FACULTAD DE HISTORIA, GEOGRAFÍA Y LETRAS

DEPARTAMENTO DE INGLÉS

INFLUENCERS'S INFORMAL DIALECTS OF THE ENGLISH LANGUAGE ANALYSED FROM A PHONOLOGICAL SEGMENTAL PERSPECTIVE FOR THE TEACHING OF ENGLISH IN THE CHILEAN CONTEXT

TESIS PARA OPTAR AL GRADO Y/O TÍTULO DE LICENCIADO EN EDUCACIÓN CON MENCIÓN EN INGLÉS Y PEDAGOGÍA EN INGLÉS

AUTORES : ARIEL ALONSO PEREIRA SEPÚLVEDA

FRANCO ANDRÉS PÉREZ CABELLO

MARCELO ALEXIS SALAS JARA

LUCIANO ALONSO TORRES GALLARDO

PROFESORA GUÍA : LUZ PAMELA DÍAZ SILVA

SANTIAGO DE CHILE, MARZO DE 2022

FACULTAD DE HISTORIA, GEOGRAFÍA Y LETRAS

DEPARTAMENTO DE INGLÉS

INFLUENCERS'S INFORMAL DIALECTS OF THE ENGLISH LANGUAGE ANALYSED FROM A PHONOLOGICAL SEGMENTAL PERSPECTIVE FOR THE TEACHING OF ENGLISH IN THE CHILEAN CONTEXT

TESIS PARA OPTAR AL GRADO Y/O TÍTULO DE LICENCIADO EN EDUCACIÓN CON MENCIÓN EN INGLÉS Y PEDAGOGÍA EN INGLÉS

AUTORES : ARIEL ALONSO PEREIRA SEPÚLVEDA

FRANCO ANDRÉS PÉREZ CABELLO

MARCELO ALEXIS SALAS JARA

LUCIANO ALONSO TORRES GALLARDO

PROFESORA GUÍA : LUZ PAMELA DÍAZ SILVA

SANTIAGO DE CHILE, MARZO DE 2022

AUTORIZACIÓN

Anexo 1: AUTORIZACIÓN PARA REPRODUCCION SIBUMCE

Se solicita esta autorización a los autores de la investigación con el fin de alojar y publicar el trabajo en el Repositorio Digital SIBUMCE, a fin de dar libre acceso electrónico a las tesis, memorias y seminarios generados en la UMCE y así contribuir a su difusión, preservación digital y mayor visibilidad en la comunidad académica y público interesado.

UNIVERSIDAD METROPOLITANA DE CIENCIAS DE LA EDUCACION
SISTEMA DE BIBLIOTECAS – DIRECCION DE INVESTIGACION

IDENTIFICACION DE TESIS/INVESTIGACION

Título de la tesis, memoria o seminario : Influencers's Informal Dialects of the English Language Analysed from a Phonological Segmental Perspective for the Teaching of English in the Chilean Context

Fecha : 28-12-2021

Facultad : Facultad de Historia, Geografía y Letras

Departamento : Departamento de Inglés

Carrera : Pedagogía en Inglés

Título y/o grado : Licenciado en Educación con mención en Inglés y Pedagogía en Inglés

Profesor guía/patrocinante : Luz Díaz

AUTORIZACIÓN

Autorizo a través de este documento, la reproducción total o parcial de este trabajo de investigación para fines académicos, su alojamiento y publicación en el repositorio institucional SIBUMCE del Sistema de Bibliotecas UMCE.

Ariel Pereira	Franco Pérez	Marcelo Salas
Nombre/Firma	Nombre/Firma	Nombre/Firma
Luciano Torres		
Nombre/Firma	Nombre/Firma	Nombre/Firma

Santiago de Chile, 28 de Diciembre 2021

Imprima más de una autorización en caso de que los autores excedan la cantidad de firmas para este documento

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my family for their immense support; family is one of the greatest supports there are, and I am truly blessed to have that. I would like to thank our guiding teacher for being so supportive during the current pandemic and every teacher I have had; whether good or bad, each one of you was an experience and a guidance about how to be (or not to be). Lastly, I would like to thank Ariana Grande for motivating me with her music while I worked on this thesis.

Ariel

This journey has been long, and I would like to thank the people who have helped me get where I am today. My mother Maritza, who has encouraged me since the beginning; my friends I have met throughout the years at UMCE, who always had my back; the many teachers I have had the pleasure of working with, who taught me everything I know; Elsie, who provided entertainment and company on many occasions; and finally, my dear friend all the way in the United States (you know who you are). Thanks, I could not have done this without you.

Franco

To my family for their unconditional love and support, and especially to my mother Milka who never let me fall down: Each and every one of my achievements is yours. To my friends who I can always rely on to have my back: Your friendship is the most valuable of my possessions. To Polimá Ngangu Orellana and Claudio Montaño Ceura: Your music has heartened me countless times. Finally, to all the teachers who gave me a voice of courage whenever I needed it and without expecting anything in return: You are actual warriors of society, and I sincerely admire you. Thank you so much.

Marcelo

If I have to write to whom I am thankful, I will rather mention those who I cannot thank personally; therefore, I would like to thank Alexandra Elbakyan, the creator of Sci-Hub, who made this whole project possible. Also, I would like to thank all those podcasters and bands that accompanied me from 9 am until 11 pm nonstop. Finally, I want to thank all those authors and researchers who are part of this study as without their previous work, none of this would have been possible.

Luciano

Table of Contents

1. Introduction 1
2. Objectives 4
   2.1 General Objectives 4
   2.2 Specific Objectives 4
3. Literature Review 5
   3.1 Social Media. 5
      3.1.1 Social Media Users. 7
      3.1.2 Influencers. 9
         3.1.2.1 Linguistic Features of influencers 10
   3.2 Register & Dialect. 11
      3.2.1 Register. 11
         3.2.1.1 Types of register 14
      3.2.2 Dialect. 14
         3.2.2.1 Accent 17
   3.3 Phonetics & Phonology. 18
      3.3.1 Articulatory Phonetics. 19
         3.3.1.1 Problematic consonants for Spanish speakers' learners of English 22
         3.3.1.2 Variations of sounds in speech 23
      3.3.2 Auditory Phonetics. 23
         3.3.2.1 Sound quality 25
      3.3.3 Acoustic Phonetics. 26
         3.3.3.1 Spectrograms 26
   3.4 Chilean education system. 27
      3.4.1 Chilean level of English. 29
      3.4.2 Curricula. 31
      3.4.2.1 Socioeconomic Composition. 33
4. Methodology 38
   4.1 Selection of English Sounds. 38
   4.2 Selection of Speakers and Samples 38

4.2.1 Speakers Selected 40
4.3 Procedure 41
4.4 Data Analysis 43
4.4.1 Voiced labiodental fricative /v/ 43
4.4.2 Voiced dental fricative /ð/ 51
4.4.3 Voiceless dental fricative /θ/ 59
4.4.4 Voiced alveolar fricative /z/ 67
4.4.5 Voiced postalveolar fricative /ʒ/ 73
4.4.6 Voiced postalveolar affricate /ʤ/ 79
4.4.7 Voiceless postalveolar fricative /ʃ/ 85
4.4.8 Voiceless glottal fricative /h/ 92
5. Results 99
5.1 Voiced labiodental fricative /v/ 99
5.2 Voiced dental fricative /ð/ 100
5.3 Voiceless dental fricative /θ/ 101
5.4 Voiced alveolar fricative /z/ 102
5.5 Voiced postalveolar fricative /ʒ/ 103
5.6 Voiced postalveolar affricate /ʤ/ 103
5.7 Voiceless postalveolar fricative /ʃ/ 104
5.8 Voiceless glottal fricative /h/ 105
6. Discussion 106
7. Limitations 113
8. Conclusion 114
9. References 116

ABSTRACT

Currently, Chilean students do not get to achieve the expected level of proficiency in English as a foreign language proposed by MINEDUC, studies show teachers do not have B2 proficiency in the language, and students do not have enough English lessons and exposure to the language. This research aimed to solve the problem of the exposure to the target language that students receive by producing material and making pedagogical suggestions based on Influencers, a phenomenon whose influence over students and young learners has increased greatly in the last 15 years. For this purpose, a list of challenging English consonants for Spanish speakers was compiled. The analysis consisted of an acoustic study of the segmental features and manner of articulation of English and American Influencers using the speech analysis software *Praat.* Then, the material and suggestions were created based on the samples analyzed, emphasizing allophonic variations from casual English.

