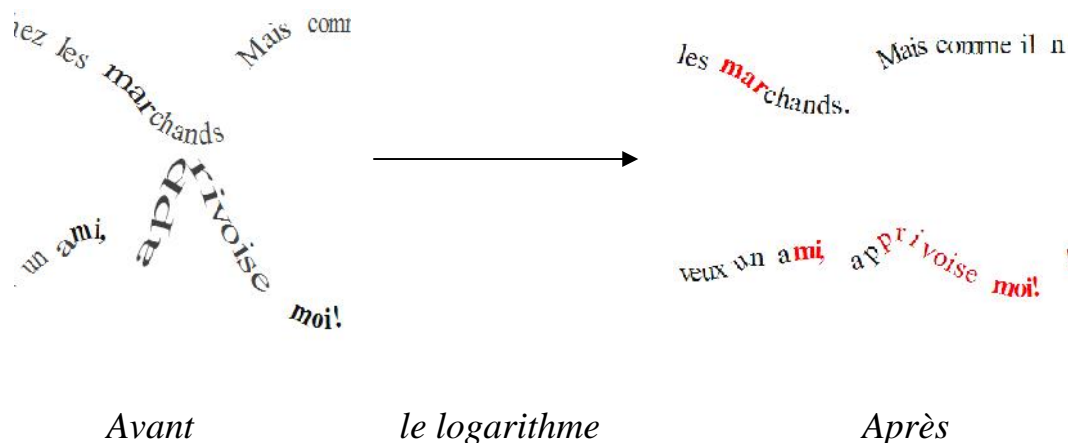
Une fois que l'on a suffisamment étudié les spécifications du SVG, on a commencé à les utiliser en créant un code manuellement (cf. Annexe 3) pour, après, pouvoir recréer de la manière automatique la même image et en l'améliorant avec du PHP.

Les tests ont été nombreux dont la preuve est la mini carte d'évolution du projet représentée dans l'Annexe 3. A savoir, on peut les regrouper de la manière suivante:

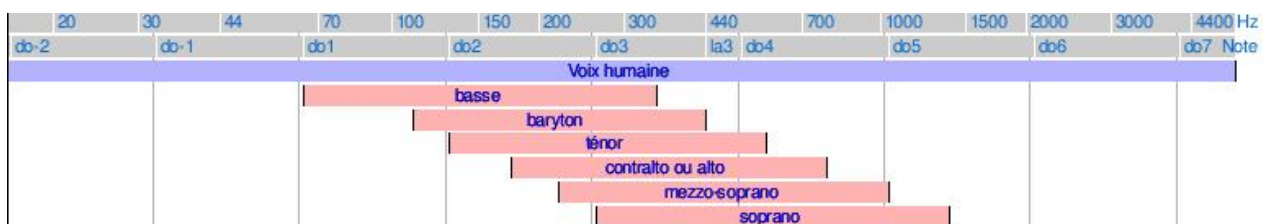
- ***Le chemin textuel.*** On a fait plusieurs tentatives de bien représenter le chemin sur lequel le texte doit être collé. Le problème principal concernait le choix des bons attributs lors de la définition du chemin dans le *path*. Parce que chaque syllabe doit être sur sa place conformément aux paramètres associés à cette syllabe. Comme on le voit sur les images des premières tentatives de la création automatique, cela n'était pas une mince affaire car, au début, on a réalisé un chemin presque sans aucune rupture et on a placé le texte entier dans une seule balise

<text>. Une solution a été trouvée qui consistait en placement de chaque syllabe dans les trois balises destinées à recevoir le texte (<text><textPath><tspan>) à fin de recevoir une maîtrise complète sur chaque syllabe. On a donc dû modifier également le chemin en le déchirant en petits morceaux chacun renvoyant à sa syllabe. La version automatique avec la courbe "rectiligne-curviligne" a apparue grâce à cette technique. Enfin, on a calculé les milieux de chaque but du path et on a effectué un autre changement, en déclarant le milieu du chemin présent comme son début et le milieu du chemin prochain comme sa fin, ce qui a lié les chemins des syllabes entre eux en créant une ligne continue du texte et a placé chaque syllabe sur sa place.

