



## Jugements sur le nombre de syllabes et coordination temporelle des gestes articulatoires

Anisia Popescu, Ioana Chitoran

UFR Linguistique et Clillac-ARP, 8 Place Paul-Ricoeur, 75013 Paris, France  
[anisia.popescu@univ-paris-diderot.fr](mailto:anisia.popescu@univ-paris-diderot.fr), [ioana.chitoran@univ-paris-diderot.fr](mailto:ioana.chitoran@univ-paris-diderot.fr)

### RESUME

---

L'article traite de la relation entre la durée acoustique des rimes et la connaissance phonologique des locuteurs concernant sur le nombre de syllabes dans un mot. L'étude présentée est une réplique et une extension d'une étude portant sur cette relation dans le cas des sesquisyllabes de l'Anglais Américain. L'extension est basée sur des résultats portant sur la coordination temporelle des gestes articulatoires des consonnes liquides en coda de syllabe. Le cas des sesquisyllabes nous permet de tester l'hypothèse d'une représentation partagée entre le contrôle moteur et la phonologie. Les résultats sont en conformité avec cette hypothèse et montrent que cette représentation a une variabilité inter-locuteurs.

### ABSTRACT

---

#### **Syllable count judgments and temporal organization of articulatory gestures.**

The paper investigates the relationship between the acoustic duration of rimes and speakers' phonological knowledge of syllable count judgments. It is a replication and an extension of an earlier study explaining this relationship in the case of American English sesquisyllables. The extension is based on results pertaining to the temporal coordination of articulatory gestures in liquid coda consonants. The case of sesquisyllables allows us to test the hypothesis that speakers share a common representation for speech motor control and phonological knowledge. Our results are consistent with this hypothesis and show that this representation varies between speakers.

---

**MOTS-CLES :** sesquisyllabes, coordination articulatoire, consonnes liquides

**KEYWORDS:** sesquisyllables, articulatory coordination, liquid consonants

---

## 1 Introduction

Les locuteurs natifs ont des intuitions robustes sur le nombre de syllabes dans un mot. Lavoie & Cohn 1999 et Tilsen & Cohn 2016 ont montré qu'en anglais américain il y a une classe de mots avec un noyau composé soit d'une voyelle tendue soit d'une diphtongue suivie d'une coda liquide (*feel*, *fire*) pour laquelle les locuteurs ont des jugements variables sur le nombre

de syllabes. Cette classe de mots a été dénommée *sesquisyllabes* (Lavoie & Cohn 1999). Tilsen & Cohn 2016 ont montré qu'il y a une corrélation entre les jugements sur le nombre de syllabes (JNS) et la durée acoustique des rimes – la durée acoustique diffère entre des mots associés à des jugements monosyllabiques et des jugements dissyllabiques. Ils proposent sur la base de ces résultats qu'une représentation commune est utilisée pour les processus phonologiques comme les JNS et le contrôle moteur de la parole.

L'étude présentée ici maintient comme hypothèse la représentation partagée entre contrôle moteur et phonologie et propose de poursuivre sa vérification par une réplique et une extension de Tilsen & Cohn 2016. Nous proposons dans cette étude l'extension de l'hypothèse au cas des codas liquides complexes, sur la base des résultats obtenus par Marin & Pouplier 2010 pour l'anglais américain. Ces résultats montrent qu'en anglais américain l'organisation temporelle des gestes articulatoires dans le cas des codas complexes liquides présente un patron de coordination qui est différent de celui des codas complexes sans liquides. En coda non liquide, les gestes articulatoires de la consonne (nasale ou occlusive) adjacente au noyau vocalique ont un patron de coordination *en opposition de phase* à la fois avec la voyelle précédente et avec les consonnes suivantes. Ce n'est pas le cas pour les consonnes liquides. Les liquides en coda présentent un patron de coordination *en phase* avec la voyelle et *en opposition de phase* avec les consonnes adjacentes. Quand la complexité phonotactique augmente (CVC : *feel* vs. CVCC : *field*) cette différence de patrons de coordination se traduit acoustiquement par un raccourcissement significatif de la durée de la voyelle dans le cas des codas complexes liquides (durée  $/i/_{\text{field}} < \text{durée } /i/_{\text{feel}}$ ). Le raccourcissement est présent mais n'est pas significatif dans le cas des nasales ou des occlusives (durée  $/i/_{\text{beams}} \approx \text{durée } /i/_{\text{beam}}$ ). Si la durée de la rime a une influence sur les JNS, le raccourcissement significatif dans le cas des mots avec des codas liquides aura une influence sur les JNS des items avec des codas complexes. Pour tester cette hypothèse nous avons répliqué Tilsen & Cohn 2016 en rajoutant des items avec des codas complexes.