Keywords: Acoustic Phonetics, Segmental Features, Influencers, TEFL, Audiovisual Material.

Actualmente, los y las estudiantes chilenas no alcanzan el nivel de competencia en inglés como lengua extranjera propuesto por el MINEDUC, principalmente porque los y las profesoras no tienen el nivel suficiente en la lengua, y los y las estudiantes no tienen suficientes horas de inglés ni exposición al idioma. Esta investigación tuvo el propósito de resolver el problema de la exposición a la lengua objetivo que los y las estudiantes reciben al producir material y hacer sugerencias pedagógicas basadas en influencers, un fenómeno cuya influencia sobre estudiantes y aprendices jóvenes se ha ido incrementando vastamente en los últimos 15 años. Para este propósito, una lista de consonantes desafiantes del inglés para hablantes del español fue compilada. El análisis consistió en el estudio acústico de las características segmentales y la manera de articulación de hablantes ingleses y estadounidenses usando el programa de análisis de discurso *Praat*. Luego, el material fue creado basado en las muestras analizadas enfatizando variaciones alofónicas del inglés casual.

Palabras clave: Fonética Acústica, Características Segmentales, Influencers, TEFL, Material Audiovisual.

1. Introduction

One of the earliest methods used for language teaching, that science knows of currently, is the Grammar-Translation method from 1750 to 1880, whose aim was mainly literary (Howatt & Smith, 2014). Nevertheless, this widely-used method failed to "treat the spoken language with the respect it deserved" (Howatt & Smith, 2014, p. 79). It was not until 1880 when the teaching of spoken language had a major role to play, as speech was "the primary foundation of all language activity" (p. 81).

Oral skills when teaching and learning English took the spotlight when the Natural method and the Berlitz method became mainstream in Europe (1880-1920). Both methods focused on the "understanding and oral 'retelling'" of information (Howatt & Smith, 2014, p. 83), but it was divided according to the age of the learners as only adult speakers were taught how to maintain conversations. It was in this period where pronunciation, the topic of interest of this dissertation, became a matter of discussion in the classroom, though it would take time for this to become truly relevant.

The scientific period (1920-1970) methodologically improved what was done in the past; drills were carefully selected to teach English, though it did not feel like real-life communication. Importance on this matter appeared in the Communicative period (1970-2000+), when it was proposed that the goal of teaching English was to communicate ideas in the target language and the development of communicative competencies (Richards, 2006). This disregarded different aspects related to language acquisition, such as grammatical competences and raw knowledge about the language (Sonsaat and Levis, 2017). Because of this situation, Hymes (Kang et. al., 2017, p.507) argued "that a speaker-hearer's knowledge not only involved competence but also knowledge of appropriate use of the language in normal interaction." Effective communication goes hand-in-hand with intelligibility, which is the degree to which a message is understood by a listener (Munro & Derwing, 1995). Several aspects can produce a loss of intelligibility, not only pronunciation. For instance, as illustrated by Levis (2018), if a British English speaker talks about the *boot* of a car to an American speaker, there would be

"misunderstanding, or even partial or delayed understanding" (p. 12), because the American speaker calls that same thing a *trunk*; in this case, intelligibility is affected by word choice. Nonetheless, loss of intelligibility is "more closely tied to pronunciation deviations" (p. 12). Taking the previous example, if a speaker pronounces *boot* as *voot*, not only may the word be unfamiliar to the listener, but also would the speakers have to decode it to make it understandable; and even if they were able to do this, the amount of time it takes to process what was said changes the nature of the interaction.

Currently in Chile, English is being taught following the principles of CLT since 2012 when the Bases Curriculares de Educación Básica de Inglés were modified by MINEDUC (Herrada et. al., 2013). According to the Ministry of Education (henceforth MINEDUC), students should finish school with a B1 level of proficiency in English, according to the Common European Framework for Languages, CEFR. Level B1 involves being able to speak fluently and accurately on familiar topics. However, according to Herrada et. al. (2013), most of the time, students do not achieve this level, and even teachers do not follow what CLT proposes, and instead they teach with inappropriate techniques and material.

This dissertation aimed to create pedagogical suggestions with material extracted from open access social media sites with the goal of generating open access material to teach English consonants in the Chilean classroom. The material was based on the audiovisual samples extracted from social media of influencers, who are users that stand out in online platforms because of their reach and how well known they are amongst young people.

The material was accompanied by pedagogical suggestions on how to use the material to teach the consonants selected, as well as information on how to select audiovisual material for teaching purposes. With the intent to tackle the aforementioned weaknesses, the pedagogical suggestions were focused only on problematic consonants for a learner of English in Chile, and not all English sounds. The suggestions and material aimed to improve the learner's proficiency

on the production of the english sounds and their pronunciation, emphasizing the segmental features of the English sounds as well as the changes that these features go through when produced in a casual register.

2. Objectives

2.1 General Objectives

To examine segmental features of casual dialects of English through influencers' social media material from a phonological perspective, creating a proposal for pedagogical suggestions in the EFL classroom.

2.2 Specific Objectives

To create a corpus of phrases of casual dialects of English, containing consonants that represent a challenge to acquire for EFL students.

To generate a pedagogical suggestion for applications of casual English in the EFL classroom, specifically in teaching pronunciation, listening comprehension, and speaking.

3. Literature Review

In this study, four main components need to be explained: Social media, Register and Dialect, Phonetics and Phonology, and the Chilean education system. These four serve as the core components, as the aim is to showcase a possible new material for Chilean EFL teachers using influencers' social media clips, specifically from *Youtube*, as a corpus in the teaching of pronunciation and listening skills. While there has been plenty of research in the field of using audiovisual material to teach pronunciation (Brito et al., 2021; Rahmi, 2018; Kathirvel & Hashim, 2020; Adela, 2017; and Nuraeni, 2018), very few studies have considered using influencers' audiovisual material to teach pronunciation.

3.1 Social Media.

Social media is a worldwide phenomenon that can be found in almost every aspect of daily life, from the platform used to communicate with family members to the platform used to watch the new trending series. Due to this situation, it is complicated to pinpoint exactly what social media is; however, authors have tried to define it because of its relevance in people's everyday lives. A general definition of social media is given by Barbara Mróz-Gorgon and Kamila Peszko (2016) stating the following:

> The concept of social media encompasses a huge range of communication. It can be defined as a set of relationships, behaviors, feelings, empiricism, and the interaction between consumers, brands, in which there is a multidirectional communication, exchange of experience with advanced communication tools. (p. 38)

Although this is a broad definition, it allows us to understand that interaction between and among users is an essential aspect of social media and that said aspect is not limited only to communication between two users, but also with brands and large groups of people. This possibility of wide social interaction is what separates it from the traditional media, like TV and Radio, as in these last two the input only comes from one side to the other, and not vice versa. While Mróz-Gorgon and Peszko also place a considerable emphasis on the economic role that

social media has, it is not as relevant for this investigation as this research aims to analyze social media from a pure linguistic point of view and not from a sociolinguistic one.

A more concise and relevant definition for this investigation is given by Kaplan and Haenlein (2010), who define social media as "a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of User Generated Content" (p. 61). Additionally, Kaplan and Haenlein, like other authors, explain concepts closely intertwined with Social Media, like Web 2.0 and User Generated Content. These new terms and concepts that derive from Social Media are due to Social Media being a modern object of study; therefore, it has not been deeply researched and there is no consensus on what it actually is and is not. For the sake of the investigation, the two concepts previously mentioned, Web 2.0 and User Generated Content, will be briefly covered to create a common ground.

Firstly, Web 2.0 is a term coined in 2004 used to describe "a new way in which software developers and end-users started to utilize the World Wide Web" (Kaplan & Haenlein, 2010, p. 60); in other words, the content published in web pages was no longer stale, instead, it was evolving constantly through the modification of users. This last aspect is important, as its main difference with its previous incarnation is that web pages like blogs and wikis are constantly being developed by their own collaborative communities. In addition, the authors define Web 2.0 as "the ideological and technological foundation" (p. 61) in which users interact in social media.

Secondly, the concept of User Generated Content, or UGN, is defined as "the sum of all ways in which people make use of Social Media" (Kaplan & Haenlein, 2010, p. 61). Conversely, Wyrwoll (2014) states that "User-generated content is content published on an online platform by users," (p. 15) emphasizing that "published" means that the concept is not for a specific

receiver, but it is directed to a general audience (any user) or a limited audience (certain members of the community, but not a specific one). This definition creates a difference between UGN and Web 2.0 as Wyrwoll states that there are many authors that use these two terms interchangeably.

