Chloé Petridis, M2 Langue et informatique

**Rapport projet 1 – Transcription de parole spontanée**

**Introduction**

Ce projet a pour objectif d'analyser la phonologie d'une langue à travers la transcription d'un extrait audio, en se focalisant sur les phénomènes de disfluences afin de révéler les subtilités phonétiques et phonologiques propres à cette langue. Dans mon cas, j'ai choisi de travailler sur un extrait radio en anglais américain, avec un accent spécifique : celui du Texas. Cette variété de l'anglais est riche en spécificités phonologiques, puisqu'elle fait partie des accents parlés dans les États du Sud des États-Unis, une région souvent désignée sous le nom de Deep South.

L'accent texan présente des caractéristiques langagières intéressantes qui méritent une étude approfondie, notamment en ce qui concerne la prononciation de certains sons et les variations lexicales et syntaxiques. Mon intention à travers ce travail est de démontrer qu'une même langue peut varier significativement non seulement d'un pays à l'autre, mais également au sein d'une même région, et qu'un accent particulier peut contenir de nombreuses subtilités phonologiques uniques. L'analyse des disfluences, telles que les hésitations, les répétitions, ou encore les pauses longues, permet également d'observer comment ces phénomènes influencent le rythme et la fluidité du discours, tout en révélant des aspects propres à la parole spontanée.

La question principale que ce projet cherche à explorer est la suivante : quelles sont les spécificités phonétiques et phonologiques de l’accent texan et comment ces caractéristiques influencent-elles la production de la parole spontanée, notamment à travers les disfluences ?

En d’autres termes, je me demande dans quelle mesure les disfluences présentes dans le discours oral avec un accent texan reflètent des particularités inhérentes à cet accent, ou s’il s’agit de phénomènes universels dans la parole humaine.

Pour répondre à cette problématique, la méthodologie adoptée se divise en plusieurs étapes.

Tout d’abord, l’extrait que j’ai choisi provient d’une émission radiophonique américaine, où les locuteurs ont un accent texan. Il s’agit d’un discours spontané, ce qui permet d’analyser les disfluences naturelles.

Ensuite, à l’aide du logiciel Praat et de l’API Whisper, j’ai effectué une transcription fine de l’audio, en annotant manuellement les différentes disfluences rencontrées (hésitations, répétitions, allongements, rires etc.).

Puis les disfluences ont été classifiées dans un tableau Excel et leur position syntaxique a été relevée pour déterminer s’il existe des schémas récurrents dans leur apparition.

Mon intention à travers ce travail est de démontrer qu'une même langue peut varier significativement non seulement d'un pays à l'autre, mais également au sein d'une même région, et qu'un accent particulier peut contenir de nombreuses subtilités phonologiques uniques. L'analyse des disfluences, telles que les hésitations, les répétitions, ou encore les pauses longues, permet également d'observer comment ces phénomènes influencent le rythme et la fluidité du discours, tout en révélant des aspects propres à la parole spontanée.

La question principale que ce projet cherche à explorer est la suivante : quelles sont les spécificités phonétiques et phonologiques de l’accent texan et comment ces caractéristiques influencent-elles la production de la parole spontanée, notamment à travers les disfluences ?

En d’autres termes, je me demande dans quelle mesure les disfluences présentes dans le discours oral avec un accent texan reflètent des particularités inhérentes à cet accent, ou s’il s’agit de phénomènes universels dans la parole humaine.

Pour répondre à cette problématique, la méthodologie adoptée se divise en plusieurs étapes.

Tout d’abord, l’extrait que j’ai choisi provient d’une émission radiophonique américaine, où les locuteurs ont un accent texan. Il s’agit d’un discours spontané, ce qui permet d’analyser les disfluences naturelles.

Ensuite, à l’aide du logiciel Praat et de l’API Whisper, j’ai effectué une transcription fine de l’audio, en annotant manuellement les différentes disfluences rencontrées (hésitations, répétitions, allongements, rires etc.).

Puis les disfluences ont été classifiées dans un tableau Excel et leur position syntaxique a été relevée pour déterminer s’il existe des schémas récurrents dans leur apparition.