Dans une autre étape du développement du programme, on a remarqué que quelques syllabes montaient brusquement trop haut ou descendaient de la même façon trop bas°:



La solution de ce problème était dans les différences des voix humaines dont l'intensité varie d'une voix à l'autre, par exemple pour le diapason du baryton est de 110 à 440 Hz (330Hz) pour le soprano de 250 à 1500 Hz (1250Hz).



Source: http://fr.wikipedia.org/wiki/Registre_%28musique%29

Du coup, on a ajouté un logarithme naturel comme une condition pour traitement de l'intensité. Après quelques changements prochains dans l'algorithme de la création du chemin, on a réussi de recevoir le résultat final de la ligne surnommée "Curviligne".

Cependant, il existe une autre représentation possible comme "les créneaux". La combinaison de ces deux représentations montrées dans le rapport présent avec l'affichage d'autres paramètres utilisés dans notre projet multiplie le nombre des visualisations possibles. C'est pourquoi on a décidé d'en créer une liste (cf. Annexe 4) qui représente une mini-recherche sur le type de chemin car on a dû toujours suivre l'idée que le texte affiché doit être bien lisible et aider les apprenants à avoir une bonne prosodie. Cette mini-recherche après des enseignants de notre faculté a montré la nécessité du développement d'une interface interactive pour percevoir mieux toutes les possibilités possibles de la représentation prosodique d'un texte à fin d'en choisir les meilleures.

5. L'interface utilisateur

Notre interface dispose des paramètres que l'utilisateur peut changer et voir directement les résultats de ces manipulations, à savoir, une nouvelle présentation du texte (cf. Annexe 6). Dans la partie interactive de l'interface l'utilisateur peut changer la mode générale de la représentation (créneaux / curviligne), le texte à affiché, le fréquence fondamentale, la longueur des lignes, la largeur des lettres, la distance entre els lignes, de visualiser/cacher la courbe, la proéminence ou la durée. Une partie avec du texte affiche tous ces changements.

Cette interface est destinée à trouver la(les) meilleure(s) forme(s) de la représentation de la courbe intonative et du texte sur cette courbe. Cette(Ces) forme(s)-là doit être la(les) mieux adaptée(s) pour l'acquisition facile d'une bonne prosodie par l'apprenant. Pour atteindre ce but, cela sera important d'inviter les différent gens, experts ou apprenants ou l'autre, à répondre aux questions suivantes^o:

1. Est-ce que la visualisation de plusieurs de paramètre (par exemple: de la fréquence fondamentale (f_0), durée et proéminence) au même temps ne perturbe-t-elle pas la lecture? Si, oui, comment serait-il mieux de le faire? Ou quelle paramètre serait-il mieux d'afficher premièrement (au début du travail sur ce son), quelle – deuxièmement, troisièmement, etc.?
2. Pour faire une recherche plus poussée, dans la version prochaine, on envisage d'intégrer la police comme un paramètre changeant: la taille de police et le police lui-même. On propose d'ajouter le Times New Roman et Comic Sans MS.
 - o A Votre avis, quels polices serait-il mieux d'ajouter au teste?
 - o Quels d'autres paramètres faut-il ajouter pour une bonne recherche ou peut-être éliminer quelques uns déjà existants?

A cette étape-là, il faut trouver également le meilleur tableau des paramètres pour après présenter ce projet à un nombre plus grand des gens (pour faire une recherche plus globale). En plus, accuellement notre programme a quelques *limitations*°:

1. si on mit le paramètre "Largeur des lignes" de 0-370 et de 710-750, le texte ne passe correctement pas sur la nouvelle ligne;
2. le même problème du passage incorrect à la nouvelle ligne est constaté si la valeur du paramètre "Largeur des lettres" est trop grande et la largeur des lignes est trop petite;
3. la durée n'est pas assez beau et correctement visualisée;
4. Pour un bon travail, il faut ajuster la résolution de la fenêtre à 200%.

6. Les conclusions

Le prototype développé a déjà servi l'objet d'une mini-étude, décrit en partie dans le chapitre précédent, sur la représentation de son interface parmi les experts et quelques suggestions ont été reçues qui sont présentées ci-dessous°:

- Les caractères par défaut sont très petits, il faudrait les agrandir;
- Le bouton pour revenir aux valeurs par défaut est nécessaire car on n'arrive plus à revenir à quelque chose de lisible après avoir manipulé les paramètres;
- Le contraste entre les 2 rouges n'est pas assez flagrant. Il faut ajouter une option pour grouper 2 niveaux de prééminence 0-1-2 vs. 3-4;
- La désactivation de la durée ne change que la largeur des lettres. Il faudrait rapprocher les dernières.