Having stated what social media is, the history of social media is much simpler and more direct in comparison to it. Kaplan and Haenlein (2010) give a brief rundown of the history of Social Media.

> By 1979, Tom Truscott and Jim Ellis from Duke University had created the Usenet, a worldwide discussion system that allowed Internet users to post public messages. Yet, the era of Social Media as we understand it today probably started about 20 years earlier, when Bruce and Susan Abelson founded ‘‘Open Diary,’’ an early social networking site that brought together online diary writers into one community. The term ‘‘weblog’’ was first used at the same time, and truncated as ‘‘blog’’ a year later when one blogger jokingly transformed the noun ‘‘weblog’’ into the sentence ‘‘we blog.’’ The growing availability of high-speed Internet access further added to the popularity of the concept, leading to the creation of social networking sites such as MySpace (in 2003) and Facebook (in 2004). This, in turn, coined the term ‘‘Social Media,’’ and contributed to the prominence it has today. (p.60)

It is worth noting that the history of Social Media is marked by what people have done in it and with it. Therefore, it is logical to determine who these Users of Social Media are and their role within the different platforms.

3.1.1 Social Media Users.

Social Media Users, or simply Users, are the main actors in what social media is, as they practically compose it and are responsible for UGC, a defining factor of social media. They are

people from all around the world, people with an account on Youtube, Instagram, or a Facebook profile, and the number of people using these apps grows constantly and at an incredible rate. According to data retrieved from BLACKLINKO, from 2019 to 2020, there was an increase of 10.9% year-on-year in the number of social media users, being that in 2019 the number of social media users was 3,484 billion and in 2020 was 3,960 billion.

This large number of users gives space to several different types of users, and therefore, a number of studies of said users. These studies acknowledge and characterize the users depending on their traits, as in "Who interacts on the Web?: The intersection of users' personality and social media use" study (Correa et al., 2010), or a label that characterizes a number of users, such as Gong et al. (2018) with silent users and content creators. This user community is constantly active in social media, which leads to different interactions and possibilities, as noted by Page et al. (2014):

> The exchange of information which takes place on social media sites when members discuss their shared interests and activities can in some cases become a form of 'goods' with economic value, exchanged with online advertisers who may use data mining applied to social media sites to target campaigns to particular customers. (p. 13)

This economic opportunity led to the rise of a particular group of social media users within the different apps, called Influencers. As Abidin (2016, p. 38) comments, "Influencers are characterized by their intention to generate profit and their emphasis on portraying and promoting "lifestyle" choices." (p. 38). This definition is too simplistic, considering the importance of Influencers for investigation; therefore, we will devote the next section to characterize and determine who Influencers are.

3.1.2 Influencers.

Influencers are individuals within social media who influence a large group of users. According to Poulopoulos et al. (2018), the term first arose with the boom of social media. The group also states that influencers "can manipulate others in various ways; in particular—and in relation to cultural heritage." (p. 240) In addition, they state how they influence other users of social media by simply visiting certain locations, generating trends, or facilitating interactions through the creation of communities. Concerning this, Dimos et. al. (2015) states that influencers are the most important users in social media in terms of economic activity, due to how they influence their followers into consuming certain brands or products. In relation to this, Abidin (2016) makes the following statement:

> Influencers are everyday, ordinary internet users who accumulate a relatively large following on blogs and social media through the textual and visual narration of their personal lives and lifestyles, engage with their following in "digital" and "physical" spaces, and monetize their following by integrating "advertorials" into their blog or social media posts and making physical appearance at events. A pastiche of "advertisement" and "editorial", advertorials in the influencer industry are highly personalized, opinion-laden promotions of products/services that influencers personally experience and endorse for a fee. (p. 1)

Dimos et. al. (2015) states the existence of three dimensions in which influencers can be related when talking about their capacity to engage social media users; those being "audience segmentation", "competitive set", and "influencer endorsement funnel". Although focused on the influencers' importance in marketing, this defines a core aspect of influencers targeting a specific audience and being a credible and relatable individual for said audience. Considering this last statement, the differentiation between types of influencers is not a relevant aspect for the investigation as long as they are relatable to the learners, which leads to relatability and how someone on the internet can be relatable. Doctor Abidin (2016) comes to the same conclusion as she states that Influencers modify their followers' opinions through their interactions with them because of the Influencers' image being perceived as relatable.

Djafarova and Rushworth (2017) concluded that social media influencers had a bigger impact on people than traditional advertising due to people considering influencers closer and more relatable. Djafarova and Rushworth (2017) explained in their study that this situation may be because online Influencers are perceived as bloggers or models instead of how traditional celebrities are Hollywood stars.

Since influencers have a certain power related to public opinion, as previously discussed, it is also important to determine how they do it, specifically concerning language. When convincing or persuading people to do something, language plays a vital role; thus, influencers need to choose their words carefully and appropriately.

3.1.2.1 Linguistic Features of influencers

Linguistically speaking, there is little to no research related to how Influencers communicate; instead, most researchers focus either on the psychological and social aspects of Influencers or on their economic impact and how brands can benefit from this type of user. In relation to authors who have studied any linguistic features of influencers, said investigations tend to be very limited in their scope and depth; like Abidin's research (2016) that presents a general overview on Singaporean influencers' linguistic features and how they are a means to appeal to their public, or Karjo and Wijaya's study (2020), that focus specifically on the language features of male and female beauty influencers. The authors showed that beauty influencers use a somewhat colloquial language, using slang, idioms, empty adjectives, and hedging information in certain situations. This information confirmed or acknowledged that this type of influencers uses everyday language, even including expressions that would not be acceptable in a more formal context. They also stated that these types of influencers tend to limit their statements to topics related to the type of influencer they are, and they do not dive into any other topics in their interactions with their followers.

In relation to what was previously stated, the register and dialect can also depend on the type of influencer, as register and dialect are in close relation with the language, vocabulary, and expressions used by the influencer.

3.2 Register & Dialect.

Languages are never rigid or static; they are constantly changing, and thus variations appear in several areas of it throughout time. A good example of this would be the historic event in the English language called "The Great Vowel Shift," a massive sound change in which the long vowels of English shifted in pronunciation during the fifteenth to eighteenth centuries; another example would be the debate about whether to use a gender-neutral language or not, a topic that has been especially prominent in the 21st century. Yet, language does not only change according to time, it also depends on the people who use it and the geographical location of those speakers. Along with this, it is important to recognize some of the different ways in which languages vary.

3.2.1 Register.

Different methods of speech can be found depending on the context in which the communicative act takes place; when a speaker chooses a specific method of speech, said method receives the name of 'register'. According to Nordquist (2019), "In linguistics, the register is defined as the way a speaker uses language differently in different circumstances." (para. 1). In this way, register can be acknowledged as an intersection between the language and context in which the former is being used. In relation to this, Małinowski (1923) first argued that a statement "(...) is never detached from the situation in which it has been uttered." (p. 307).

Therefore, this intersection happens since, consciously or unconsciously, the context is considered each time that a communicative interaction occurs.

Thus, the context will define the level of formality in which a person communicates to others. As Nordquist (2020) states, “These variations in formality, also called stylistic variation, are known as registers in linguistics. They are determined by such factors as social occasion, context, purpose, and audience.” (Para. 1). These factors will transform the speech act depending on how much each one of them varies. Therefore, depending on these factors, the communication process takes different forms with different results, such as formal register, casual register, among others. For the purpose of this study, the casual register is further explained in the following section.

In this context, Halliday (1985) describes register as “a variety of language, corresponding to a variety of situation [*sic*],” (p. 38) a description that reaffirms the understanding of register as an intersection between these two factors. This description can be thoroughly explained through the analysis of a further definition of register that the author (1985) also states as: “(...) a configuration of meanings that are typically associated with a particular situational configuration of field, mode and tenor.” (pp. 38-39).

Halliday (1985) refers to the field of discourse as: “(...) what is happening, to the nature of the social action that is taking place: what is it that the participants are engaged in, in which the language figures as some essential component?” (p. 12) The situation in which an interaction occurs, then, can vary due to the time, place, or moment in which the participants communicate. Therefore, these elements will shape the speech act, which will be highly different when comparing, for instance, a working environment with a family meeting.

