

## LA SYLLABATION AUTOMATIQUE DE SÉQUENCES PHONÉMIQUES EN ANGLAIS – ALGORITHME, APPLICATIONS ET OUTILS

*Roussi Nikolov*

*Université de Plovdiv Païssii Hilendarski*

*Elise Ryst*

*Université Paris 8*

The presented new algorithm of automatic syllabification allows insertion of dots as syllable boundaries in phoneme sequences in English, in strict accordance with the rules applied, for example, in the *Cambridge Advanced Learner's Dictionary*. The algorithm has already been integrated into several projects, most prominently into *SARP*, into the Arpabet to IPA conversion of the *CMU Pronouncing Dictionary* and into a new computational method to evaluate the phonological distance between words, more consistent with the perceptual data and applied in the computer program *TREFL* for automatic generation of didactically useful associations between words.

**Key words:** automatic syllabification, english phonological distance

### **Introduction**

Contrairement à la syllabe, la syllabation automatique, au sens technique du terme, est facile à définir. Ce terme désigne la segmentation informatisée en syllabes d'une représentation (acoustique ou graphique) de la parole. La réalité acoustique des frontières syllabiques étant généralement contestée, la question d'une division automatique en syllabes du signal sonore ne se pose même pas actuellement. Quant à la forme orthographique des mots, la modélisation de la syllabation est peu intéressante intellectuellement en raison de l'arbitraire de l'écriture alphabétique. Les syllabes orthographiques, qui ne coïncident pas toujours avec les syllabes de prononciation (*syl.la.ble* vs /'sɪl.ə.bl/), sont notées, par exemple, dans le dictionnaire *Longman Dictionary of Contemporary English* ou encore dans le *Merriam Webster Dictionary*.

Ce travail traite le cas le plus simple, notamment celui de la représentation de la parole sous forme d'une suite de symboles phonémiques. La syllabation automatique d'une chaîne de phonèmes est par elle-même une tâche de programmation informatique, donc purement technique, puisqu'il s'agit simplement de « traduire » en langage informatique les règles phonologiques de syllabation pour une langue donnée. Nous ne partageons donc pas l'opinion selon laquelle « Automatic syllabification of words is challenging, not least because the syllable is not easy to define precisely » (Marchand & Al. 2009: 1). La syllabation automatique étant définie ici comme une sorte de traduction en langage informatique des règles de syllabation, dire que « la syllabation automatique pose des défis, en particulier parce qu'il n'est pas facile de définir la syllabe de manière précise » équivaut à dire, par parallélisme, qu'il est difficile de traduire la phrase, par exemple, « Je m'appelle Pierre. » parce que l'assertion n'est peut-être pas vraie.

C'est la syllabation qui pose des défis et non pas son automatiser. Pour autant, le rôle de la syllabation automatique, autant sur le plan théorique que pratique, n'est pas négligeable, ne serait-ce que parce qu'on connaît l'importance d'aborder une question, en l'occurrence la syllabe, sous de multiples angles<sup>1</sup>.

En effet, au-delà des applications pratiques, l'instrumentalisation des modèles formels de la syllabe sous forme d'outils d'automatisation de la syllabation devient autrement plus intéressante quand on considère l'économie relative de chaque algorithme comme une sorte de mesure de la rationalité du modèle de départ correspondant. Toutes choses égales par ailleurs, un modèle phonologique de la syllabe serait donc considéré comme étant d'autant plus fonctionnel qu'il se traduit informatiquement par une procédure de syllabation simple.

En outre, le modèle informatisé présenté ici relatif à la syllabation en anglais suggère quelques ajustements possibles des représentations phonologiques généralement adoptées, notamment concernant les notations /ɪ/ et /ʊ/ en syllabe ouverte non accentuée.

Initialement, ce travail a été motivé par une application pratique de la syllabation automatique dans le cadre d'un projet de conception de la similarité phonologique entre les mots anglais et de réalisation d'outils de l'évaluation automatique (Nikolov 2010, Ditchewa 2010). La représentation phonologique des mots anglais a été fournie par le

---

<sup>1</sup> „If you understand something in only one way, then you don't really understand it at all.“ – Marvin Minsky.

dictionnaire *the Carnegie Mellon Pronouncing Dictionary*. Afin de prendre en compte la similarité aussi bien de contenu que structurelle, cette représentation a été enrichie par la notation des syllabes.