Enfin, une attention particulière a été portée aux variations phonétiques liées à l’accent texan, en comparant la production des sons dans les disfluences avec leur réalisation dans d’autres accents anglais.

**Historique**

L'accent texan est une composante spécifique de la langue américaine. Historiquement influencé par des langues telles que l'espagnol et l'anglais britannique, mais aussi par des langues autochtones, il possède aujourd'hui des caractéristiques distinctives au sein de la famille des accents du Sud des États-Unis. Cet accent, souvent lié à des caractéristiques culturelles comme le mode de vie rural et l'autonomie, reflète aussi l'histoire coloniale et les migrations qui ont influencé le Texas. L'accent texan est souvent employé dans les médias pour représenter un caractère « authentique » ou « terre-à-terre », et il est souvent employé pour faire référence à des stéréotypes en lien avec la culture du Sud américain.

Le mémoire de Caroline Jean Story « *Dynamics of Phonological Variation in Texas English* », Texas State University (2021) explique qu’en 1836, à la fin de la révolution Texane, l’anglais a commencé à remplacer l’espagnol. Ce phénomène s’est ensuite développé grâce à l’afflux de migrants anglophones dans le Sud des États-Unis.

**Description du corpus**

Dans cet enregistrement, nous entendons deux hommes d’environ une trentaine ou quarantaine d’années, originaires du Texas, se raconter des anecdotes de leur vie. Leurs accents marquent clairement leur appartenance à la région du Texas, et leur style de langage est assez décontracté, caractéristique d'une conversation informelle. L'enregistrement a été réalisé dans un studio de radio, où les deux animateurs semblent travailler régulièrement, et il s'agit vraisemblablement d'un extrait diffusé en direct, comme le suggère leur manière spontanée de s'exprimer et l'absence de montage apparent.

Le corpus transcrit dure 3 minutes. Au cours de l’enregistrement, on observe plusieurs moments de rires et de pauses, ainsi que quelques interruptions mutuelles entre les deux interlocuteurs, qui contribuent à une atmosphère conviviale et détendue. Ce type d’interaction est propice à l’apparition de phénomènes de disfluences, car les locuteurs semblent peu préoccupés par une élocution formelle.

**Analyse lexicale**

En premier lieu, j’ai calculé les mots les plus fréquents du corpus avec Python et l’API Whisper (pour transcrire l’audio afin de pouvoir calculer les mots les plus fréquents).