## 1.1 Méthodologie

Nous avons mis en place une expérience dont le design expérimental est identique à celui de Tilsen & Cohn 2016 à quelques modifications près, expliquées ci-dessous. 20 locuteurs natifs de l'anglais américain ont participé à cette expérience, tous étudiants en première année à l'Université de Chicago. L'expérience était composée de deux tâches séquentielles : une tâche de production suivie d'une tâche de jugements sur le nombre de syllabes. Au moment de la partie production, les participants n'étaient pas au courant de la partie suivante, soit celle de jugement. La tâche de JNS a été modifiée par rapport au design original : Tilsen & Cohn 2016 utilisent une barre graduée de 1 à 2 pour solliciter les JNS. Pour éviter les complications liées à la conversion de données continues en données catégorielles peut s'avérer compliquée. C'est pourquoi nous avons opté pour un choix parmi trois options de réponse : (1) **1** pour un mot considéré monosyllabique, (2) **1.5** pour un JNS intermédiaire – un mot de plus d'une syllabe mais moins de deux syllabes, et (3) **2** pour un mot considéré dissyllabique. Comme dans Tilsen & Cohn 2016, les instructions justifiaient un choix de 1.5 pour un JNS. Trois des 20 locuteurs ont été éliminés car ils ont donné des jugements incohérents pour les items de contrôle.

Les stimuli cibles étaient composés de triplets *syllabe ouverte (CV) – coda simple (CVC1) – coda complexe (CVC1C2)* avec un noyau vocalique consistant soit d'une voyelle tendue soit d'une diphtongue suivie soit d'une consonne latérale soit d'une consonne rhotique et d'une deuxième consonne : *fee - feel – field, tie – tire – tired*. Dans le cas des codas complexes, la majorité des stimuli étaient dimorphémiques (*10 items mono-morphémiques, 41 items dimorphémiques*) : *23 formes du passé – pooled, peaked (12 avec liquides ; 6 avec nasales/occlusives) ; 18 formes du pluriel – meals, stains (6 avec liquides ; 5 avec nasales/occlusives)*. Aux stimuli cible nous avons ajouté trois types de paires de contrôles : (1) voyelle relâchée suivie d'une consonne liquide, nasale ou occlusive (*gull-gulp, bin-bins, peak-peaked*) ; (2) voyelle tendue/diphtongue suivie d'une consonne nasale ou occlusive (*zoo – zoom – zoomed, pea -peak – peaked*) et (3) des mots clairement disyllabiques (*public, bacon*).

Un enregistreur Zoom H4NPRO a été utilisé pour recueillir les données de production. L'analyse acoustique a été faite sous Praat (Boersma & Weenink 2015). Comme les frontières entre voyelle et liquide ne sont pas clairement identifiables, nous avons comparé la durée de la voyelle + la consonne suivante (VC1) pour tous les items. La mesure envisagée pour mesurer le degré de raccourcissement de la durée de la voyelle dans chaque paire coda simple - coda complexe est le ratio  $VC1_{\text{complexe}}/VC1_{\text{simple}}$  ( $D[i:t]_{\text{field}}/ D[i:t]_{\text{feel}}$ ). Les valeurs attendues se trouvent dans l'intervalle [0.5, 1]. Plus la valeur est proche de 1 moins le raccourcissement est important. Pour la tâche de jugement une application a été créée, suivant les indications de Tilsen & Cohn 2016. Contrairement à ces derniers, nous n'avons pas écarté les réponses avec des temps de réponse au-delà de 5 secondes n'ont pas été écartées. A la place nous avons enregistré les temps de réponse pour chaque item. Toutes les analyses statistiques ont été faites avec le logiciel R (R Core Team 2013) en utilisant les packages des modèles linéaires mixtes (lme4) et des modèles de régression ordonnés (ordinal).