Les résultats de ce projet de NTIC seront très utiles pour un développement futur du système lors d'un mémoire de Master où on doit réussir à°:

- Ajouter la génération d'un fichier TXT directement de l'interface;
- Ensuite, il faudrait ajouter une possibilité d'enregistrer la voix et d'en générer le fichier TXT;
- ajouter la comparaison du son enregistré avec le modèle et, si nécessaire, de pointer le problème et donner les recommandations d'amélioration, etc.

Ce projet ouvre sa nouvelle facette - la psycho-visuelle qui consistera à une recherche prochaine sur l'efficacité de cette visualisation du texte, l'efficacité de ce concept pour la correction de la prosodie. Et on espère que ce programme sera utile pour les apprenants. Mais, quand même, on a déjà réussi à produire quelques œuvres **d'art proso-SVGieniste** (cf. Annexe 7).

7. Bibliographie

La prosodie

1. Goldman Jean-Philippe, *Tutoriel Praat*, Université de Genève, Décembre 2006, <http://latlntic.unige.ch/phonetique/easyalign/tutorielpraat.pdf>;
2. Goldman Jean-Philippe, Avanzi Mathieu, Auchlin Antoine, Simon Anne Catherine, *A Continuous Prominence Score Based on Acoustic Features*, InterSpeech 2012 (Portland, Oregon, du 09/09/2012 au 13/09/2012);
3. Lacheret-Dujour Anne et Beaugendre Frédéric, *La Prosodie du français*, 1999, Editions du CNRS;
4. Racine Isabelle, *Polycopies du cours: Phonétique et phonologie du français*, ELCF: BA1/DE6, Année 2010-2011;
5. Rude Markus, *Native-like Duration Ratio of Stressed vs. Unstressed Syllables through Visualizing Prosody*, Speech Prosody, may 2012.

La programmation

1. svgbasics.com/index.html, *Tutoriel basique de la programmation en SVG*;
2. [fr.wikipedia.org, SVG, http://fr.wikipedia.org/wiki/Scalable_Vector_Graphics/](http://fr.wikipedia.org/wiki/Scalable_Vector_Graphics/);
3. [www.w3.org, SVG, http://www.w3.org/TR/SVG/](http://www.w3.org/TR/SVG/);
4. www.w3.org/TR/SVG/text.html, *Traitement de texte en SVG*;
5. [php.net, PHP, http://www.php.net/manual/fr/intro-what-is.php](http://www.php.net/manual/fr/intro-what-is.php).

Le lien vers notre projet

<http://latlntic.unige.ch/~bolyche0/Projet/prototypeCurviligne.php>

8. Annexes

Annexe 1

Ich wohne in einer **Altbauwohnung**, **Meine Wohnung ist groß**
und **hell** und sie **liegt zentral!**