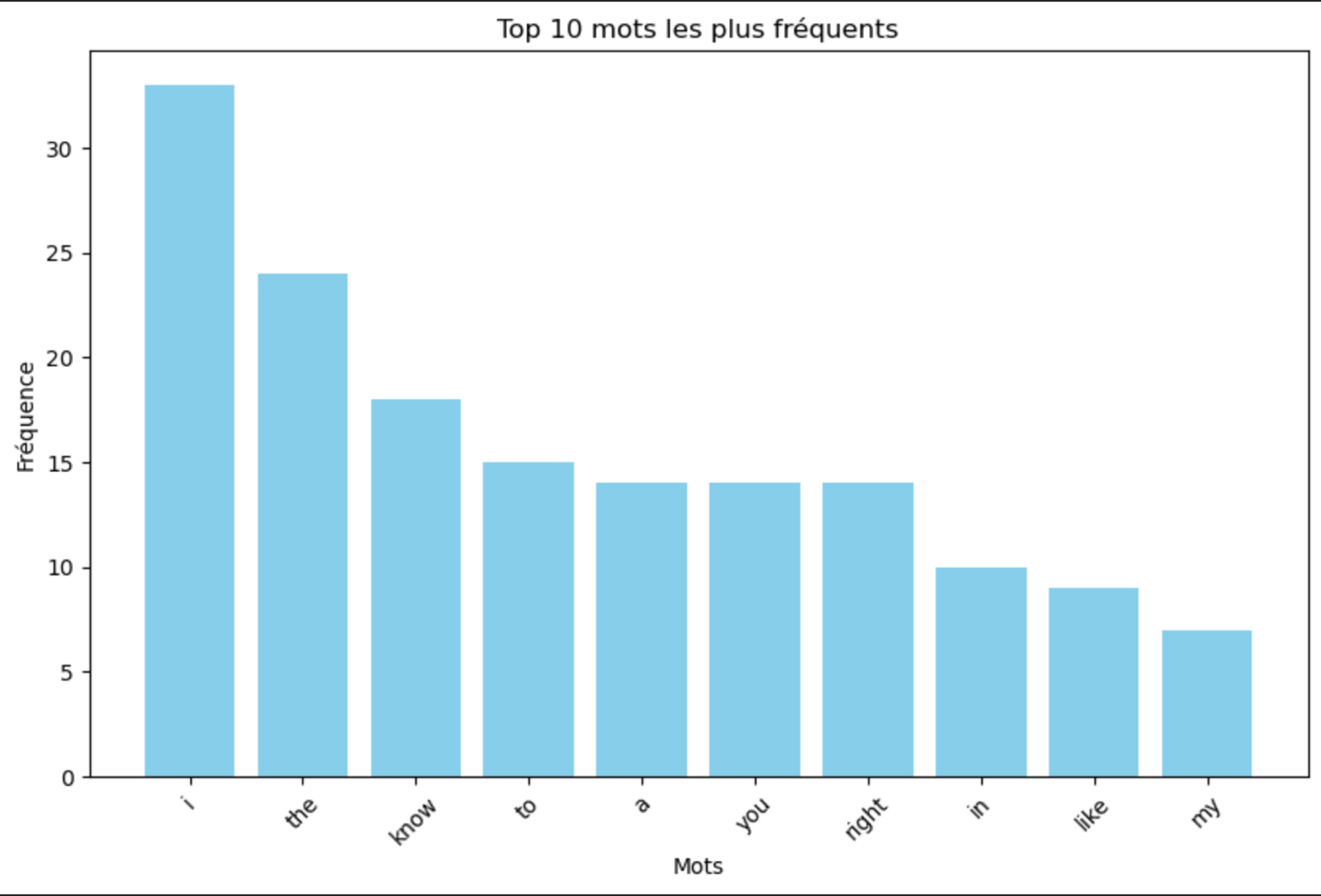


Figure 1 : Histogramme des 10 mots les plus fréquents fait avec MatplotLib

Comme le montre cet histogramme, les mots les plus fréquents sont des pronoms personnels (I, you), des articles définis (the) et indéfinis (a), des verbes (know, like), des prépositions (to, in), et des adverbes (right, like).

Et parmi ces mots, si on enlève les articles, les pronoms et les prépositions, il ne reste que des verbes et des adverbes, qui sont en fait des mots de remplissage (filler words en anglais). Si on les compte tous dans le dialogue, on peut en trouver 32. « You know » est présent 7 fois, « Like » 6 fois, « Right » et « Dude » 5 fois, « Yeah » 4 fois, « so » 3 fois, « Oh my God » 2 fois et « Really » 1 fois.

Selon le site JSTOR, la définition des filler words est la suivante : « Les mots de remplissage et les outils de conversation vont du phonémique (le niveau du son) au morphologique (les mots) en passant par le syntaxique (les phrases courtes). Ils permettent de formuler des pensées, d'écouter, de mettre en place des stratégies de politesse, d'attendre, de couvrir, d'adoucir et de signaler l'approbation ou le dédain, parmi d'autres. »

Cependant, même s’ils ont parfois un aspect parasite, ne nombreux linguistes les considèrent comme importants. En effet, certains considèrent que ces mots ont une fonction de ponctuation, ce qui permet de rendre le discours moins monotone et plus dynamique. Comme par exemple, Michael Barabro, animateur du podcast The Daily, complète que « les dysfluences telles que « hmm » peuvent être un moyen de « ponctuer les entretiens de manière à vous rappeler que deux personnes ont une véritable conversation » et d'exprimer l'intérêt sans porter de jugement, de faire preuve de curiosité et de maintenir la conversation. » (JSTOR, *Filler Words and Floor Holders: The Sounds Our Thoughts Make*)

Dans ce contexte, ces mots-là nous rappellent qu’il s’agit une conversation informelle et décontractée. Leur discours a l’air spontané et non préparé et cela montre qu’il n’y a pas de contrainte formelle dans le contexte dans lequel ils parlent, ce qui rend le dialogue authentique. Ces mots montrent également une ambiance chaleureuse et conviviale. Ils sont également utilisés pour que les locuteurs puissent organiser leur pensée, et également de marqueurs affectifs (comme le mot « dude » qui sert à témoigner de l’amitié). Ces mots servent enfin à mettre en valeur les émotions, comme pour souligner une idée ou un sentiment. Cela donne un côté complice à la relation des deux locuteurs.