Secondly, Halliday defines tenor as the following:

> The tenor of discourse refers to who is taking part, to the nature of the participants, their statuses and roles: what kinds of role relationship obtain among the participants, including permanent and temporary relationships of one kind or another, both the types of speech role that they are taking on in the dialogue and the whole cluster of socially significant relationships in which they are involved? (p. 12)

The level of confidence between the participants of the speech act, then, affects directly the way in which they communicate. The communicative process is particularly different when two people know each other, compared to a situation in which the speakers are acquaintances or strangers. The sociocultural circumstances of the speakers are also an important factor that influences the way in which communication is developed, mainly because of the knowledge and language accuracy that the participants dominate.

Finally, Halliday defines mode as the following:

> The mode of discourse refers to what part of the language is playing, what it is that the participants are expecting the language to do for them in that situation: the symbolic organisation of the text, the status that it has, and its function in the context, including the channel (is it spoken or written or some combination of the two?) and also the rhetorical mode, what is being achieved by the text in terms of such categories as persuasive, expository, didactic, and the like. (p. 12)

The type of speech that a speaker chooses, then, depends on the purpose and expectations that the speaker has. This is what determines the words and sentences and their organization, thus defining the way in which the content is communicated.

3.2.1.1 Types of register

With the several variations that exist regarding the level of formality of an utterance, it is possible to find five language registers or styles, namely the frozen, formal, consultative, casual, and intimate registers. According to Gen (n.d), "Each level has an appropriate use that is determined by differing situations." (Para. 1). A thorough description of one of the most commonly used registers and its determining situations is given by Joos (1967), who defines formal register as the words and vocabulary chosen by the speaker when producing speech, which tended to be adjusted to what the speaker wanted to say regardless of length.

It is highly important to recognize the main differences between the formal and the casual (or informal) register, which is the most important register for the purpose of this work. Joos (1967) defines casual register as the language used among friends and peers that is used in a daily context and usually includes idiomatic expressions and general ideas. These two styles are the most known registers of language and, considering how contrasting they are, they might also be the most recognizable ones. Notwithstanding, not only do communicative acts vary greatly depending on the context, but also on the way people naturally talk, i.e., the way they pronounce things. This form of language is unique to each speaker, for this reason, the concept of dialect is addressed below.

3.2.2 Dialect.

Apart from registers, it is possible to find several variations in the use of the same language. Akmajian et al. (2001) state that:

> No human language is fixed, uniform, or unvarying; all languages show internal variation. Actual usage varies from group to group, and speaker to speaker, in terms of the pronunciation of a language, the choice of words and the meaning of those words, and even the use of syntactic constructions. (p. 275)

In this sense, dialects can be found within the internal variations that the authors mention. The researchers (2001) defined dialect as the noticeable differences in the speech of groups of speakers of the same language. In this way, a dialect can be construed as any type of variation produced by two speakers of the same language; however, a thorough definition of the concept is given by Halliday et al. (2012):

> Each speaker has learnt (...) a particular variety of the language of his language community, and this variety may differ at any or all levels from other varieties of the same language learnt by other speakers (...) Such a variety, identified along this dimension is called a dialect. (p. 144)

With this definition, the authors elucidate the concept of dialect by explaining that cultural elements, such as the community to which the speakers belong, are the foremost factors that will shape the language varieties of dialect. Nevertheless, Akmajian et al. (2001) also state that:

> It is notoriously difficult, however, to define precisely what a dialect is, and in fact the term has come to be used in various ways. The classic example of a dialect is the *regional* dialect: the distinct form of a language spoken in a certain geographical area. (p. 276)

This, as the researchers explain, is the well-known idea of the concept of dialect; however, it is not the only variety that may exist in this field. They also identify a social dialect, which is defined (2001) as "(...) the distinct form of a language spoken by members of a specific socioeconomic class." Thus, even if geographical boundaries might be the most outstanding variety, sociocultural factors are also an important element regarding dialect. Therefore, despite the geographical area to which two speakers of a language might belong, their level of proficiency in that language, for example, may differ notoriously depending on their socioeconomic status (for instance due to their level of education), and then, said difference in knowledge results in a distinct form of using that same language in comparison to another speaker.

Furthermore, Akmajian et al. (2001) also state that "certain *ethnic* dialects can be distinguished, such as the form of English sometimes referred to as Yiddish English, historically associated with speakers of Eastern European Jewish ancestry." (p.276). Therefore, these three variations — regional, social, and ethnic — are to be considered as essential cultural factors which shape what linguists call dialect. However, these three types of variations do not necessarily influence language separately. Instead, as the researchers also declared: "It is important to note that dialects are never purely regional, or purely social, or purely ethnic. (...) As we will see, regional, social and ethnic factors combine and intersect in various ways in the identification of dialects." (p. 276).

Albeit the concept of dialect might be commonly understood as a less valuable type of variation in comparison to the standard form of a language, it is important to consider that this is not entirely accurate, as Akmajian et al. (2001) remark:

> In popular usage the term *dialect* refers to a form of a language that is regarded as "substandard," "incorrect," or "corrupt" as opposed to the "standard," "correct," or "pure" form of a language. In sharp contrast, the term *dialect*, as a technical term in linguistics, carries no such value judgment and simply refers to a distinct form of a language. Thus, for example, linguists refer to so-called Standard English as a dialect of English, which, from a linguistic point of view, is no more "correct" than any other form of English. (p. 277)

Another important characteristic to consider regarding dialect is the one stated by the authors (2001):

> Although dialects are often said to be regional, social, or ethnic, linguists also use the term *dialect* to refer to language variations that cannot be tied to any geographical area, social class, or ethnic group. Rather, this use of *dialect* simply indicates that speakers show some variation in the way they use elements of the language. (p. 276)

In this way, dialect is a concept that can embrace as many ways of communicating as the speakers of the language can produce, as is highlighted by the researchers (2001):

> Language variation does not end with dialects. Each recognizable dialect of a language is itself subject to considerable internal variation: no two speakers of a language, even if they are speakers of the same dialect, produce and use their language in exactly the same way. (...) The form of a language spoken by a single individual is referred to as an *idiolect*, and every speaker of a language has a distinct idiolect. (p. 277)

With this definition, it is possible to assume that the variations in language are basically endless, as every person uses language in a different way. Along with these variations, it is also possible to find the concept of accent, which will be explained in more detail below.

3.2.2.1 Accent

The concept of accent is commonly used interchangeably with dialect; notwithstanding, they do not share the same meaning. Jones (n.d.) defines accent as "a unique mode of pronunciation within a given language that is specific to a country, region, or social class." (para. 5) According to this definition, accent can be understood as a variation in language, specifically related to the way in which words are pronounced. The difference between accent and dialect relies, then, on the specific areas of language where variations occur; while dialect varies in the areas of Morphosyntax and Phonology, accent is a variation that occurs only within the Phonology area. Thus, "(...) accent is merely a part of dialect." (Jones, n.d, para. 10). In this sense, although both dialect and accent refer to "specific ways of speaking a language" (n.d, para. 7), accent is found at a lower level, and therefore it is encompassed within dialect. Taking into consideration both concepts, it is important to mention that this study was conducted based on the Standard accents of English.

Standard English is defined (Yule, 2014, as cited in Norman, 2017) as the idealized variety of English, which can be found in mass media, education, books, the news, and the type of English that is taught in schools as a second or foreign language. One suggestion as to why the standard variety of English is taught is provided by McArthur (2003), as cited in Farrell & Martin (2009), in which the author states that it is "(...) the variety most widely accepted, understood, and perhaps valued within an English speaking country." (p. 442) The key concept given by McArthur is *understood*. As an illustration, if a person were to speak with an Irish, Scouse, or Cockney accent, they might not make themselves understood to the same degree as an individual who speaks a standard accent of English.

In this regard, social or geographical features (register and dialect) are not the only types of elements that may signify a difference in a language, but their sounds and the way in which they are produced may also influence its comprehension parameters. For this reason, Phonetics and Phonology will be addressed in the following section.