## 1.2 Prédictions

Pour les mesures de durée on s'attend à des ratio  $VC1_{\text{complexe}}/VC1_{\text{simple}}$  inférieurs à 1 pour tous les types de consonnes en coda, avec un ratio plus éloigné de 1 dans le cas des liquides que dans le cas des nasales et des occlusives ( $\text{Ratio}_{\text{liquides}} < \text{Ratio}_{\text{nasales/occlusives}} < 1$ ).

Basé sur Tilsen & Cohn 2016, avec le rajout d'une deuxième consonne en coda, on s'attend à une augmentation des JNS entre les items CVC et CVCC indépendamment du type de consonne (liquide, nasale, occlusive). La différence de JNS observée entre les codas simples liquides d'une part et les codas simples nasales ou occlusive d'autre part sera préservée dans le cas des codas complexes.

Par contre, si on tient compte des résultats de Marin & Pouplier 2010, le rajout de la deuxième consonne va déclencher un raccourcissement significatif de la voyelle dans le cas des liquides mais pas dans le cas des nasales ou occlusives. Ceci résulterait dans une durée raccourcie de la rime dans le cas des codas complexes liquides. On s'attend alors à ce que les locuteurs augmentent les JNS dans le cas des nasales ou occlusives, mais pas dans le cas des liquides. Ces prédictions sont illustrées dans le Tableau 1.

Un choix de 1.5 pour un JNS pourrait s'avérer contre-intuitif pour les participants. Même si les locuteurs n'optent pas tous pour un jugement intermédiaire ( $JNS=1.5$ ) dans le cas des codas liquides, le temps de réponse pourrait indiquer une différence de représentation entre les liquides et les nasales ou occlusives : on s'attend à des temps de réponse plus longs dans le cas des liquides.

A part la durée et le temps de réponse nous avons considéré trois autres paramètres qui pourraient avoir une influence sur le choix de JNS : la fréquence lexicale des items (d'après le Corpus COCA, Davies 2008), la composition morphémique (*mono- ou di-morphémique*), le type du deuxième morphème (formes du passé ou du pluriel) et l'orthographe (*le nombre de voyelles écrites non-consécutives: field – 1V ; pooled – 2V*).

	<i>JNS attribué : 1 simple</i>	<i>JNS prédit : complexe</i>	<i>JNS attribué : 1.5 simple</i>	<i>JNS prédit : complexe</i>	<i>JNS attribué : 2 simple</i>	<i>JNS prédit : complexe</i>
Coda liquide	1 $\sigma$ feel	1 $\sigma$ field	1.5 $\sigma$ feel	1.5 $\sigma$ field	2 $\sigma$ -	-
Coda nasale	1 $\sigma$ pain	1.5 $\sigma$ pained	1.5 $\sigma$ -	-	2 $\sigma$ -	-
Coda occl.	1 $\sigma$ peak	1.5 $\sigma$ peaked	1.5 $\sigma$ -	-	2 $\sigma$ -	-

Tableau 1 : Prédications des JNS pour les codas complexes en fonction des JNS attribués aux codas simples correspondants (basé sur Tilsen & Cohn 2016 et Marin & Pouplier 2010)

## 2 Résultats

Nous allons présenter les résultats en trois étapes. En 2.1 nous allons confirmer les résultats de Marin & Pouplier 2010 portant sur le raccourcissement acoustique de la voyelle devant les codas liquides. La section 2.2 portera sur les jugements sur le nombre de syllabes (JNS). Finalement, en 2.3 nous allons tester la corrélation entre les deux types de résultats.

### 2.1 Durée acoustique de la voyelle

Une comparaison de modèles linéaires mixtes avec *Consonne\_Coda* comme facteur fixe et *Locuteur*, *Item*, *Fréquence* comme facteurs aléatoires montre qu'il y a un effet global du type de consonne coda sur le ratio  $VC1_{\text{complexe}}/VC1_{\text{simple}}$  ( $p= .025$ ). Les séquences V + latérale et V + rhotique ont une durée acoustique significativement plus courte dans des items CVC1C2 que dans des items CVC1 (*Figure 1 gauche*). Les paires avec des séquences V + nasale ou V + occlusive ne montrent pas cette différence ( $p=0.17$ ). Il n'y a pas de différence entre les latérales et les rhotiques ( $p= .57$ ). Ces résultats sont en accord avec Marin & Pouplier 2010. Une analyse par locuteurs montre que le degré de raccourcissement est variable entre locuteurs. Certains locuteurs ne présentent pas de raccourcissement significatif dans le cas des liquides. Cette variabilité inter-locuteurs devrait se retrouver dans l'attribution des JNS.