Enfin ces mots témoignent de l’identité des locuteurs.

Cette atmosphère décontractée se traduit également par des rires entre les dialogues, ce qui montre encore une fois que les locuteurs sont familiers et qu’il ne s’agit pas d’une conversation formelle.

Ce dialogue contient également un lexique spécifique au contexte. Comme par exemple, on retrouve beaucoup de mots référant aux activités personnelles et à la vie de famille. Cela révèle également des loisirs que l’on trouve couramment chez les américains. En effet, ils parlent de sport, en particulier de baseball et de football américain, et de faire du cheval, qui sont des activités faisant partie du folklore Texan (notamment l’équitation). On retrouve également beaucoup de mots d’argot courant en anglais américain comme « buddy », « dude », « man », « oh my god ». Ces mots sont très souvent employés chez les américains, notamment dans les cercles sociaux amicaux.

Une expression typique du Sud des États-Unis est également employée, qui est « bless her heart ». Cette expression sert à exprimer sa sympathie ou sa gratitude envers une personne. Dans ce dialogue, le locuteur emploie cette expression pour exprimer la reconnaissance qu’il a envers sa femme. Cependant, cette expression peut également être employée de manière condescendante. Comme par exemple en 2016, la gouverneure de la Caroline du Nord avait employé cette expression en réponse à une attaque de Donald Trump sur le réseau social Twitter (CNN, *‘Nikki Haley response to Trump attack: ‘Bless your heart’*, 2016).

Une autre expression témoignant la bienveillance et la familiarité des conversations informelles aux États-Unis est « I know your pain ». Cette expression a pour but d’exprimer son empathie envers quelqu’un qui subit quelque chose de désagréable, en particulier une situation pénible. C’est tout à fait le cas dans ce dialogue car le locuteur explique qu’il a besoin de prendre du temps pour lui et ne pas avoir sa femme derrière lui.

Dans une discussion plus formelle, un registre moins familier et plus soutenu aurait été utilisé, d’autant plus que les locuteurs n’auraient pas partagé de détails aussi intimes de leur vie.

**Analyse syntaxique**

Dans ce dialogue, beaucoup de phrases sont entrecoupées ou laissées incomplètes, soit pour laisser l’autre locuteur parler, soit à cause des rires, soit à cause d’hésitations. Ces constructions de phrases sont typiques des conversations informelles où l’on a souvent une pensée rapide, où il n’y a pas un effort de construction grammaticale très important. On pourrait penser que ces phrases incomplètes sont surtout fréquentes chez les locuteurs non natifs, mais elles sont également très courantes chez les locuteurs natifs, notamment lors de discussions informelles. Cela montre un discours spontané où les locuteurs ne sont pas obligés de suivre strictement les règles grammaticales.

On trouve également quelques questions formulées de façon informelle, comme « Riding, riding what ? » où l’on ne retrouve pas la construction standard avec l’inversion du verbe et du sujet ou même la présence d’auxiliaires. Des questions simplifiées, parfois ne se résumant qu’en un mot (« Really ? ») sont également présentes, ce qui montre l’oralité du discours.

Cela témoigne d’une volonté d’aller à l’essentiel, de vouloir échanger des informations de manière rapide. Ce qui n’est pas le cas dans une conversation formelle où l’on va faire attention au formalisme du discours.

Pronoms déictiques utilisés  
"You know", "I mean" : Il est fréquent d'utiliser ces expressions afin d'introduire ou de clarifier un point sans recourir à des phrases approfondies. Par exemple, "You know, I do bother", "I know your pain" utilise "you know" pour indiquer la collaboration et la compréhension. Les signes de familiarité établissent une connexion entre les interlocuteurs sans accentuer la syntaxe.