Source: Rude, 2012

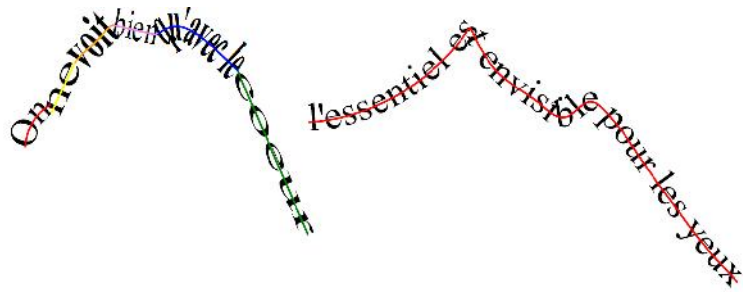
Annexe 2

Une partie du fichier TXT utilisé

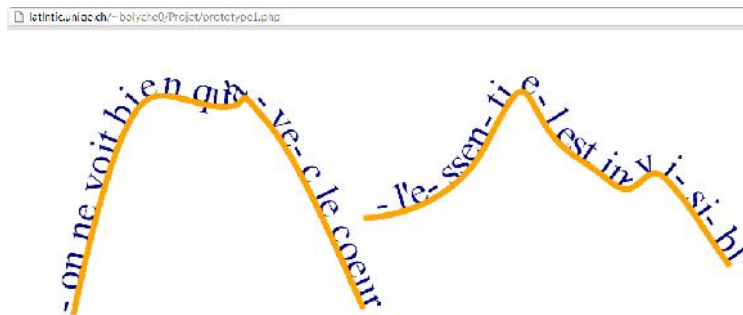
-	0	0.245	235	43	
On	0.245	0.355	218	79	0
ne	0.355	0.475	257	83	0
con-	0.475	0.635	285	82	1
nait	0.635	0.785	312	84	1
que	0.785	0.915	325	78	1
les	0.915	1.025	268	81	0
choses	1.025	1.335	261	78	3
que	1.335	1.455	252	76	0
l'o-	1.455	1.545	199	72	0
n ap-	1.545	1.665	241	76	0
pri-	1.665	1.895	228	74	0
voise,	1.895	2.255	191	72	2
dit	2.255	2.475	223	74	0
le	2.475	2.575	221	77	0
re-	2.575	2.705	211	72	0
nard	2.705	2.965	157	68	4
-	2.965	3.656	190	43	

Annexe 3 évolution du projet

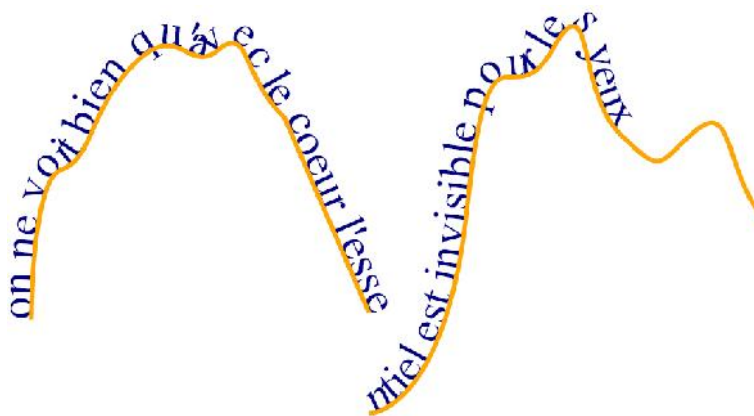
La version faite à la main



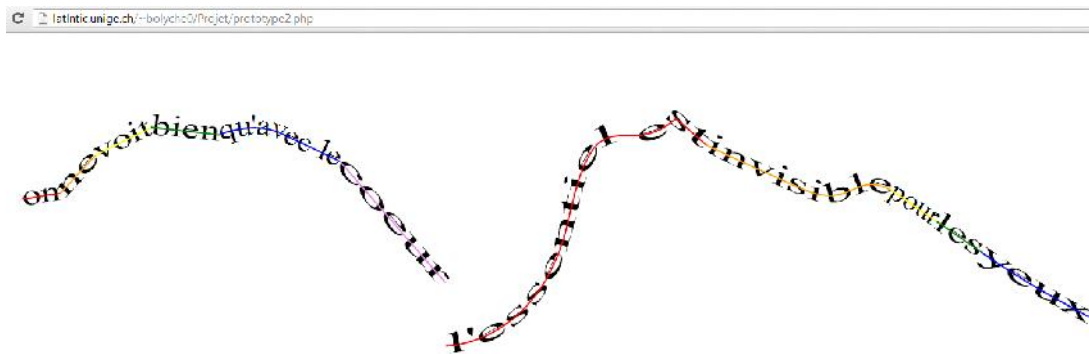
La première version automatique avec des tirets