Une même analyse sur la durée totale des rimes (ratio  $VC1C2_{\text{complexe}}/VC1_{\text{simple}}$ ) indique qu'il y a une différence significative ( $p = .031$ ) entre l'allongement de la durée de la rime pour les items avec des séquences V + liquide d'une part, et les items avec des séquences V + nasale ou V + occlusive d'autre part (Figure 1 droite). L'allongement des items avec des consonnes nasales ou occlusives est plus grand. Ceci n'est pas surprenant si on prend en compte le raccourcissement significatif de  $VC1_{\text{complexe}}$  dans le cas des liquides.

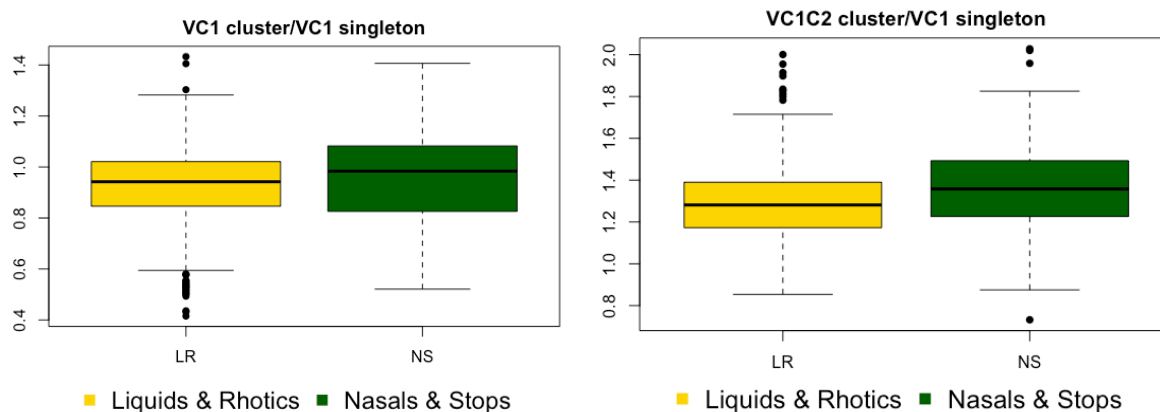


Figure 1 :  $VC1_{\text{complexe}}/VC1_{\text{simple}}$  (gauche) et  $VC1C2_{\text{complexe}}/VC1_{\text{simple}}$  (droite) en fonction du type de consonne en coda (liquide ou nasale/occlusive)

## 2.2 Jugements sur le nombre de syllabes (JNS)

Les résultats des jugements des locuteurs sont plus variables que ceux des données de production. On peut pourtant faire des généralisations. La figure 2 montre le nombre total des trois types de JNS (1, 1.5 ou 2 syllabes) attribué aux différents degrés de complexité phonotactique (*coda simple vs. coda complexe*) et aux types de consonne en coda (*latérale, nasale, rhotique ou occlusive*). On peut observer que dans le cas des codas simples des JNS supérieurs à 1 ( $JNS=1.5$ ,  $JNS=2$ ) sont attribués exclusivement aux items avec une coda liquide. Ces JNS sont attribués par un même groupe de locuteurs qui sont cohérents dans leurs réponses. Ce résultat confirme la variabilité des JNS rapportée pour les sesquisyllabes (Lavoie & Cohn 1999, Tilsen & Cohn 2016).

La figure 3 montre les temps de réponse pour les JNS en fonction de la complexité phonotactique et du type de consonne en coda. On observe que le temps de réponse est plus long pour les codas complexes (*boîtes de gauche pour chacune des quatre couleurs*) que pour les codas simples (*boîtes de droite pour chacune des quatre couleurs*). Le temps de réponse pour les codas liquides est plus long que pour les codas nasales et occlusives, dans le cas codas simples aussi bien que des codas complexes. Une comparaison de modèles linéaires mixtes montre qu'il y a un effet du type de consonne en coda sur le temps de réponse: pour les nasales et les occlusives le temps de réponse est plus court ( $p=.02$ ,  $p=.03$  respectivement), mais il n'y a aucune différence entre les latérales et les rhotiques ( $p=0.22$ ).