Concernant les références vagues et généralités, les phrases telles que "all them five producers" ou "for those who dont know" utilisent des pronoms déictiques ("them", "those") pour désigner des individus ou des objets sans fournir de détails. C'est le cas dans des échanges où les interlocuteurs partagent un contexte commun et n'ont pas besoin de préciser chaque référence par exemple.

Des ellipses sont également fréquentes, telles que "dude, I'm done man" au lieu de "I am done, man". C'est un trait caractéristique de l’argot américain, où l'efficacité est plus importante que la structure.  
 On oublie également souvent les auxiliaires et les articles dans le discours informel afin de simplifier la syntaxe. Par exemple, "all them are five producers" pourrait être traduit de manière plus officielle par "all of them are five producers" si on voulait construire une phrase grammaticalement correcte.

Une utilisation des conjonctions afin de relier des fragments est également courante.  
On utilise souvent des conjonctions telles que "so" et "and" pour relier des phrases ou des fragments de phrases, sans toujours suivre une logique syntaxique rigoureuse. À titre d'exemple, "And man I get to be a free out of the house without..." utilise le terme "and" pour ouvrir une pensée, une structure courante dans le langage informel.  
Effet de la Parataxe : Un effet de parataxe est créé par l'utilisation de nombreuses conjonctions telles que "and" ou "but" sans subordination, où les idées sont placées côte à côte sans hiérarchie syntaxique. Cela témoigne d'une réflexion spontanée, typique de la conversation.  
Ensuite, les formes contractées telles que "I'm", "it's", "he's" sont présentes partout et accentuent le ton informel. Il est fréquent dans la langue parlée de raccourcir les mots pour favoriser la fluidité.  
Utilisation de termes implicites : On utilise parfois des contractions informelles, telles que « gonna » au lieu de « going to » ou « wanna » au lieu de « want to ». Ces structures reflètent la spontanéité de la parole et diminuent la complexité syntaxique.

Enfin, il y a l’utilisation de répétitions syntaxiques pour indiquer l'insistance.  
Parfois, les locuteurs répètent des mots ou des phrases afin de souligner, d’insister, comme : "What, what dude, I live in a row house". La répétition a pour but de consolider l'intention ou de gagner du temps pour organiser la suite du discours.

Ici dans ce contexte, le locuteur a voulu insister sur son incompréhension. En effet, son interlocuteur dit qu’il a un grand terrain, et locuteur répète « what » pour insister sur le fait qu’il ne comprend pas pourquoi il dit ça alors qu’il vit en ville.  
Influence sur l'échange : L'effet de cette répétition est celui d'une parole qui se construit progressivement et qui témoigne d'une approche collaborative de l'interaction, où l'un des locuteurs peut reprendre un mot pour mieux se faire comprendre.

**Analyse phonétique**

Dans cette analyse, j’ai procédé à une analyse des formants F0, F1 et F2. Cela servira à connaître le pitch de la voix des locuteurs et l’ouverture des voyelles.

Concernant la fréquence fondamentale, celle-ci avoisine les 150 Hz, ce qui est la plage moyenne de fréquence chez les voix d’hommes. La fréquence fondamentale peut également refléter des émotionnels. Or, si celle-ci est à 150 Hz, cela signifie que le ton est plutôt calme. De fortes variations de fréquence fondamentale pourraient être dues à de l’excitation ou à de la colère.

Ensuite pour le premier formant, la fréquence moyenne est de 893,35 Hz, ce qui est caractéristique des voyelles ouvertes telles que /a/ ou /ae/ (comme dans "cat"). Cela laisse entendre que les locuteurs ont la possibilité d'utiliser des voyelles plus ouvertes, ce qui est courant dans certaines langues, en particulier avec un accent texan.  
Dans l'accent texan, il est possible que certaines voyelles telles que /æ/ (dans "man") soient prononcées de manière plus ouverte et peuvent également être diphtonguées. Un F1 assez élevé pourrait suggérer l'emploi de ce genre de voyelle, ce qui renforce l'authenticité de l'accent.