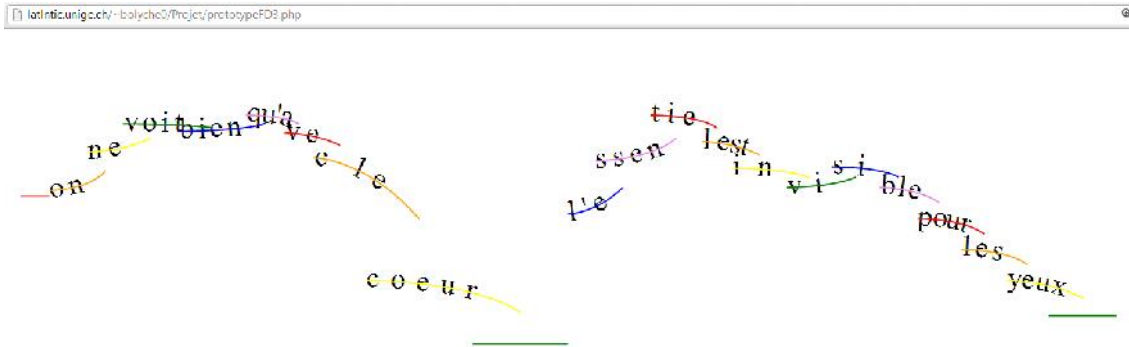
La version automatique sans les tirets avec la courbe continue



La version automatique avec la courbe rectiligne-curviligne



La version automatique avec la courbe "curviligne-créneaux"



Une des versions finales



Annexe 4 mini-recherche

La proéminence

On ne connaît que les choses que l'on **apprivoise**, dit le **renard**. Les hommes n'ont plus le temps de rien **connaître**.

Ils **achètent** des choses toutes faites chez les **marchands**. Mais comme il **n'existe** point de **marchands d'amis**, les hommes n'ont plus **d'amis**. Si tu veux un **ami**, **apprivoise moi!** Adieu! dit le **Renard**. Voici mon secret.

Il est **tres** simple: on ne voit bien qu'avec le **coeur**.

La durée et la proéminence

On ne **connaît** que **les** choses que l'on **apprivoise**, dit le **renard**. Les hommes n'ont plus le temps de **rien connaître**. Ils **achètent** des choses toutes faites chez les **marchands**. Mais comme **il n'existe** point de **marchands d'amis**, les hommes n'ont plus **d'amis**. Si tu veux un **ami**, **apprivoise moi!** **Adieu!** dit le **Renard**. Voici mon **secret**. Il est **tres** simple: on ne voit bien qu'avec le **coeur**.

"Créneaux"

Fréquence fondamentale et la durée

on ne voit bien qu'a
c le ve
l'est in vi si ble
l'e
pour les yeux
coeur

Fréquence fondamentale et la proéminence

chettent toutes
Ils a des choses faites chez les mar Mais comme il n'existe point de marchands d'a
chands pri
mis, les hommes n'ont plus d'a Si tu veux un a mi, ap voise A dicu! dit le Renard Voi cret

Fréquence fondamentale et la durée et la proéminence

On ne connaît que les choses que l'on appri voise, dit le re Les hommes n'ont plus le temps de
ard
chettent toutes
Ils a des choses faites chez les mar Mais comme il n'existe point de
chands
maître

"Curviligne"

Fréquence fondamentale

on ne voit bien qu'avec le coeur l'essentiel est invisible pour les yeux

Fréquence fondamentale et la durée

on ne voit bien qu'avec le coeur l'essentiel est invisible pour les yeux

Fréquence fondamentale et la proéminence, version 1

On ne connaît que les choses que l'on apprivoise, dit le renard. Les hommes n'ont plus le temps de rien connaître

Ils achètent des choses toutes faites chez les marchands. Mais comme il n'existe point de marchands d'a-

mis, les hommes n'ont plus d'amis. Si tu veux un ami, apprivoise moi'. Adieu! dit le Renard. Voici mon secret

Il est très simple: on ne voit bien qu'avec le coeur

Fréquence fondamentale et la proéminence, version 2

On ne connaît que les choses que l'on apprivoise, dit le renard. Les hommes n'ont plus le temps de rien connaître.

Ils achètent des choses toutes faites chez les marchands. Mais comme il n'existe point de marchands d'amis, les hommes n'ont plus d'amis.

Si tu veux un ami, apprivoise moi! Adieu! dit le Renard. Voici mon secret.

Il est très simple: on ne voit bien qu'avec le coeur.