### **Un aperçu sur la théorie de la syllabe et les règles de syllabation en anglais**

La syllabe, utile pour les descriptions linguistiques (beaucoup d'études en phonologie, en phonétique articulatoire et en psycholinguistique), reste une vaste question de discussion scientifique quant à sa nature, ses rôles et ses frontières. Il existe de nombreuses définitions de la syllabe, mais on leur trouve toujours *a posteriori* au moins une exception dans une langue du monde. Par exemple, pour de nombreux auteurs et pédagogues, toutes les syllabes de toutes les langues ont une voyelle pour noyau. Pour le français, la règle « une voyelle équivaut à une syllabe » fonctionne (Ditcheva, dans cette même revue, décrit les règles de syllabation du français). Pourtant, en anglais, il est attesté, qu'il y a des « consonnes syllabiques », c'est-à-dire des consonnes qui occupent la position de noyau, telles que /l/ et /n/ dans <bottle> et <button>, respectivement ; sans parler d'études sur le tachelhit berbère (Ridouane 2008) qui examinent des mots entiers composés uniquement de consonnes, sans aucune voyelle au niveau phonologique. Ainsi, ces consonnes syllabiques, indiquées dans les transcriptions par un petit trait en dessous de la consonne (ex : /l̩/), peuvent être traitées par le système de segmentation automatique comme s'il s'agissait de voyelles, pour qu'elles gardent leur comportement de noyau de syllabe dans la suite de règles adoptée.

En linguistique, aucune théorie de la syllabation n'est unanimement acceptée, qu'elle soit considérée comme *a priori* applicable à toutes les langues ou *simplement* à l'anglais. Certaines langues peuvent se contenter d'une règle ou deux, alors que l'anglais en nécessite souvent cinq ou plus. L'ajout d'exceptions à la règle principale permet ici de prendre en compte les particularités de la langue anglaise en gardant la même règle principale pour toutes les langues. Même si des approches compétitives se contredisent sur de nombreux points, Treiman et Danis (1988) proposent très raisonnablement trois axes généraux sur lesquels il y a consensus quelle que soit la langue : 1) chaque syllabe contient au moins une voyelle ou une consonne syllabique ; 2) la syllabation suit, le plus souvent, des règles phonotactiques : si une consonne ou une séquence de consonnes est illégale en début de mot, elle est aussi illégale en début de syllabe ; 3) il y a

en début de syllabe accentuée le plus possible de consonnes formant une suite phonotactiquement légale.

La plupart des approches linguistiques s'accordent sur la syllabation des séquences VC'V de l'anglais (où V=voyelle inaccentuée, C=consonne possible en position initiale de syllabe, et 'V=voyelle accentuée) pour les syllaber ainsi : V.C'V (frontière syllabique placée avant la consonne quand l'accent porte sur la dernière syllabe), par exemple <machine> /mə'ʃi:n/. Cependant, les linguistes sont en désaccord en ce qui concerne la syllabation des séquences 'VCV en anglais (où 'V =voyelle accentuée, C=consonne possible en position initiale ou finale de syllabe, et V=voyelle inaccentuée) où la séquence 'VCV peut être syllabée 'V.CV, 'VC.V ou 'VC.CV, en fonction des règles de syllabation adoptées :

« city »	:	ci-ty	ou	cit-y	ou	cit-ty
?						
C'VCV	:	C'V.CV	ou	C'VC.V	ou	
C'VC.CV	?					

Ainsi, ce sont bien les règles de syllabation qui sont ici difficiles à choisir.

Dans les années 1970-80, la segmentation des mots VCV en syllabes pouvait déjà dépendre, selon l'auteur, de différents critères comme la vitesse d'élocution, le contexte vocalique et accentuel (place de l'accent, qualité de la 1<sup>ère</sup> voyelle quand accentuée), du niveau linguistique (phonétique ou phonologique). Certains linguistes (comme Pulgram 1970), ont proposé des listes de règles régissant la syllabation des mots, sous forme de règles successives. Pour certains auteurs, la nature de la voyelle accentuée en anglais détermine la syllabation des séquences 'VCV. Si la voyelle accentuée d'une séquence 'VCV est longue et qu'elle peut être en position finale de syllabe, la consonne est alors attribuée à la syllabe qui la suit. Si la voyelle accentuée d'une séquence 'VCV est brève ou que la consonne qui la suit est illégale en position initiale de syllabe, la consonne est alors attribuée à la syllabe qui la précède. Kahn (1976) adopte la même approche, sauf qu'il marque aussi une différence en fonction du débit de parole : si le locuteur parle lentement, il syllabera tous les mots V.CV,