La position de la langue vers l'avant ou l'arrière dans la bouche est souvent liée au deuxième formant (F2).  
Un F2 à environ 1950 Hz est caractéristique des voyelles précédentes telles que /i/ (comme dans "see") ou /e/. Cela peut suggérer une propension à former certaines voyelles en plaçant la langue en avant.  
L'accent texan et d'autres accents du sud des États-Unis présentent parfois des voyelles antérieures diphtonguées ou prononcées avec un F2 élevé, ce qui crée un son distinctif par rapport aux accents d'autres régions anglophones.

A blue line graph with white text

Description automatically generated

Figure 2 sur la fréquence fondamentale (F0) au fil du dialogue

Cette figure montre la variation de la F0 au fil du temps d’enregistrement du dialogue. Les pics de F0 (entre 400 et 600 Hz) montrent de fortes accentuations ou des émotions fortes comme des rires, ou une forte intonation ou prosodie. Les segments montrant une F0 basse montrent soit un silence, soit une conversation plus posée.

Cela représente bien le dialogue où il y a des rires, des variations d’intonations et de prosodie, et des moments plus calmes, qui sont plutôt associés aux moments où le premier locuteur raconte ses anecdotes.

A chart of a diagram

Description automatically generated with medium confidence

Figure 3 : Diagramme sur la fréquence des 2 premiers formants (F1, F2)

Cette figure montre la fréquence de F1 et F2 au fil de l’enregistrement. Le premier formant (F1) indique l’ouverture de la bouche pendant la prononciation des voyelles.

Plus la fréquence de F1 est élevée, plus les voyelles sont ouvertes comme /a/ alors qu’une fréquence plus basse indique une voyelle fermée comme /i/.

Le second formant montre plutôt la position de la langue dans la bouche. Plus la fréquence de F2 est élevée, plus les voyelles sont produites avec la langue vers l’avant (comme /i/ et /e/). Plus la fréquence de F2 est basse, plus les voyelles sont produites avec la langue vers l’arrière (comme /u/).

Ce qui correspond aux résultats évoqués si dessus. Même si pour cette analyse, nous n’avons pas isolé chaque voyelle pour mesurer les formants, mais que nous avons juste calculé une moyenne de ces formants dans l’enregistrement, cela nous indique quand même une tendance à prononcer les voyelles de façon assez ouvertes comme nous l’indique le F1 et à prononcer des voyelles plus antérieures selon le F2.

En ce qui concerne la distribution des formants, la plupart des points sont situés entre 500 et 1500 Hz pour F1 et entre 1000 et 3000 Hz pour F2, ce qui est caractéristique d'un large éventail de voyelles typiques. Ces valeurs laissent entendre une variété de voyelles de plus fermées (F1 plus bas) à plus ouvertes (F1 plus élevé), et de voyelles postérieures (F2 plus bas) à antérieures (F2 plus élevé).  
La diversité de la répartition des points indique que les locuteurs peuvent utiliser des voyelles différentes, ce qui pourrait être lié à la nature spontanée de la conversation.

Les voyelles produites par les locuteurs texans semblent légèrement décalées par rapport à un accent américain standard. Par exemple, autour de 900 Hz pour F1 et 1950 Hz pour F2, on observe une concentration de points qui correspond à des voyelles plus ouvertes et antérieures, souvent liées à des caractéristiques spécifiques de l'accent texan.

Les teintes du schéma, symbolisant le temps, indiquent que les composants changent de manière régulière tout au long de l'enregistrement. Cela peut suggérer des modifications régulières des voyelles et des adaptations articulatoires en fonction de la parole. Par exemple, on pourrait représenter les voyelles diphtonguées caractéristiques de l'accent texan par des points avec des variations de fréquence, visibles dans la répartition des couleurs.

**Conclusion**

L'étude des caractéristiques phonétiques, phonologiques et syntaxiques de l'accent texan dans un contexte de parole spontanée a donné l'opportunité d'approfondir différents aspects de la linguistique appliquée. En se basant sur un enregistrement audio d'une conversation informelle entre deux locuteurs texans, nous avons réalisé une série d'analyses quantitatives et qualitatives afin d'approfondir notre compréhension des caractéristiques vocales, lexicales et stylistiques liées à cet accent du territoire. Cette recherche met en évidence l'importance de la variété linguistique, en mettant en évidence comment les subtilités de l'accent texan contribuent à une identité régionale spécifique dans le cadre de l'anglais américain.  
  