Annexe 6

L'interface de l'utilisateur

Prosodie en SVG Curvilinear

latintic.unige.ch/~bolyche0/Prjct/prototypeCourv_FL.php?TexteAfficheTemp=texte1&FTemp=30&largeurTemp=900&longueurLettreTemp=10&strokeTemp=Non&prominTemp=non&dureeTemp=spacing&distan

Afficher en créneaux Avant de commencer, lisez, s'il vous plaît, les consignes de recherche

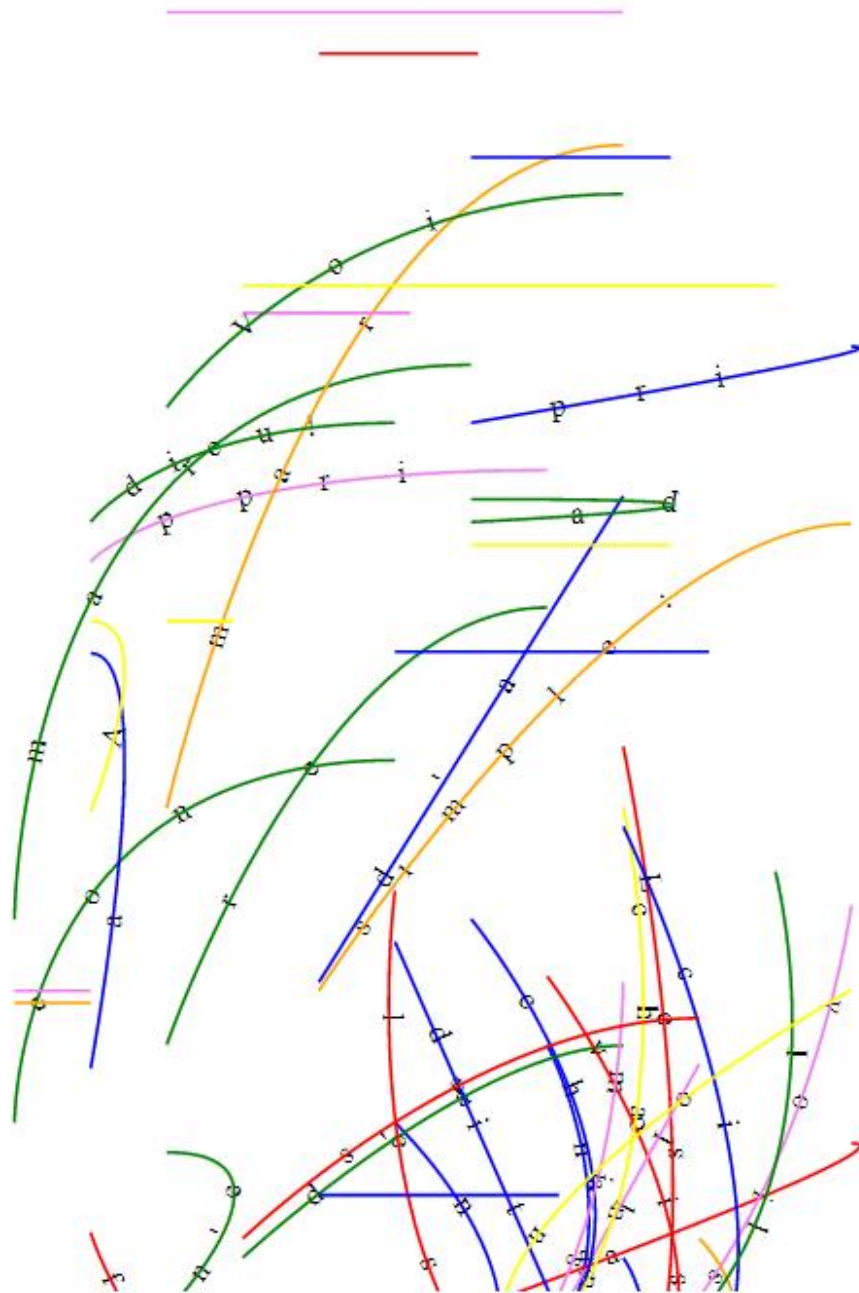
	Texte à affiché	F0 (pitch)	Longueur des lignes	Largeur des lettres	Voir Courbe	Proéminence	Durée	Distance entre les lignes
Paramètres à manipuler	Choisissez un texte	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="radio"/> Oui <input type="radio"/> Non	<input type="radio"/> Rouge <input type="radio"/> Noir-gris <input type="radio"/> Non	<input type="radio"/> Oui <input type="radio"/> Non	<input type="text"/>
Paramètres actuels, à ne pas changer	texte1	30	900	10	Non	non	spacing	150