indépendamment de la place de l'accent ou de la qualité des voyelles. D'ailleurs, pour Hoard (1971), « toutes les séquences 'VCV sont syllabées 'V.CV » (ex : <ci-ty>).

A l'époque des premières études sur la syllabe, les linguistes n'avaient pas à leur disposition l'instrumentation et les outils informatiques que nous connaissons aujourd'hui. Pourtant aujourd'hui encore, divers principes de syllabation sont proposés dans des manuels d'anglais oral et dans les dictionnaires de prononciation. Souvent, les syllabations proposées par le dictionnaire *Longman* (2008, 3<sup>ème</sup> édition) sont différentes de celles proposées par le dictionnaire *Cambridge* (2011, 18<sup>ème</sup> édition), justement parce que les règles adoptées sont différentes, inférées soit prioritairement en fonction de l'orthographe et de la segmentation en morphèmes (ex : <help.ing> /<sup>h</sup>helpɪŋ/, *Longman*) ou en fonction de règles phonologiques, par exemple de sonorité croissante en début de syllabe (ex : <hel.ping> /<sup>h</sup>hel.pɪŋ/, *Cambridge*). Il n'y a pas encore, à notre connaissance, de dictionnaire de prononciation, qui présente une segmentation syllabique présentant une transcription phonologique impliquant une consonne ambisyllabique, même si à l'écrit beaucoup de dictionnaires découpent les mots à consonne doublée entre les deux (ex : <sil-ly>, <diz-zy>) ; probablement car pour beaucoup de phonologues, chaque phonème ne doit occuper qu'une position syllabique et chaque frontière syllabique ne doit intervenir qu'entre deux phonèmes. Le système de segmentation automatique en syllabes présenté ici permet de présenter des découpes alternatives et donc, si cela est souhaité, une segmentation incluant le marquage de consonnes ambisyllabiques.

Parmi les différentes approches de syllabation, il est récurrent de faire appel à des règles phonologiques. La notion de sonorité pour définir la syllabe, remonte à peu près au début du 20<sup>ème</sup> siècle, avec Saussure (1916, « sonority scale ») notamment. La syllabe est alors décrite comme une enveloppe de sonorité croissante puis décroissante. La question de la sonorité croissante ou des pics de sonorité dans les principes de syllabation a ensuite été réadaptée et rediscutée dans de nombreux travaux, dont ceux de Clements (1990, « sonority cycle »). Il est reconnu que le principe de sonorité s'applique, *a priori*, à toutes les syllabes de toutes les langues du monde, mais rares sont les langues qui l'appliquent dans tous les contextes. Pour les mots en anglais comprenant plusieurs consonnes en position intervocalique, comme <dolphin>, la règle classique de sonorité croissante suffit généralement à leur syllabation (liquide de sonorité supérieure à la fricative qui suit : première sonorité croissante juste avant la fricative

<▲dol▲phin▲> /<sup>l</sup>dɒl.fɪn/). Les cas les plus problématiques de mots à plusieurs consonnes intervocaliques à syllaber en anglais sont les mots ayant une fricative /s/ suivie d'une occlusive sourde, /p/, /t/ ou /k/ dans l'agrégat consonantique où /s/ est parfois appelé « extra-syllabique », par exemple <dusky> /<sup>l</sup>dʌski/. Le problème est résolu ici par l'adoption de règles telles que l'attribution d'une position spéciale dans l'échelle de sonorité pour le /s/, analysé alors comme de moindre sonorité que l'ensemble des fricatives et occlusives.