Tout d'abord, l'étude phonétique a mis en évidence des données essentielles concernant les caractéristiques vocales des locuteurs.

Le pitch de la voix, l'ouverture des voyelles et la position de la langue dans la bouche ont été évalués à l'aide des mesures de fréquence fondamentale, ainsi que des formants F1 et F2. Les voix masculines ont une fréquence fondamentale moyenne de 149.97 Hz, ce qui suggère une intonation généralement calme, ce qui suggère que les locuteurs se trouvent dans un environnement détendu. En outre, les fréquences de F1 (893 Hz) et de F2 (1949 Hz) sont en accord avec l'emploi de voyelles ouvertes et antérieures, ce qui est fréquent dans les accents du Sud des États-Unis, notamment au Texas. Le timbre et la position des voyelles dans cet accent sont à l'origine d'un son distinctif qui peut être considéré comme typiquement texan.

Quant à l'analyse des disfluences, nous avons souligné la présence de nombreux mots-complètes tels que « like », « you know », « right » et « dude ». Ces termes, souvent utilisés dans les discussions informelles, reflètent une interaction sociale amicale et décontractée. Leurs rôles dépassent les simples interruptions du discours, car ils contribuent à la structuration des échanges et à la formation de liens entre les interlocuteurs. Ces termes sont également utilisés comme indicateurs d'hésitation et offrent aux locuteurs la possibilité de gérer la fluidité de leur discours, tout en soulignant leur appartenance à un registre de langage familier. L'utilisation fréquente de termes tels que "oh my god" ou "vraiment" apporte également des éléments d'accent et d'émotion, ce qui contribue à l'authenticité de la conversation.

D'un point de vue syntaxique, l'analyse a révélé une prédominance de phrases incomplètes, d'hésitations et de fragments de phrases, caractéristiques des échanges informels. Des questions simples, parfois réduites à un mot (« Really? »), et des formes contractées (« I'm », « it's », « he's ») permettent de rendre le discours plus fluide. Ces traits sont fréquents dans le langage oral, où l'efficacité est souvent privilégiée par rapport à la complexité grammaticale. Selon cette recherche, il est démontré que les locuteurs utilisent une structure syntaxique moins formelle, ce qui accentue l'aspect informel et la spontanéité de l'intervention. Il y a également des expressions régionales telles que "bless your heart", ce qui renforce l'identité culturelle et linguistique texane dans le discours.

Pour résumer, cette recherche sur l'accent texan et les différences dans la parole spontanée souligne la diversité et la complexité de la langue orale dans un environnement régional. Elle met en évidence comment des caractéristiques linguistiques particulières, telles que la fréquence fondamentale, les éléments de formation, les mots complémentaires et les structures syntaxiques, jouent un rôle non seulement dans l'identité individuelle des locuteurs, mais également dans celle de leur communauté. Malgré l'influence de la standardisation de l'anglais américain, l'accent texan conserve des caractéristiques qui le différencient et qui peuvent être utilisées par les locuteurs en fonction du contexte social. De cette manière, cette étude permet d'approfondir notre compréhension de l'anglais américain régional et des dynamiques linguistiques qui influencent les interactions humaines dans la diversité culturelle.

**Bibliographie**

1. Story, Caroline Jean (2021). *Dynamics of Phonological Variation in Texas English*, Texas State University.
2. JSTOR. *Filler Words and Floor Holders: The Sounds Our Thoughts Make*.
3. CNN (2016). "Nikki Haley response to Trump attack: ‘Bless your heart’."
4. Michael Barbaro, animateur du podcast *The Daily* (The New York Times)
5. Python et l’API Whisper.
   * Citée pour la transcription de l'audio, afin de calculer la fréquence des mots utilisés dans le corpus.
6. Matplotlib.
   * Utilisé pour générer l'histogramme des mots les plus fréquents dans le corpus.