3.3 Phonetics & Phonology.

To the everyday person, Phonetics and Phonology may seem similar or even the same concept, but there are core differences between the two that make them completely different fields of language, yet related to one another. Gut (2009) described them both as fields of linguistics concerned with the study of pronunciation, though each one with a different focus. On the one hand, Gut (2009) defines Phonetics as the field of study concerned with the production of speech sounds involving three different areas: articulatory phonetics, focused on the organs and muscle movements required to produce speech; acoustic phonetics, focused on the transmission of these speech sounds between speaker and listener; and auditory phonetics, focused on the way the sounds are perceived by the listener. In a similar vein, Ogden (2017) defines Phonetics as "the systematic study of the sounds of speech, which is physical and directly observable," stating that linguists may not consider phonetics as part of linguistics due to it being a physical phenomenon instead of abstract, as language is. The same three areas

(articulatory phonetics, auditory phonetics, and acoustic phonetics) are defined by Ogden (2017) in a similar fashion but with one different concept that replaces “auditory phonetics” with “perception”. The author defines articulatory phonetics as “how speech sounds are made in the body,” acoustic phonetics as “the physical properties of the sounds that are made,” and perception as “what happens to the speech signal once the sound wave reaches the listener’s ear” (2017, p. 2)

On the other hand, Ogden (2017) defines Phonology as the field of study that deals with sound systems, a sound system being dependent on each language. A more detailed definition is given by Gut (2009), who states that Phonology is concerned with the way sounds make up patterns in individual languages, and is divided into two areas: segmental phonology, which deals with individual sounds; and suprasegmental phonology, which deals with larger units, such as syllables and utterances, as well as aspects like stress and pitch. Additionally, McMahon (2002) states that phonology is the domain of “a language-specific selection and organization of sounds to signal meanings.” (p. 2) In relation to this, McMahon (2002) states that “phonologists are interested in the sound patterns of particular languages, and in what speakers and hearers need to know, and children need to learn, to be speakers of those languages: in that sense, it is close to psychology”. (p.2)

Having clarified the difference between phonetics and phonology, it is important to note that phonetics, as stated by Ogden (2017), has different fields and points of view, being the three main ones Articulatory Phonetics, Auditory Phonetics and Acoustic Phonetics.

3.3.1 Articulatory Phonetics.

According to Ogden (2017, p. 198), articulatory phonetics is “the aspect of phonetics which looks at how the sounds of speech are made with the organs of the vocal tract such as the

vocal folds, tongue, lips, velum, etc". These sounds of speech are defined to have certain characteristics or features that distinguish them. Pennington & Richards (1986) defined these characteristics as segmental features, which are the minimal units of sound in phonetic terms called phonemes, and they concluded that a basic understanding of them can help acquire a new language.

Phonemes in the English language can be classified according to what characteristics they possess; voicing, point of articulation, and manner of articulation (Ogden, 2017). Based on Ogden (2017), voicing occurs when the vocal folds located in the larynx vibrate while the sound is produced at the same time. According to Ogden (2006, p. 9) "sounds which are accompanied by voicing are called voiced sounds, while those which are not are called voiceless sounds."

According to Ogden (2006), the point of articulation refers to what parts of the oral tract, or articulators, are being used to articulate and shape the sound. Depending on how they are articulated, these consonants can be classified as bilabial (sounds made with the lips), labiodental (upper teeth against the lower lip), dental (back of the upper teeth), alveolar (alveolar ridge), postalveolar (behind the alveolar ridge), palatal (middle of the tongue raised towards the hard palate), velar (tongue back towards the soft palate) or glottal (sounds made at the glottis, between the vocal folds).

In relation to the manner of articulation, first, there are **vowels**. Considered one of the two major types of segments (the other being consonants) and defined by Ogden (2017) as sounds produced without a constriction in the vocal tract, meaning the air is free to pass down with no interruption. English vowels are classified depending on how they are produced, only one vowel sound or in groups of vowel sounds; monophthongs are sounds that consist of only one vowel, diphthongs consist of two and triphthongs consist of groups of three vowel sounds.

Then we have the **approximants**, which Ogden (2017) defined as the sound formed when two articulators (parts of the oral tract responsible for producing speech) are close together. While they are similar to vowels, they work like consonants.

Plosives are consonants made with complete closure of the vocal tract, hence their name, as the sound is released in what is typically called an explosion. They are formed in a process with three phases: closing, where the articulators come together to stop the air from escaping; hold, where the air is retained and pressure is built by the lungs trying to expel the air; and release or plosion, where the air is finally released, which results in a burst of noise.

Fricatives are sounds produced when two articulators are in close proximity yet leaving a small gap for air to pass through, which results in friction noise. Sometimes sounds can begin as plosives and end as fricatives, in which case they are called affricates.

Finally, **nasals** are sounds produced when the oral tract is completely closed. To produce them, air must flow through the nasal cavity while the velum is lowered.

Each phoneme is said to carry a functional load, which King (1967) defined as the contrast between two different phonemes, and measured by the number of minimal pairs that can occur between them. The concept is relevant to this research because looking at the functional loads of certain phonemes in English can provide insight into how mispronouncing phonemes might affect intelligibility and comprehensibility: When these high functional load phonemes are not pronounced correctly, the comprehensibility of the speaker is lowered much more than when low functional load phonemes are mispronounced (Munro and Derwig, 2006). The authors explain that the functional load of a phoneme is dependent on how frequent the sound is in the language, how the sound changes or does not change from accent to accent, and to what degree the message is understood by a listener when the sound is replaced by another; a higher functional load means that the sound does not change from accent to accent, and if it

changes to another sound, it highly affects intelligibility, and vice versa. Munro and Derwig (2006) analyze how sounds like /n/ and /l/ present a higher functional load than others like /d/ and /ð/, as the former lead to problems regarding intelligibility while the latter does not.

3.3.1.1 Problematic consonants for Spanish speakers' learners of English

Plenty of research has been conducted regarding the English sounds that are problematic for Spanish speakers. According to Arellano & Cisterna (2009), the most difficult sounds for them are the following: /t, d, v, b, θ, ð, ʃ, ʒ, s, z, ʧ, ʤ, h, ŋ, r, j, w/. Similarly, Gómez & Sánchez (2016) stated that /t, d, ʤ, v, ð, z, ʃ, ʒ, h, ŋ, r, j, w/ do not have an exact equivalent phoneme in Spanish, thus deserving particular attention. Gómez & Sánchez, however, excluded the phoneme /θ/ as they were from Spain, a country that makes a distinction between /s/ and /θ/. As Finch & Lira (1982) point out, the /θ/ used in most of Spain is an acceptable variation of the English /θ/. The rest of the Spanish speakers do not have the /θ/ phoneme. However, authors such as Coe (2001) would disagree with the previous lists to some extent. He analyzed all consonants and stated that /p, b, f, θ, t, d, s, ʧ, k, g, m, n, ŋ, l, r, j, w/ have an equivalent or a near equivalent sound and do not pose a serious difficulty. However, the author considered *Catalán*, an accent from Spain. He further declared that /v, ð, z, ʃ, ʒ, ʤ, h/ do pose a difficulty; by looking at the list of these phonemes, none of them is part of the Spanish system, hence the difficulty for Spanish speakers to learn them. This is in accordance with the Contrastive Analysis Hypothesis proposed by Lado (1957), who "believes that the degree of difference between the two languages also correlated with the degree of difficulty." (Joze, 2015, p. 1106).

It appears that there is no clear consensus yet about the phonemes that are problematic for Spanish speakers. Nevertheless, for the purposes of this research, the Contrastive Analysis Hypothesis (CAH) had a major role to play; in fact, all the sounds selected are not a part of the Spanish Phonetic System, thus they should be more difficult to acquire than the rest. Segmental

features are not the only thing to consider in the teaching of Phonetics & Phonology, as they could change when in contact with other sounds.

3.3.1.2 Variations of sounds in speech

Some phenomena regarding the phonemic system occur in conversational speech, as Ogden (2017) implied when pointing out how sounds change when affected by other sounds. Ogden named different characteristics of language that are only present when in connected speech, such as assimilation, glottal stops, and silent pauses. The first of these, assimilation, according to Ogden, is how a consonant sound affects the place of articulation of another; for example, in the phrase "own bonfire," the /n/ sound becomes /m/ due to the influence of the /b/ sound in "bonfire". The second concept, glottal stops, is the "adduction of the vocal folds usually before the oral closure is made" (2017, p. 107). The last concept, silent pause, is a natural phenomenon according to the author as there are times in a conversation when no speaker is talking.

As it was mentioned, there are problematic phonemes for Spanish speakers' learners of English and there are variations to these sounds in connected speech. If not careful, the meaning of an utterance may not reach a listener, creating problems in the communication process. For instance, if someone were to speak with a strong non-English accent, with a lot of mistakes both grammatically and phonetically, the intelligibility of this speaker would be affected.