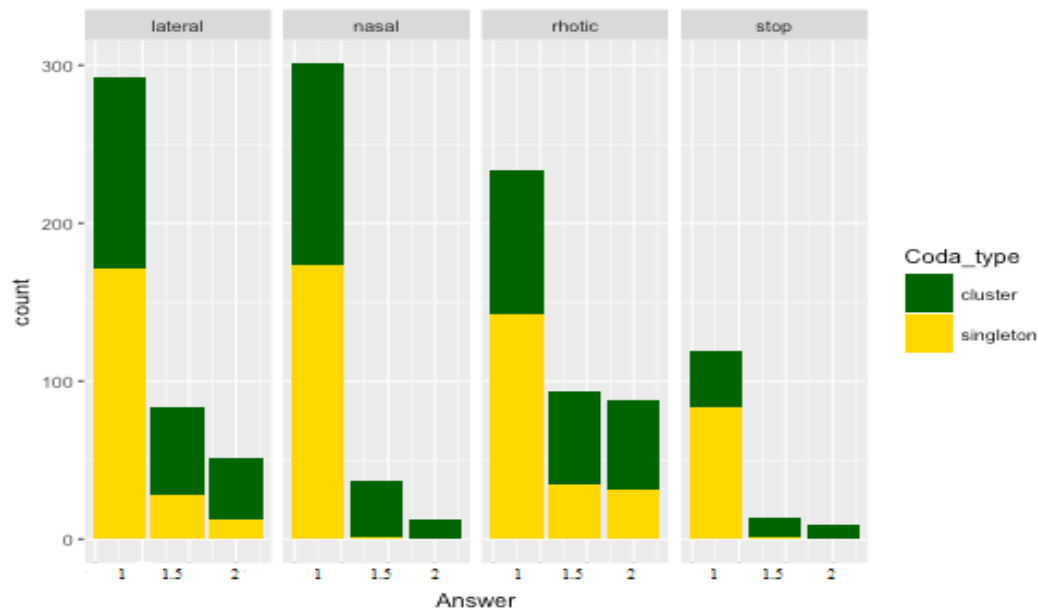


FIGURE 2: Nombre total de JNS (1, 1.5 resp. 2) attribué aux différents types de coda (simple ou complexe) et types de consonne coda (latérale, nasale, rhotique ou occlusive)