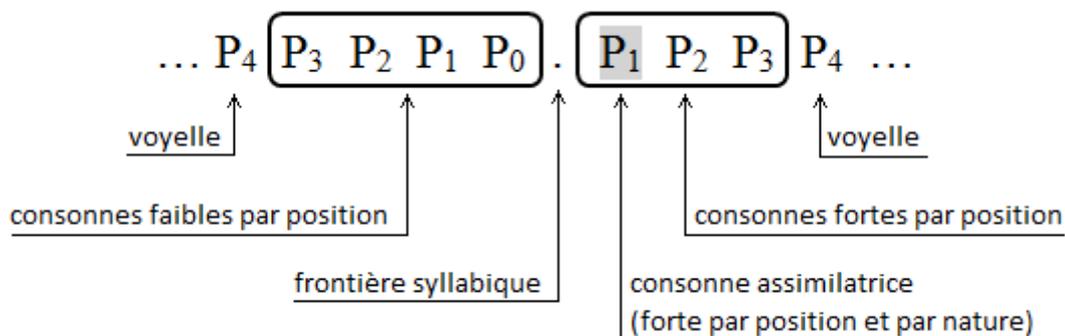
La syllabe étant le plus souvent définie comme unité phonologique, il est plutôt rare qu'on lui reconnaisse une existence phonétique. Des outils utilisés en phonétique (dont *PRAAT*) permettent déjà une syllabation automatique de transcriptions phonologiques (pouvant elles-mêmes être automatisées) pour plusieurs langues dont le français, mais ceux-ci ne sont pas encore efficaces pour l'anglais et demandent à l'utilisateur d'être familiarisé avec l'application de scripts. Par conséquent, l'originalité de ce travail porte davantage ici sur le fait de révéler, au travers d'un outil ergonomique, des règles de syllabation cohérentes et économiques applicables à autant de langues que possible dont l'anglais.

### **Règle de base et ajustements phonotactiques**

Pour toute séquence intervocalique (les consonnes syllabiques étant considérées comme des voyelles), cette méthode de syllabation automatique consiste en deux procédures successives : 1) remplacement des phonèmes dans une structure linéaire, c'est-à-dire dans une suite de positions dont chacune est associée à des caractéristiques segmentales prédéfinies et où la frontière syllabique est déterminée aussi, et 2) prise en considération des contraintes phonotactiques propres à l'anglais.

On voit apparaître une régularité stricte suite à l'application de la deuxième procédure – chaque fois que la prise en considération des contraintes phonotactiques remet en cause la règle de base, la frontière syllabique se déplace à droite d'un phonème.

La structure intervocalique est basée sur un modèle des lois de position dans la chaîne segmentale en français :



**Fig. 1.** *Modèle des séquences intervocaliques permises en français*

Cette suite de positions est une généralisation des séquences intervocaliques possibles en français, où chaque position est associée à des caractéristiques segmentales particulières (intrinsèques : tableau 1, et positionnelles : consonnes faibles ou fortes, assimilatrices ou assimilables). Le point entre P<sub>0</sub> et P<sub>1</sub> représente la frontière syllabique. L'indice de position indique la catégorie de phonèmes susceptibles d'occuper la position correspondante. Le remplacement d'une séquence intervocalique doit se faire de droite à gauche, c'est-à-dire qu'en procédant successivement de droite à gauche, chaque consonne occupera la position correspondante le plus à droite ; certaines positions resteront vides, en ce sens qu'elles ne seront associées à aucune consonne.

**Tableau 1 :** Restrictions de contenu des positions segmentales

Position	Phonèmes
P <sub>0</sub>	la fricative /s/
P <sub>1</sub>	les bruyantes (les occlusives et les fricatives)
P <sub>2</sub>	les consonnes nasales et les liquides
P <sub>3</sub>	les semi-consonnes
P <sub>4</sub>	les voyelles

On peut envisager la séquence P<sub>4</sub> P<sub>3</sub> P<sub>2</sub> P<sub>1</sub> P<sub>0</sub> P<sub>1</sub> P<sub>2</sub> P<sub>3</sub> P<sub>4</sub> comme l'écriture concise d'un ensemble de chaînes de phonèmes, plus spécialement l'ensemble des séquences intervocaliques possibles. Or, *l'écriture concise* [grâce à l'utilisation de caractères génériques et opérateurs] *d'un ensemble de chaînes* [de caractères] est la définition précise de ce qu'on appelle en informatique une expression régulière (Wall & Al. 2001: 31). Le langage des expressions régulières, méthode puissante

et efficace de traitement du texte, permettra donc de donner une expression informatique immédiate du modèle phonologique représenté sur la fig. 1.