Afficher

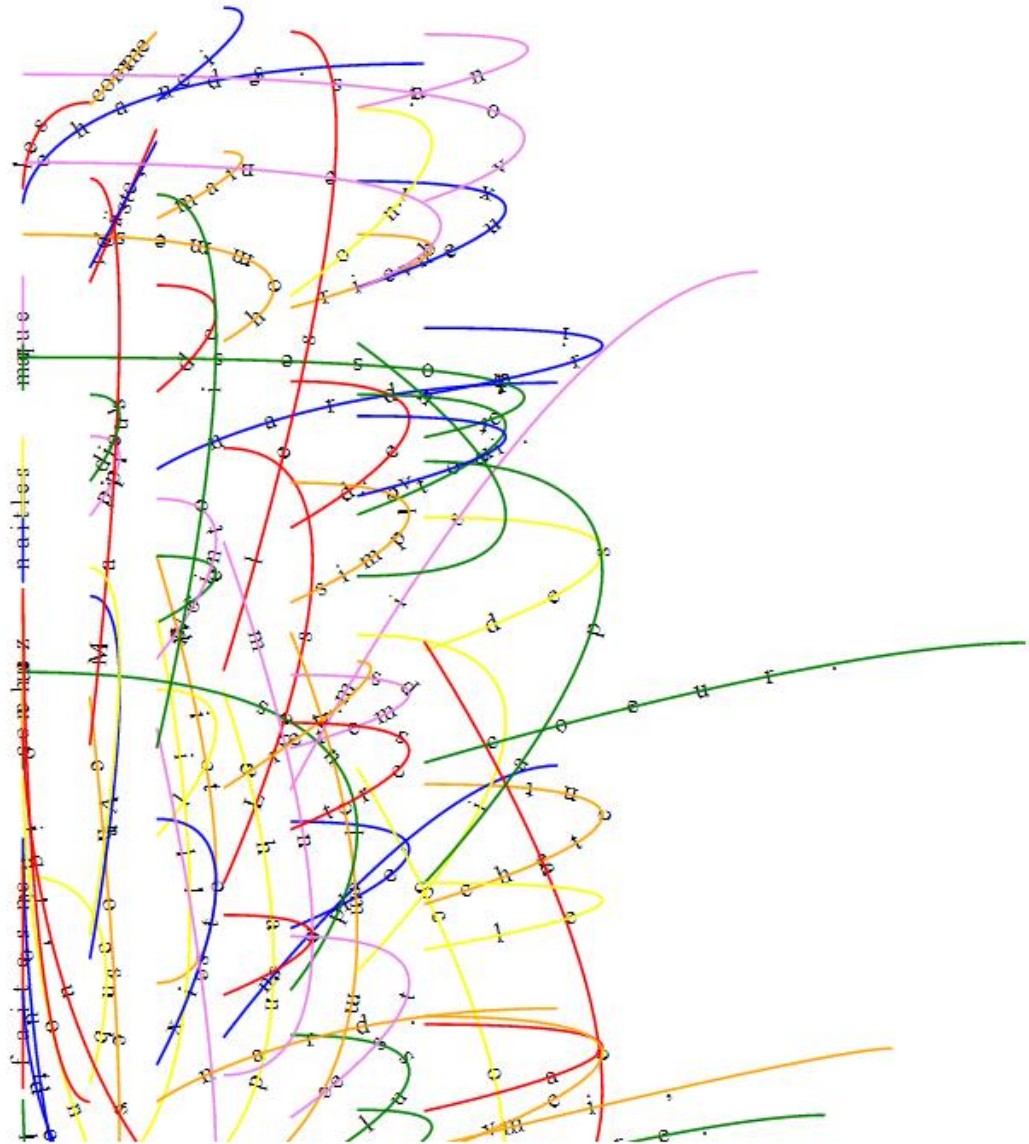
onne voit bien qu'avec le coeur l'essentiel est invisible pour les yeux

Annexe 7

L'art proso-SVGieniste



Expression régulière pour les chiffres doit avoir "`\W`"



Une belle erreur de la saute à la nouvelle ligne