3.3.2 Auditory Phonetics.

Auditory phonetics is the "part of phonetics, speech sciences and psychology more generally that investigates how the sounds of speech are recognised, organized and classified." (Ogden, 2017, p. 201) This concept is strictly related to Intelligibility as this concept deals with how the listener will understand from an utterance. Munro and Derwig (1995) defined it as "the

extent to which the native speaker understands the intended message" (p. 76), explaining that it is very important when understanding L2 speech, as high intelligibility will make it easier to communicate in the target language. The pair state that the concept is related to that of accentedness, namely the degree to which an L2 learner's accent differs from a particular standard, as the pair state that there is connection between how strong an accent is and how intelligible is what is being produced by the speaker. On the other hand, comprehensibility is defined by Munro and Derwig (1995) as a judgment of how easy or difficult an utterance is to understand from the perspective of a native speaker.

The previous definitions differ from Smith and Nelson (1985, p 334), who give a brief definition of the two plus adding the concept of interpretability, defining intelligibility as "utterance recognition," comprehensibility as "utterance meaning," and interpretability as "meaning behind the utterance". The information provided by the pair could be considered outdated by this point as it was stated in 1985, but allows us to see the similarities and how some concepts evolve with time while others maintain their meaning, such as intelligibility having a similar definition between both authors.

Munro and Derwig (1995) state that despite the common belief that accent is very important when it comes to intelligibility, there is evidence that strong accents do not equal less intelligibility. On the contrary, Kang et al. (2018), as well as Jung (2010), claim that intelligibility is also affected by factors such as attitudes, previous experiences, and even the culture of the listener.

Another element to consider is the tempo of speech, which according to Ogden (2017), consists of the speed at which a speaker stresses syllables in an utterance. When faced with native speakers, learners of the language might find it difficult to understand what they are saying because of how fast they speak. Indeed, this is considered as one of the hardest things when learning English (Yurtbaşı, 2015). It is of paramount importance that, as the author states,

teachers slow down their pace so that students benefit from their teaching, and it has been shown that "speech rate does affect the intelligibility of the speaker's utterance." (p. 1). In terms of intelligibility, the author states that the tempo and its loudness (intensity) are key factors. The latter term relates to sound quality and it is another aspect to bear in mind; as part of this study aimed for the analysis of speech from influencers, it is necessary to define what sound quality is, in order to explain why the material for the study was selected.

3.3.2.1 Sound quality

Several elements or features of sound can be addressed when sound quality is being studied; according to Raake (2016), "Two sounds can be distinguished based on their loudness, pitch, timbre, spaciousness and duration. The totality of these features describes the (...) character of a sound (...). In essence, this reflects a multidimensional view of auditory events".

According to Raake (2016, p. 6), several features (such as loudness or duration) can be considered when differentiating two sounds as they describe the "character of a sound", which would signify a multidimensional view of sounds. Furthermore, as the author explains: "Often times a so-called ideal-point of the multidimensional feature space can be found that represents the best-possible sound quality." (p. 6). Therefore, "good sound quality" in this sense is understood as an average point within the different features of sound, that allow a high level of intelligibility in the audiovisual material. Thus, taking this definition into account, some of these features were considered in this study as key factors for the selection of videos with good enough sound quality.

The importance of taking sound quality into consideration for this study is explained by Newman & Schwarz (2018), who carried out an experiment that involved identical recordings of conference talks and radio interviews about science, but one of them was in low quality and one of them was in high quality. Their experiment showed that the listeners who were exposed

to the low quality recording perceived it negatively, while those who listened to the high quality version had overall positive impressions. This shows that audio quality does indeed influence listeners.

3.3.3 Acoustic Phonetics.

Acoustic phonetics deals with the physical and acoustic characteristics of sounds. According to Ogden (2017), acoustic phonetics uses technology to transform sounds and the changes of air pressure into static images that allow their analysis. This field of phonetics emcompasses the study of sounds through different charts and images like waveforms and spectrograms. One definition given by Ogden about waveforms is the following:

> Waveforms are a kind of graph. Graphs have an x-axis, which runs horizontally, and a y-axis, which runs vertically. In waveforms of speech, the x-axis represents time and is usually scaled in seconds or milliseconds, while the y-axis shows (to simplify a great deal) amplitude, a representation of loudness. (2017, p.32)

According to the author, spectrograms convey more information than waveforms and they are strictly related to the research topic so they will be discussed at length in the next section.

3.3.3.1 Spectrograms

According to Hagiwara, a spectrogram is “a visual representation of an acoustic signal” (2009, para. 6) that expresses amplitudes at various frequencies. A spectrogram usually displays “degrees of amplitude (represented light-to-dark, as in white=no energy, black=lots of energy), at various frequencies (usually on the vertical axis) by time (horizontal).” (Hagiwara, 2009,

para. 6). Gavaldà (2021) expresses that spectrograms are computer generated charts of a sound frequency or signal.

According to Gavaldà, N. (2021), a spectrogram is a tool that can be utilized for both representing and analyzing visual and audiovisual signals; however, it is worth noting that Gavaldà does not explain how visual signals can be analyzed and represented, and this fact is not mentioned by any other author. As used by Hagiwara (2009), Ogden (2017) and Ladefoged (2021), the amplitude shown on a spectrogram carries characteristics of the sounds produced by a speaker, such as the point of articulation of the sound.

In this chapter, the field of Phonetics and Phonology was covered in its three main parts: Articulatory Phonetics, Auditory Phonetics, and Acoustic Phonetics. In Chilean schools, this field is often left behind, even though the teaching of the English sounds is incorporated in the scholar curriculum. An in-depth review of the system, as well as the students' level of English, will help clarify why Phonetics and Phonology are being overlooked.

3.4 Chilean education system.

The Chilean education system is divided into four stages: kindergarten, primary, secondary, and tertiary education. For the purpose of this research, kindergarten and tertiary educational level will not be taken into account, as we will be focusing from 7th to 12th grade only, which corresponds to primary and secondary education. The reason for the focus of the research on 7th to 12th grade is that students share a similar level of English. In fact, 7 out of 10 students from 11th grade do not achieve the level of English expected for 8th grade, which is A2 (Bustos, 2019).

Educación Básica or primary education, which goes from first to eighth grade, is oriented towards tackling physical, affective, cognitive, social, cultural, spiritual, and moral dimensions (Unidad de Currículum y Evaluación, 2012).

Educación Media or secondary education, which goes from ninth to twelfth grade, expands and deepens on the formation received earlier in primary education, developing necessary skills to practice active citizenship (DFL 2, Diario Oficial de la República de Chile, art. 20, 2010). These two levels are mandatory, of which the State must finance a free system to ensure access to them to all the population (Decreto 100, 2005).

Equally, education is provided by four types of establishments: i) public; ii) private subsidized; iii) private; and iv) business corporations (Darville & Rodríguez, 2007). Before the big change in the public education system, public institutions were financed by the State, alongside contributions by the corresponding local municipalities. Private subsidized establishments are also financed by the State, but the extra contributions are paid by the students' families. Private organizations are financed only by families by paying an enrollment fee, as well as monthly fees. Finally, business corporations are financed by public resources given through covenants (Darville & Rodríguez, 2007).

For this study, only public and private-subsidized schools will be considered for the pedagogical suggestion, as the business corporation institutions (called technical schools) have a job-oriented curriculum and private schools have a solid English teaching curriculum, ensuring that the majority of their students get an English certification; indeed, more than 80% of students with high SES get to certify their level, and mostly all high SES students attend private schools (Bustos, 2019).

The considered establishments coexist within the educational system but differ on their core elements, as they have a different level of English, socioeconomic composition, and curricula. Ultimately, all these differences impact on the students' English proficiency. A general notion of the Chilean level of English is required before deepening on socioeconomic

composition and curricula, as the last two themes are reasons for the level of English proficiency.

3.4.1 Chilean level of English.

Worldwide, Chile ranks 37 out of 100 countries, according to the 2020 EF English Proficiency Index. Chile scores 523 points out of 800 and is in the top 3 of Latin American countries, with Argentina first in rank, with 566 points, and Costa Rica second in rank, with 530 points. Nonetheless, Argentina is the only Latin American country with a high proficiency, corresponding to a B2 CEFR level where Chile is considered to have a moderate proficiency, corresponding to a low B2 CEFR level.

Table 3. EF EPI averages scores and English proficiency of every continent.

Region average	EF EPI score	Proficiency
Europe	550	High
Asia	492	Low
Latinamerica	480	Low
Africa	497	Low
Middle East	441	Very low

Source: EF EPI 2020.