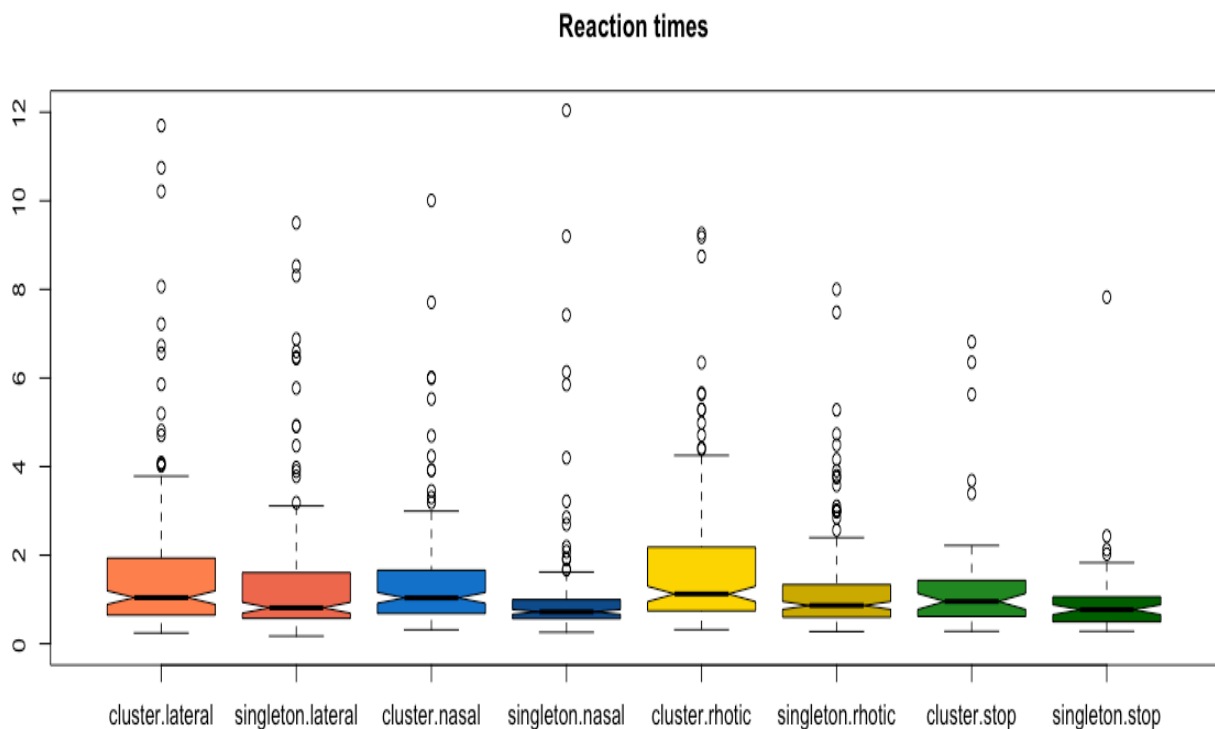


FIGURE 3: Temps de réponse (en secondes) pour les JNS pour différentes consonnes coda (latérale, nasale, rhotique ou occlusive) et type de coda (simple vs. complexe)

Même si les locuteurs n'optent pas tous pour un jugement intermédiaire ( $JNS=1.5$ ) dans le cas des codas liquides, les temps de réponse plus longs montrent que ces mots constituent une classe à part. On observe une différence entre les sesquisyllabes et leurs homologues avec des codas nasales ou occlusives.

Nous avons considéré la possibilité que la fréquence peut aussi influencer l'attribution des JNS. Nous avons trouvé que de manière générale le même groupe restreint de locuteurs attribue des jugements supérieurs à 1.5 aux sesquisyllabes. Une analyse des logarithmes des fréquences à partir du corpus COCA (Davies 2008-) montre que ce groupe devient plus grand lorsque le mot jugé est moins fréquent.

Dans le cas des codas complexes, la majorité des stimuli sont dimorphémiques. Un test  $\chi^2$  montre qu'il y a une différence entre les items monomorphémiques et dimorphémiques ( $\chi^2 = 44.711$ ,  $df = 4$ ,  $p = 4.566e-09$ ). Les locuteurs attribuent plus de jugements supérieurs à 1 aux items avec des codas complexes formées de deux morphèmes. Pour le type de morphème on trouve que les locuteurs donnent plus de jugements supérieurs à 1 si le deuxième morphème est une marque du passé (*-ed*) que si c'est une marque du pluriel (*-s*) ( $X\text{-squared} = 13.886$ ,  $df = 2$ ,  $p = .00096$ ). Une cause possible pour la différence entre les deux types de morphèmes pourrait venir de l'orthographe, i.e. la présence ou l'absence d'un graphème en plus (*peeked* vs. *peaks*). Nous avons donc considéré le nombre de voyelles non consécutives comme paramètre. Un test  $\chi^2$  montre que le graphème supplémentaire a une influence sur les jugements attribués aux items avec des codas complexes, indépendamment du type de consonne en coda. Des modèles cumulatifs mixtes confirment que le nombre et le type de morphèmes ont une influence sur les jugements des locuteurs. Ces résultats ont déterminé le choix d'inclure l'orthographe et la structure morphologique comme facteurs aléatoires dans les modèles cumulatifs présentés en 2.3.

Les résultats portant sur les jugements sur le nombre de syllabe sont donc influencés par la fréquence lexicale, le nombre des morphèmes et l'orthographe. Cependant, les résultats portant sur l'influence de la consonne en coda et les temps de réponse pour chaque JNS montrent que les locuteurs ont une représentation particulière pour la classe des items avec des codas liquides. Dans la section suivante nous allons présenter les résultats sur la corrélation entre la durée acoustique des rimes et le JNS.

