

Analyse des relations entre la durée des voyelles, leur aperture et le débit de la parole

Objectif

Ce travail étudie comment le débit de parole et l'aperture influencent la durée des voyelles orales. L'analyse vise à identifier des relations entre ces caractéristiques en vue d'améliorations potentielles de la synthèse vocale et des technologies de la parole.

Hypothèse

Nous proposons l'hypothèse que le débit de la parole et le degré d'aperture influencent significativement la durée des voyelles en français. Un débit plus rapide raccourcit leur durée, tandis que les voyelles ouvertes (/a/, /ɑ/) ont une durée plus longue que les voyelles fermées (/i/, /y/, /u/).

Base de données

L'analyse se produit sur des données comprenant des informations sur la durée de la production des voyelles dans différentes conditions. Le premier tableau contient des données de phrases complexes lues lentement et avec expression, en faisant attention à la prononciation des sons. Les données du deuxième tableau ont été tirées du discours cohérent au débit plus élevé. La base de données inclut deux tableaux et comprend 326 occurrences uniques. Les données utilisées pour cette analyse proviennent de deux ensembles distincts, appelés data1 (débit lent) et data2 (débit rapide). Chacun de ces ensembles contient plusieurs variables, dont les principales sont APERTURE, DEBIT et DUREE :

1. **APERTURE** : Une variable représentant les différents niveaux d'ouverture des voyelles orales. Pour l'analyse, nous avons divisé trois niveaux d'aperture : ouvert (1), mi-ouvert / mi-fermé (2), fermé (3).
2. **DEBIT** : Cette variable représente deux niveaux de débit du discours (rapide ou lent).
3. **DUREE** : C'est une variable numérique qui représente la durée des voyelles orales (en secondes). Cette variable est utilisée comme la variable dépendante dans l'analyse, en particulier dans les tests ANOVA.

Outils

- Excel
- Rstudio

Le script R avec les explications du travail est déposé sur GitHub.

Visualisation des données (RStudio)

Après avoir traité les données, nous avons visualisé la variation de la durée dans chaque ensemble de données pour en tirer des informations. Ces histogrammes montrent la distribution de la variable dans le discours du débit lent (Figure 1) et celui du débit rapide (Figure 2). Sur deux graphiques, l'axe X représente la durée, l'axe Y représente la fréquence des occurrences. En général, on observe que la durée diffère de 0,1 seconde, mais le discours plus rapide crée des valeurs aberrantes plus significatives (où la prononciation du son dépasse 0,75 seconde).

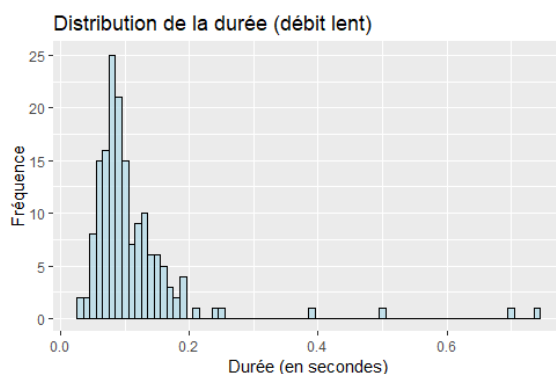


Figure 1. Distribution de la durée à un débit de la parole lent

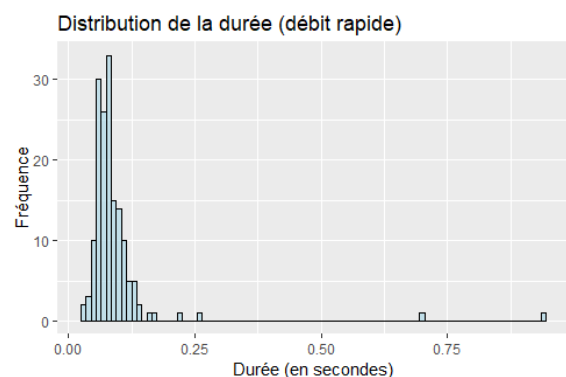


Figure 2. Distribution de la durée à un débit de la parole rapide

En outre, nous avons examiné la variation de la durée en fonction de l'aperture et du débit. L'historgramme « La distribution de la durée selon l'aperture » montre les répartitions de la durée pour chaque catégorie d'aperture (ouvert (1), mi-ouvert / mi-fermé (2), fermé (3)) dans la base de données combinée (Figure 3). L'axe X représente la durée, l'axe Y représente la fréquence des occurrences. On observe que les voyelles mi-fermées et mi-ouvertes (2) sont les plus fréquentes dans la base de données, et leur prononciation dépasse rarement 0,2 seconde. Les voyelles ouvertes et fermées (1, 3) ont des fréquences similaires, mais légèrement

inférieures à celles des voyelles mi-fermées et mi-ouvertes. Les voyelles ouvertes (1) ont tendance à présenter les valeurs aberrantes.