Ce modèle s'adapte facilement à la variabilité de syllabation en français (par exemple [vɛs.tiz] ou [ve.stiz] ; [tɛk.nik] ou [te.knik]) et, comme on le verra, aux contraintes phonotactiques en anglais. Sa formalisation informatique permet de réaliser un outil de syllabation automatique souple et extrêmement simple. En effet, le code source du programme de syllabation automatique en français (voir ci-dessous) ne contient en fait qu'une ligne de code, sans compter les deux instructions relatives à l'entrée et la sortie des données (la chaîne de caractères), reproduisant très strictement le modèle phonologique. (L'utilisation des variables statiques P<sub>0</sub>, P<sub>1</sub>, P<sub>2</sub>, P<sub>3</sub> et P<sub>4</sub>, représentant les catégories de phonèmes, n'est pas obligatoire mais contribue à la clarté de la présentation.) :

Code source du programme de syllabation automatique en français  
(en langage *AutoHotkey\_L*)

```
InputBox, string
```

```
P0 = (s)?
```

```
P1 = (s|p|b|t|d|k|g|f|v|s|z|ʃ|ʒ|m|n|ɲ|ŋ)?
```

```
P2 = (r|l)?
```

```
P3 = (j|ɥ|w)?
```

```
P4 = (i|y|u|e|ø|o|ɛ|œ|ɔ|a|ɑ|iː|yː|uː|øː|oː|ɛː|œː|ɔː|ɑː|ɑː)
```

```
string := RegExReplace(string, "U") P3 P2 P1 P0 P1 P2 P3 P4,  
"$1$2$3$4.$5$6$7$8", "", -1, RegExMatch(string, "P") P4, L)  
+ L)
```

```
MsgBox % string
```

Les différentes propositions de syllabation de mots français (Laeufer 1992) se présentent généralement chacune sous forme d'un ensemble de règles. Rédiger un programme informatique de syllabation automatique consiste, comme il a été dit, à traduire ces règles en un langage de programmation. Cette approche a été adoptée, par exemple, par Christophe Pallier (Pallier, 2004). Nous avons montré ici qu'il est plus économique de

partir d'une règle générale en ajoutant, éventuellement, des exceptions, par exemple :

/-.tɪ-/ → /-t.l-/(atlas, athlète...)

En anglais, même avec un nombre plus important des cas en dehors de cette même règle générale (à savoir six, voir l'extrait du code source ci-dessous), un algorithme fondé sur le modèle de la fig. 1 s'avère efficace.

```
string := RegExReplace(string, "(^|[^æɔə])([æʌɒɪʊ])\.(.)",
"$1$2$3.") ; happy, city
string := RegExReplace(string, "ï", "ɪ")
string := RegExReplace(string, "ö", "ʊ")
string := RegExReplace(string, "\.([td])l", "$1.l") ;
needless, effortless
string := RegExReplace(string, "\.stʃ", "s.tʃ") ; textual US
string := RegExReplace(string, "\.stju\.", "s.tju.") ; textual
UK
string := RegExReplace(string, "([^ɪ])ə\.([lr])([ei])", "$1ə$2.$3") ;
period; financially, blackberry, cooperation
string := RegExReplace(string, "ə\.nər.i", "ən.ər.i") ` ` ;
dictionary
```

Le code source complet des programmes de syllabation automatique en anglais est accessible à l'adresse suivante : <http://web.uniplovdiv.bg/rousni/autosyll/>; il peut être comparé avec celui, par exemple, de Joshua Tauberer (2011). En outre, un programme dérivé de ce code (autosyll.exe) est téléchargeable à la même adresse ; son mode d'emploi se résume ainsi : sélectionner une séquence de phonèmes, appuyer et relâcher rapidement la touche *Ctrl* pour l'anglais ou *Shift* pour le français ; la transcription du mot s'affichera, avec des points pour noter les frontières syllabiques, ainsi que les règles spécifiques appliquées (s'il y en a). Pour réaliser plus facilement le test de l'algorithme de syllabation, on peut sélectionner la transcription des mots proposée, par exemple, par le dictionnaire en ligne de *Cambridge University Press* : les symboles autres que phonémiques (points, accents, barres obliques) seront effacés dans un premier temps pour procéder ensuite à l'application des règles de syllabation automatique.