### 2.3 Corrélation durée acoustique et JNS

Pour analyser la corrélation entre la durée acoustique et les JNS nous avons utilisé un modèle de régression ordonné cumulatif (*package ordinal*) avec des facteurs aléatoires (*Type de consonne coda*, *Locuteur*, *Complexité phonotactique*). Les résultats montrent que la durée totale des rimes est un bon prédicteur pour les jugements sur le nombre de syllabes ( $p < 2e-16$ ). Des JNS supérieurs à 1 sont attribués aux items avec des rimes plus longues (indépendamment du type de coda, simple ou complexe). Pour toutes les données (CV, CVC, CVCC) ensemble les jugements monosyllabiques (JNS=1) sont attribués aux items avec les rimes les plus courtes. Par rapport à la complexité phonotactique, les JNS varient entre locuteurs. Nous observons deux stratégies différentes. Certains locuteurs attribuent les mêmes JNS aux items par paires CVC - CVCC, d'autres augmentent la valeur du JNS de 1 syllabe/1.5 syllabes pour un item CVC à 1.5 syllabes/2 syllabes pour sa contrepartie complexe CVCC. La différence entre les deux groupes est le degré de raccourcissement des séquences VC1 : dans le cas où le JNS reste inchangé entre CVC et CVCC, il y a un degré

plus fort de raccourcissement de VC1 ; dans le cas où le JNS augmente entre CVC et CVCC le raccourcissement est moins marqué. Les deux stratégies sont illustrées dans le tableau 2. Il n’y a donc pas de stratégie universelle d’attribution des JNS, mais à l’intérieur d’une stratégie précise les prédictions sont confirmées.

	<i>JNS attribué pour CVC</i>	<i>JNS attribué pour CVCC</i>	<i>Degré de raccourcissement de la voyelle</i>
STRATEGIE 1	1 (ou 1.5)	1 (ou 1.5)	haut
STRATEGIE 2	1 (ou 1.5)	1.5 (ou 2)	bas

Tableau 2 : Illustration des deux stratégies utilisées pour l’attribution des JNS en fonction du degré de raccourcissement de la voyelle

Il y a donc une corrélation entre les jugements et la production. La durée totale de la rime est un paramètre qui est pris en compte lors de l’attribution des jugements. Le fait que les deux tâches expérimentales sont indépendantes soutient l’hypothèse d’une représentation commune du contrôle moteur de la parole et des jugements phonologiques.

### 3 Conclusion

Dans cette étude nous avons montré une corrélation entre la durée des rimes et les intuitions des locuteurs sur les jugements sur le nombre de syllabe dans un mot. Nous avons répliqué les résultats de Tilsen & Cohn 2016 et nous avons trouvé des résultats similaires dans l’extension de la tâche aux codas simples vs. complexes. Nous avons montré qu’il y a une grande variabilité inter-locuteurs dans l’attribution des jugements sur le nombre de syllabes. Pourtant, cette variabilité est cohérente avec la variabilité de production. Nos résultats vont donc dans le même sens que ceux de Tilsen & Cohn 2016 et restent consistants avec l’hypothèse d’une représentation partagée entre les processus gérant le contrôle moteur et la phonologie.

Malgré l’accord entre les résultats des deux études on ne peut pas conclure définitivement sur le statut de cette représentation. Nous avons vu, par exemple, que le rôle de l’orthographe ne peut pas être ignoré dans la tâche de jugement. En même temps, le paramètre acoustique de la durée de la rime reste une mesure indirecte de la coordination temporelle. Nous envisageons pour continuer, une étude articulatoire qui nous permettrait de tester de manière plus directe le lien entre la coordination temporelle en production et les jugements phonologiques.



## Références

BATES D., MAECHLER, M., BOLKER, B., WALKER, S., (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1), 1-48.

BOERSMA P., WEENINK, D., (2018). Praat : doing phonetics by computer [Computer program]. Version 6.0.08, retrieved in 2015 from <http://www.praat.org/>

CHISTENSEN R., H., B., (2015). ordinal : Regression Models for Ordinal Data. R package, <http://www.cran.r-project.org/package=ordinal/>

DAVIES M., (2008- ). The Corpus of Contemporary American English (COCA) : 560 million words. Available online at [http://corpus/byu.edu/coca](http://corpus.byu.edu/coca)

LAVOIE L.,M., COHN, A., (1999). Sesquisyllables of English : the Structure of vowel-liquid syllables. *In proceedings of the XIVth International Congress of Phonetic Sciences* 109-112

MARIN S., POUPLIER, M., (2010). Temporal Organization of Complex Onsets and Codas in American English : Testing the Predictions of a Gestural Coupling Model. *Motor Control*, 14, 380-407.

R CORE TEAM, (2013). R : a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>

TILSEN S., COHN, A., (2016). Shared representations underlie metaphonological judgments and speech motor control, *Laboratory Phonology : Journal of the Association for Laboratory Phonology*, 7(1) : 14, 1-13.