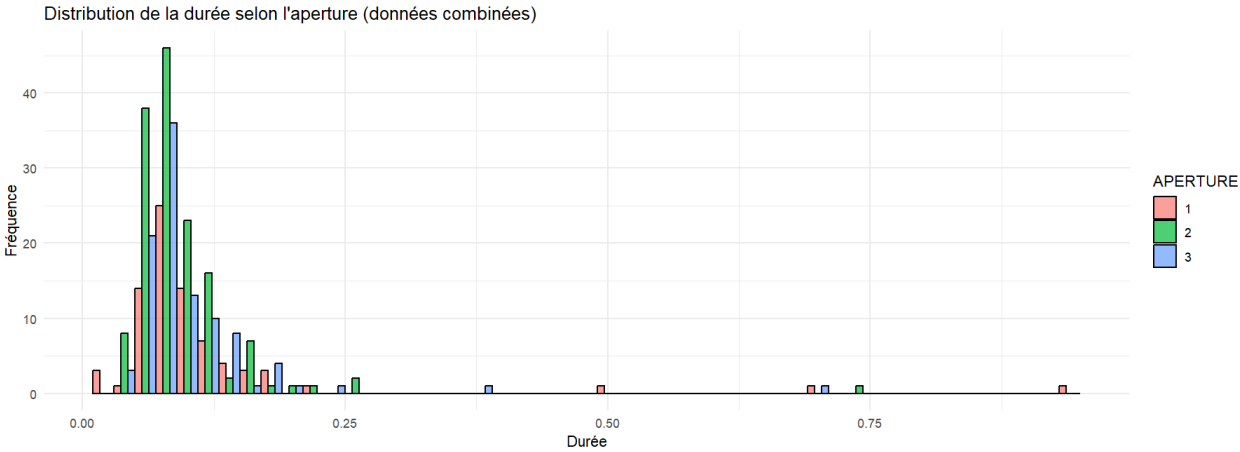


Figure 3. Distribution de la durée selon l’aperture dans les données combinées.

L’histogramme « La distribution de la durée selon le débit » illustre comment la durée des voyelles change avec le débit de la parole (Figure 4). L'axe X représente la durée des voyelles, l'axe Y représente la fréquence des occurrences. On observe une diminution de la durée des voyelles à mesure que le débit augmente. La majorité des durées, pour les deux débits, se situent entre 0,05 et 0,15 secondes. Les voyelles en débit rapide sont plus fréquentes dans les intervalles les plus courts (de 0,05 à 0,10 secondes).

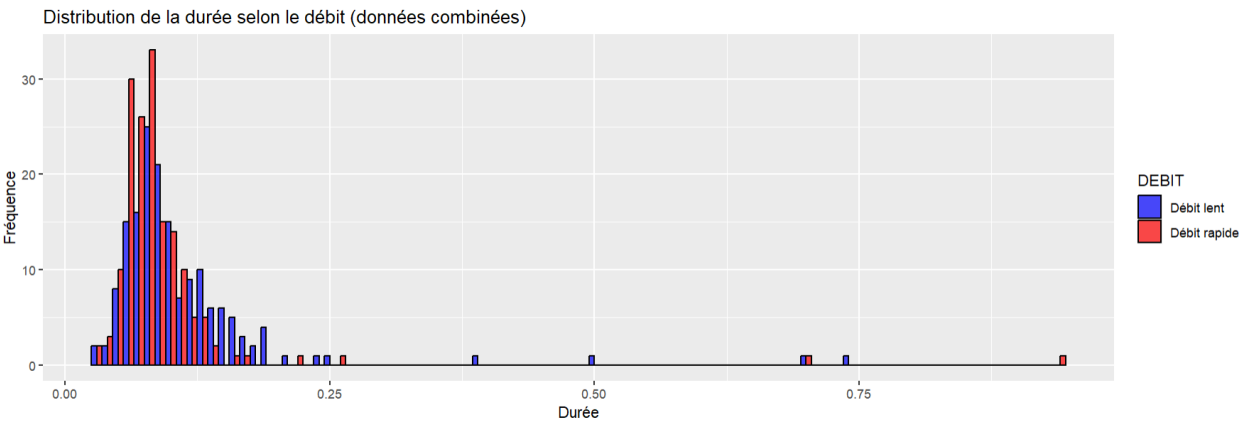


Figure 4. Distribution de la durée selon le débit dans les données combinées.

De plus, nous avons observé les répartitions de la durée en fonction du débit et de l’aperture à la fois. Le graphique « scatter plot » révèle comment le débit et l’aperture

interagissent pour influencer la durée (Figure 5). L'axe X représente l'aperture (1 - ouvert, 2 - mi-fermé/mi-ouvert, 3 - fermé), l'axe Y représente la durée des voyelles. Le graphique souligne les différences entre les niveaux d'aperture différents. Par exemple, les voyelles ouvertes se prononcent un peu plus longtemps, visibles par les points situés plus haut sur l'axe Y. Le point le plus intéressant est les valeurs aberrantes (de 0,9 seconde) qui appartiennent au débit rapide. On peut interpréter ces données comme la tendance à prolonger les voyelles ouvertes pendant l'hésitation.