Il est à noter que ces règles sont pratiquement fondées sur des informations concernant exclusivement les phonèmes. Pour autant, on ne peut nier le rôle que joue l'accentuation dans la syllabation en anglais, puisque le timbre des voyelles et l'accentuation sont interdépendants.

Cette interdépendance n'est cependant pas toujours manifeste dans la transcription phonétique. Ainsi, les voyelles /ɪ/ et /ʊ/ sont les seules en anglais qui n'apparaissent jamais en finale de mot mais peuvent se trouver en finale de syllabe (en dépit du principe des contraintes phonotactiques). Cette « incohérence » de syllabation s'explique par une confusion de notation : « il existe également un *ɪ* et un *ʊ* non accentués qui ne se confondent pas avec *ɪ* et *ʊ* bien qu'on n'utilise qu'un seul signe pour les transcrire » (Viel 1992: 49). (Nos propres observations expérimentales montrent d'ailleurs qu'en syllabe ouverte les /ɪ/ et /ʊ/ non accentués se confondent assez souvent avec le schwa.) Il est donc impossible de déterminer les frontières syllabiques à partir d'une suite de symboles phonémiques comme /kə'ɪkju:ləm/ qui ne fournit aucune information directe sur l'accentuation/la centralisation concernant les voyelles /ɪ/ et /ʊ/. En matière de règles de syllabation en anglais, le recours à des règles d'accentuation s'impose donc afin de lever l'indétermination du timbre des voyelles /ɪ/ et /ʊ/, à moins de partir d'une transcription plus informative au niveau segmental, comme c'est le cas du *Longman Dictionary of Contemporary English* :

**curriculum** /kə'ɪkju:ləm/      **citizen** /'sɪtɪzən/

C'est le cas aussi du dictionnaire *The CMU Pronouncing Dictionary* où l'accentuation des voyelles dans les mots fait partie intégrante de la notation même de chaque voyelle et détermine ainsi son timbre.

Face au mode de transcription où le timbre/l'accentuation des voyelles /ɪ/ et /ʊ/ n'est pas explicitement noté dans les formes phonémiques à syllaber (par exemple les trois /ɪ/ dans /sɪləbɪfɪkeɪʃən/), nous suggérons le recours à une seule règle d'approximation issue de considérations rythmiques pour éviter le recours aux nombreuses règles d'accentuation traditionnelles et traiter assez correctement la syllabation automatique des mots contenant /ɪ/ et/ou /ʊ/ : les voyelles /ɪ/ et /ʊ/ sont généralement non accentuées en voisinage d'une voyelle accentuée, c'est-à-dire autre que /ɪ/, /i/ (en finale de mot), /ʊ/ ou

/ə/. Cette règle n'est pas valable pour les mots composés comme *off-putting* /,ɒf'pʊt.ɪŋ/. D'ailleurs elle est sans effet, par exemple, sur le second /ɪ/, pourtant non accentué, dans le mot *signification* /,sɪg.nɪ.fɪ'keɪ.ʃən/ et *particular* /pə'tɪk.jʊ.ləʳ/, mais une fois la règle rythmique appliquée, pour deux occurrences successives des voyelles /ɪ/ et /ʊ/, la seconde sera non accentuée.

À titre d'exemple d'application pratique, le dictionnaire *The CMU Pronouncing Dictionary*<sup>2</sup> sera converti en API, avec notation des frontières syllabiques (voir annexe 3). Ce dictionnaire de prononciation de l'anglais américain, utilisant un code de transcription appelé *Arpabet* où chaque phonème est représenté par une séquence de caractères ASCII, rend compte de l'accentuation des voyelles, ce qui facilite largement la syllabation automatique. Par exemple, le /ʊ/ dans *ambushes* est centralisé (car en syllabe ouverte non accentuée), alors qu'il ne l'est pas en syllabe fermée comme dans *ambush* et *ambushed* :

```
AMBUSH AE1 M B UH2 SH
AMBUSHED AE1 M B UH2 SH T
AMBUSHES AE1 M B UH0 SH IH0 Z
```