As it can be observed, Europe is the only country with an average high English proficiency. Moreover, despite that Latin America has a low proficiency, Chile manages to get a moderate level. However, this data from EF EPI considers adults only, but it provides a broad context of Chile's overall English proficiency.

Concerning Chilean students, the results present a considerable gap between students from private schools and those from private-subsidized and public schools. 7.340 eleventh-grade students from 137 schools took an English test, based on CEFR standards, that ranged from 0 to 100 points, evaluating listening and reading skills with 25 questions each. The results showed that the average was 51 points and that only 32% of students achieved the basic and intermediate level; in other words, around 7 out of 10 students do not achieve the level expected for 8th grade (Bustos, 2019). By 8th grade, students are expected to achieve a level of A2; upon graduation, students are expected to achieve a B1 level (MINEDUC, 2016). MINEDUC further explains that by the end of 10th grade, students should have a lower B1 level. All in all, 7 out of 10 11th grade students do not even achieve an A2 level. On the other hand, the English test mentioned earlier gave insightful information correlating students' socioeconomic status with their English proficiency.

Table 4. English level gap by socioeconomic status.

Socioeconomic status	Percentage of students between A2 to B1 level
High	85%
Upper middle	54%
Middle	39%
Low middle	14%
Low	9%

Source: Bustos (2019).

It is worth noting that approximately 1 student out of 10with a low SES has a level between A2 to B1 as opposed to around 8 out of 10 with a high SES.

This is further noted by the SIMCE results in 2012. SIMCE is a standardized test applied nationally to every school that "seeks to assure that all students achieve a similar level in educational quality and equity." (Vallejos, 2012, p. 6). Out of the top 100 schools with better

results in English, 99 were private and 1 was private-subsidized; however, in the end, all of them were private (Alarcón, 2013). From this test, only 18% of students achieved a satisfactory level in the test, enabling them to get international certification (Pérez, 2013). Meanwhile, 83,3% of students with a high SES got to certify their level as opposed to 0,8% of students with a low SES. On the whole, students' SES impact their English proficiency. Considering the SIMCE results, private schools have the bigger lead while state schools are yet to show their satisfactory results. To explain these results, schools' curricula give some insights, as it is shown that high class students have more English lessons per week than that of low class students, correlating hours per week with English levels of proficiency. However, the amount of hours per week is not the only factor. Almost a third of students claimed that they had more than 3 hours per week (and thus scoring 11 points more than the average), factors such as the teacher actually speaking english while in class and the level of certification of the teacher are also relevant. Students whose teachers speak English in class scored 11 more points in the test than those whose teachers spoke the entire class in Spanish (59 vs 48 points). Conversely, around 40% of teachers had A2 certification, which is a worry considering that the required CEFR level for in-practice teachers is B2 and in-service teachers C1.

3.4.2 Curricula.

For many Latin American countries, the ideal target would be for students to graduate at 18-years-old with a minimum of B1 on the CEFR scale (Hernández-Fernández & Rojas, 2018). Notwithstanding, this is expected to be achieved by having around two to three 45-minute lessons per week. As stated by the authors, three 45-minute lessons per week equals 78 hours of instruction per year, which is equivalent to three weeks in an intensive language school; ultimately, this is not enough to "develop the understanding, vocabulary, system knowledge, and the spoken proficiency that is required to policymakers' targets and goals." (Kamhi-Stein et al, 2017, as cited in Hernández-Fernández & Rojas, 2018, p. 33).

As three 45-minute lessons are not getting the expected results, a broad comparison can be made among Chile, Latin America, and OECD countries in order to correlate hours per week

and level of proficiency. For Chile, it will also be provided information about hours per week given in public and private-subsidized schools and private schools, with the purpose of checking any difference between them. The table below shows foreign-language hours given per week.

Table 3. English hours per week: a comparison among Chile, Latinamerica and OECD countries.

	Hours per week
OECD average	3.6
Latin American average	3
Chilean average: state schools and private-subsidized	3
Chilean average: private schools	8

Source: OECD (2020); Fernandez, Rojas & British Council (Mexico) Staff (2018); De Améstica (2013).

Latin America and Chile are more than half an hour per week behind the OECD average. Bear in mind that, for OECD countries, the term is foreign-language lessons instead of English ones. This is because not only English is taught in non-English speaking countries, but many other languages as well, but the weekly hours are the same for any language being taught. It is also worth noting the wide gap between the public and private-subsidized average and the private average. Mary Jane Abrahams, president of the English Teachers Association in Chile, claims that public schools have 3 hours a week while private schools have 8. However, according to Pavez (2008), the vast majority of non-bilingual private schools have between 3 and 10 English lessons per week. The Ministry of Education (2018), in its national study plan, sets 3 minimum English hours per week. Nevertheless, in the curricular bases by the Ministry of Education (2018), it is stated that establishments can develop their own plans and programs as the objectives proposed in the bases do not use the total school time, reassuring the value of plurality and flexibility of curricula options.

As previously mentioned, students with a high SES have a much greater percentage of English certification, which is more than 80%, than those with a low SES, which is below 10%. Coincidentally, students with a high SES mainly attend private schools, students in the mid spectrum attend private-subsidized schools, and those in the lower spectrum attend public schools, thus correlating socioeconomic status, the type of school students attend, and English level of proficiency. To further demonstrate this, the following section will deepen on segregation by socioeconomic status and the hindrance for learning outcomes.

3.4.2.1 Socioeconomic Composition.

Chile has high levels of inequality as expressed by its Gini index of 0.47, placing Chile in the 24th spot in terms of inequality of a total of 159 countries with available data (Lambeth, Otero, & Vergara, 2019). The Gini index is "a measure of the distribution of income across a population" (Hayes, 2021, para. 1), which reflects a country's economic inequality. It ranges from 0 to 1, with 0 representing perfect equality and 1 representing perfect inequality. Moreover, the income of the 10% richest people in Chile are 26 times higher than the 10% poorest people (González, 2018). This, ultimately, is reflected in the education system, as seen in Table 1 and Table 2 below where it is shown that the richest and the poorest students are mostly segregated based on their socioeconomic status (private and public, respectively).

Valenzuela, Bellei & de los Ríos (2010), as cited in Bellei (2013), put together a table depicting the Duncan Dissimilarity Index of the 30% most vulnerable students' socioeconomic level. This D index is "generally used to describe the overall extent of uneven distribution of two populations." (Sakoda, 1981, p. 245).

Table 1. National school segregation of students of low socioeconomic level. Duncan Index: 30% of less socioeconomic level.

	2005	2006	2007	2008
4th grade	0.53	0.53	0.54	0.54
8th grade			0.53	
10th grade		0.50		0.50

Source: Valenzuela, Bellei, & de los Ríos (2010), as cited in Bellei (2013).

Valenzuela (2008), as cited in Bellei (2013), concluded that Chile has the highest Duncan Index for the 30% poorest students among 57 countries considered; along with Thailand, these two countries are the only ones that their D index is higher than 0.50 where most countries have a D index of 0.3 and 0.4. The same applies to the richest students: of the 57 countries, Chile is the most segregated when it comes to the 30% of students with the highest socioeconomic status, with numbers greater than 0.6; i.e., about 60% of the richest students would need to be moved to other establishments in order to reach equality.

To further demonstrate segregation by income, González (2018) put together a table of the percentages of students enrolled in public, private-subsidized, and private schools with their corresponding socioeconomic status.

Table 2. Percentage distribution of students from 10th grade by socioeconomic status (SES) and administrative dependency, 2013.