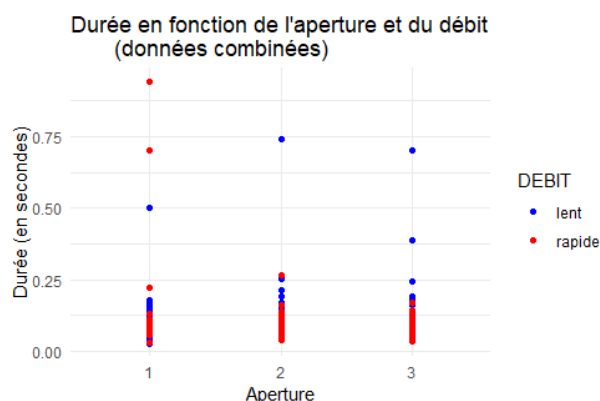


Figure 5. Distribution de la durée en fonction de l'aperture et du débit dans les données combinées.

Vérification des hypothèses

Les tests statistiques nous permettent de déterminer si les différences observées dans les données sont statistiquement significatives ou non. Comme nous avons choisi deux variables indépendantes, nous avons effectué l'ANOVA (Analyse de la variance) à deux facteurs pour vérifier l'hypothèse. Nous avons choisi l'aperture et le débit en tant que facteurs d'analyse. Avant le test, nous avons défini l'hypothèse nulle - la durée des voyelles ne dépend pas de leur aperture ni de leur débit. L'hypothèse nulle représente toujours que les différences entre les groupes ne sont pas statistiquement significatives ($p > 0,05$). Au cas où le test la conteste, il justifie l'hypothèse alternative : la durée des voyelles dépend de leur aperture ou de leur débit : les voyelles ouvertes ont tendance à la durée plus longue que les autres, et le débit peut augmenter leur durée.

L'ANOVA montre que la valeur p pour l'aperture dépasse 0,05, donc les différences ne sont pas significatives. (Tableau 1). Par conséquent, nous n'avons pas trouvé de preuve suffisante pour rejeter l'hypothèse nulle. Bien qu'il y ait une tendance à une différence dans la durée des voyelles en fonction du débit, cette différence n'est pas statistiquement significative.

selon les critères classiques. La valeur p pour l'interaction entre aperture et débit est également supérieure à 0,05, mais proche du seuil de significativité. Selon l'ANOVA, nous ne pouvons pas conclure que la durée des voyelles est influencée de manière significative par l'aperture ou le débit de la parole. En même temps, cela pourrait indiquer que la taille de la base de données n'était pas suffisamment sensible pour détecter des différences plus subtiles, qui pourraient être apparues dans d'autres conditions.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
APERTURE	2	0.0264	0.013221	1.729	0.1790
DEBIT	1	0.0274	0.027403	3.585	0.0592
APERTURE :DEBIT	2	0.0419	0.020969	2.743	0.0659
Residuals	318	2.4309	0.007644		

Tableau 1. Résultats de l'ANOVA à deux facteurs

Après avoir effectué l'ANOVA, nous avons décidé de réaliser un test-t pour analyser spécifiquement l'effet du débit sur la durée des voyelles. Le test-t est adapté pour comparer les deux groupes (débit lent et rapide), ce qui nous permet de vérifier si la différence dans la durée des voyelles entre ces deux groupes est statistiquement significative. La valeur p du test-t est inférieure au seuil de significativité ($p < 0,05$), et nous pouvons conclure qu'il y a une différence significative entre la durée des voyelles dans la parole rapide et la parole lente.

Résultats

En conclusion, nous avons trouvé une différence significative dans la durée des voyelles selon le débit de la parole. Pourtant, l'ANOVA n'a pas confirmé l'influence combinée du débit et de l'aperture. Les résultats des tests soulignent que, bien que l'analyse globale par l'ANOVA ne détecte pas d'effet significatif, l'analyse plus ciblée par test-t met en évidence une différence significative pour l'effet du débit seul.

Discussion

Les tests statistiques montrent que le débit influence la durée des voyelles, mais des limitations, comme le manque de données, affectent la fiabilité des conclusions. Bien que nous ayons effectué des tests pour examiner l'impact du débit de la parole et de l'aperture des voyelles sur leur durée, la taille de notre échantillon reste relativement modeste. Cela peut

limiter la puissance des tests. D'ailleurs, des données extraites peuvent être influencées par la qualité d'enregistrements.

Pour améliorer la fiabilité des résultats et confirmer nos hypothèses, il serait nécessaire d'élargir la base de données, notamment en incluant davantage d'enregistrements, ainsi que des données supplémentaires concernant des variables potentielles non explorées. Un échantillon plus large augmenterait la puissance statistique et représenterait mieux les différentes pratiques de parole.

Application pratique

L'analyse a montré que le débit de la parole influence la durée des voyelles de manière significative, ce qui peut être pris en compte dans le travail sur des algorithmes de reconnaissance vocale pour différents types d'utilisateurs (par exemple, pour les personnes qui parlent plus vite ou plus lentement). De plus, la connaissance de l'influence du type des voyelles sur la durée de leur prononciation peut être utilisée dans le développement d'interfaces utilisateur vocales.