Il en est du même du /ɪ/ : des signes différents sont utilisés en fonction 1) de la forme sous-jacente de la voyelle et 2) de l'accentuation, par exemple :

```
REMIND R IY0 M AY1 N D
CITIZEN S IH1 T IH0 Z AH0 N
```

La syllabation automatique est donc largement facilitée par cette notation explicite de l'accentuation, mais, en revanche, d'autres difficultés de conversion Arpabet - API apparaissent en raison, par exemple, de l'opposition entre /u/ et /ʊ/ (qu'on observe en syllabe ouverte non accentuée devant respectivement une voyelle ou une consonne), dont l'absence dans *The CMU Pronouncing Dictionary* est illustrée par les deux mots suivants :

```
ACCENTUATE AE0 K S EH1 N CH UW0 EY0 T
```

<sup>2</sup> <http://www.speech.cs.cmu.edu/cgi-bin/cmudict>

ACCUSATION AE2 K Y UW0 Z EY1 SH AH0 N

Notre objectif à été de générer automatiquement une liste en texte plein des mots anglais (133354 dans *The CMU Pronouncing Dictionary*) accompagnés de leur transcription phonétique en API et en conformité exacte avec celle qui est proposée, par exemple, par *Cambridge Advanced Learner's Dictionary*. Cette nouvelle forme du dictionnaire de prononciation constitue la base de données d'un outil de gestion d'associations, didactiquement utiles, sur la base de similarités phonologiques, y compris structurelles, entre mots anglais.

RÉFÉRENCES BIBLIOGRAPHIQUES

- Clements 1990:** Clements, G. N. The role of the sonority cycle in core syllabification. // *Papers in Laboratory Phonology I*. Cambridge: Cambridge University Press, 1990, 283–333.
- Ditcheva 2010:** Ditcheva, M. Formalisation d'un aspect poétique du discours par évaluation de la distance phonologique entre mots. // *Plovdiv University Paissii Hilendarski, Scientific Papers*, vol. 48, Book 1, Part B, 2010 – Philology, 407–412.
- Hoard 1971:** Hoard, J.W. Aspiration, tenseness, and syllabification in English. // *Language*, N°47, 1971, 133–140.
- Kahn 1976:** Kahn, D. *Syllable-based generalizations in English phonology*. Bloomington: Indiana Univ. Linguistics Club, 1976.
- Laeufer 1992:** Laeuffer, C. Syllabification and resyllabification in French. // *Theoretical analyses in Romance linguistics*. Linguistic Symposium on Romance Languages, Ohio State University. Amsterdam: John Benjamins, 18–36.
- Marchand 2009:** Marchand, Y., C. R. Adsett, and R. I. Damper. Automatic syllabification in English: A comparison of different algorithms. // *Language and Speech*, vol. 52, N°1, March 2009, 1–27.
- Nikolov 2010:** Nikolov, R. Similarité de prononciation entre mots – développement théorique et application. // *Plovdiv University Paissii Hilendarski, Scientific Papers*, vol. 48, Book 1, Part B, 2010 – Philology, 413–416.
- Pallier 2004:** Pallier, C. *Syllabation des représentations phonétiques de Brulex et de Lexique*. <<http://www.pallier.org/ressources/syllabif/syllabation.pdf>>.
- Pulgram 1970:** Pulgram, E. *Syllable, word, nexus, cursus*. La Hague: Mouton, 1970.
- Ridouane 2008:** Ridouane, R. Syllables without vowels: phonetic and phonological evidence from Tashlhiyt Berber. // *Phonology* 25, 2008, 321–359.

- Saussure 1916:** Saussure (de), F. *Cours de linguistique générale*, Lausanne: Payot, 1916.
- Tauberer 2011:** Tauberer, J. *The P2TK automated syllabifier*. <<https://p2tk.svn.sourceforge.net/svnroot/p2tk/python/syllabify/syllabifier.py>>.
- Treiman 1988:** Treiman, R. & Danis, C. Syllabification of Intervocalic Consonants. // *Journal of memory and language*, 27, 1988, 87–104.
- Viel 1992:** Viel, M. *La phonétique de l'anglais*. Paris: Que sais-je ?, 1992.
- Wall 2001:** Wall, L., Christiansen, T., Orwant, J. *Programmation en Perl*. O'Reilly Media, Inc., 2001.