Students			
SES	Public	Private-subsidized	Private
Low	14.3	5.4	a
Low middle	18.4	16.3	a

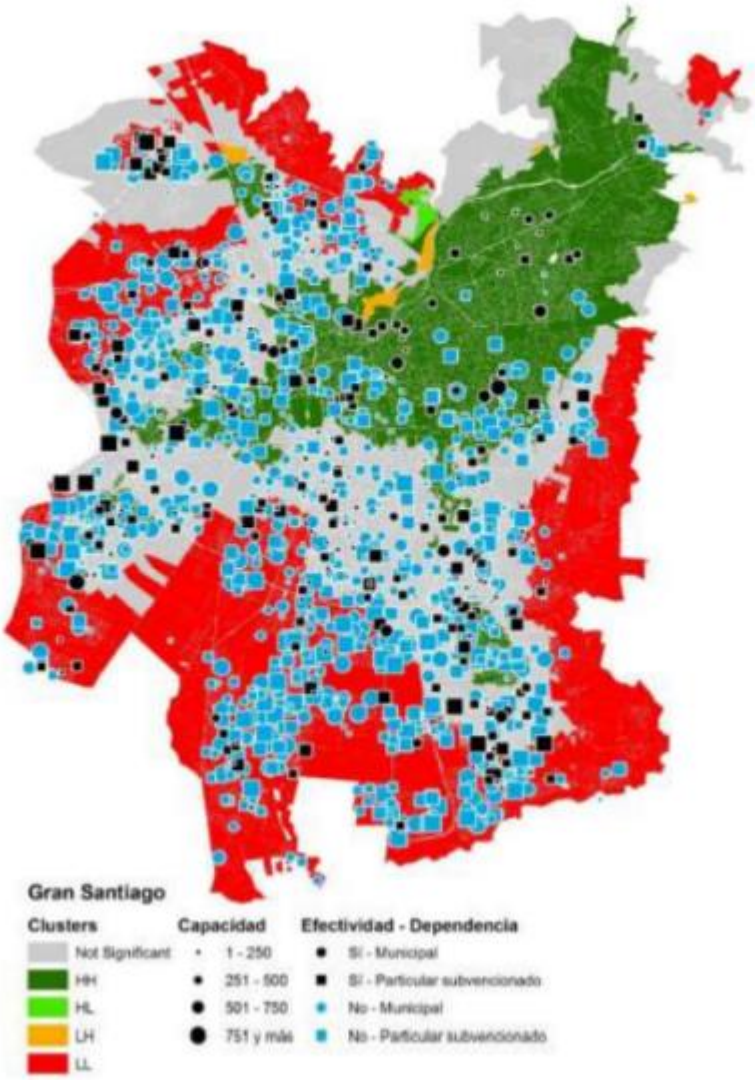
Middle	5.4	18.9	a
Upper middle	1.2	11.5	a
High	a	0.5	7.6
Total	**39.3**	**52.7**	**8**

Source: Agencia de Calidad de la Educación (2013), as cited in González, (2018).

From table 2, we can see the crystal-clear segregation from the private sector, where only 8% of students are enrolled, of whom the vast majority come from a high class; thus, the private education sector is directed towards high classes. Private-subsidized schools are directed towards middle classes; if all of them are put together, the percentage of middle-class students enrolled in private-subsidized establishments is 46.7%. Finally, public schools are directed towards the low middle and low classes. Concerning the "a" value, it was designated by González: the values there were so low that he decided to assign them the letter 'a'.

It is also important to mention that not only schools are segregated, but territories as well. Rodríguez (2016) determined the access students from *Educación Básica* have to "effective establishments." By effective, the author refers to schools that get at least 30% of their students to achieve an acceptable level in reading and mathematics for 4th grade, according to MINEDUC and learning standards. He concluded, as it can be seen in Figure 1, that the cluster HH — which shows a high concentration of students with high accessibility to effective schools in the same territory in the capital of Chile, *Santiago* — encompasses territories with a high SES, and also including less economically disadvantaged cities like *Estación Central, la Florida, Maipú, Conchalí,* and *Quilicura.* The cluster LL — which shows the concentration of students with a low coverage of effective establishments — equals 40% of the total student population from *Santiago,* of which more than 70% are part of the groups D and E (the poorest groups in the socioeconomic scale), and that are concentrated in peripheral areas with high socioeconomic segregation.

Figure 1. Santiago's map of clusters on access to effective schools. The type of school is represented by its form (public are circles, private-subsidized are squares). The color indicates effectiveness (black are effective schools, blue are not). The size represents the capacity of schools.



To conclude and to demonstrate how segregation impacts on students' learning outcomes, segregation by income has "negative impacts on children's future educational and economic outcomes." (Owens, 2016, para. 8). For instance, according to Boser (2017), students attending high-poverty schools have a 68% chance of graduating, whereas students attending low-poverty schools have a 91% chance. Additionally, the author further states that students

across the socioeconomic spectrum are better prepared for post-secondary success when they have been educated in diverse schools and have learned alongside peers who come from all walks of life. In the case of Chile, students who come from vulnerable families have a much higher chance of dropping out of school than those with a better financial status. As reported by Olivares (2018), a study conducted by Luis Gajardo, sociologist from *Universidad Central*, showed that 14% of students from 15 to 19 years of age who were in the lowest spectrum dropped out of school, whereas only 2% were in the upper spectrum.

4. Methodology

The aim of this study was to design a website that contains videos of influencers, in which challenging English consonants for Spanish Speakers of English are present. The purpose of designing it was to create a source of free-access material for EFL teachers to use in the teaching of pronunciation and listening skills in Chilean classrooms.

For the development of this work, a specific selection of consonants was considered, a number of influencers were chosen based on several parameters, and then the sounds picked were extracted from their videos to be analyzed using the speech analysis software Praat. Audiovisual material is available to watch on the website as a complement to the observation of the pronunciation of the sounds. Additionally, all the audiovisual material, as a corpus, will be available in Annex 1.

4.1 Selection of English Sounds.

The list of sounds contemplated consonants that pose a challenge to the Chilean learners of English based on what Coe (2001) proposed, which included sounds that are not part of the Spanish phonetic system, which are /v, ð, z, ʃ, ʒ, ʤ, h/. To this list, the phoneme /θ/ will be added, as it is a sound that is present in the accent from Spain, and it is an acceptable substitute for the English /θ/ (Lira & Ortiz, 1982), but is not in other Spanish accents. What Coe proposed goes in accordance with what is proposed by CAH (Lado, 1957). Consequently, the list considered in this study was the following: /v, ð, z, ʃ, ʒ, ʤ, h, θ/.

4.2 Selection of Speakers and Samples

The selection of speakers and audiovisual material for the study was done through the consideration of several parameters, namely the place of birth and upbringing of the influencers,

their accents, the intelligibility of their speech, and the possibility of seeing them articulate the sounds.

The very first factor to contemplate was to ensure that the speaker or content creator effectively represented the role of an influencer (according to the definition of the term explained in the 'Influencers' section of this study). Likewise, the nationality of the speakers, or more specifically, the status of being a native speaker of the language or not, was another factor we considered. The study only contemplated native speakers to eliminate the variable of L1 interference; in other words, if the mother tongue of a speaker is English, there should not be any foreign sound influencing the speaker's pronunciation.

Regarding the influencers and their material, two aspects were considered: accent and the intelligibility of their speech. For the first one, only SSBE (Standard Southern British English) and SAE (Standard American English) were taken into account. For the second, the main parameters for intelligibility were tempo of speech, the sound quality of their videos, and possible variations of sounds in their speech. These parameters were considered to the extent that their presence or influence allowed the correct understanding of the speakers' speech. For instance, even if any problems regarding sound quality were present in the videos (such as background noise, saturation of sound, echo, etc), the videos were still selected as long as what the speaker was saying could be understood.

Concerning the selection of the audiovisual material, two main elements were considered. First, the presence of the expected sounds: it was strictly necessary that the speaker produced, within their speech, at least one of the English sounds aimed for this study; otherwise, the material was useless for the purpose of the research. Secondly, the possibility of watching the face, specifically the mouth and lips, of the speaker at the moment of producing the sounds. This is an important feature, since seeing the way the speaker pronounces the sounds allows for the visual confirmation of the place and manner of articulation.

4.2.1 Speakers Selected

The speakers selected for this analysis were 10 in total, 5 born in the United States and 5 born in the United Kingdom. The speakers selected were the following:

- CDawgVA - United Kingdom - CDawgVA - Youtube
    - https://www.youtube.com/channel/UCPsZ_0SkFdi551iYTG04R2g
- Daniel Howell - United Kingdom - Daniel Howell - YouTube
    - https://www.youtube.com/channel/UCGjylN-4QCpn8XJ1uY-UOgA
- Emilia Clarke - United Kingdom - @emilia_clarke - Instagram
    - https://www.instagram.com/emilia_clarke/?hl=es-la
- Emma Watson - United Kingdom - @emmawatson - Instagram
    - https://www.instagram.com/emmawatson/?hl=es-la
- Tomska - United Kingdom - TomSka & Friends - Youtube
    - https://www.youtube.com/channel/UC3tMH8u6yG3mSxi-qpfmpkA
- Artsy Madwoman - United States - Artsy Madwoman - YouTube
    - https://www.youtube.com/channel/UC1DVxeBCg2Omad3JFVXh2_w
- Elise Ecklund - United States - Elise Ecklund - YouTube
    - https://www.youtube.com/channel/UCUsCq3Hq5haa6tTph41hpJQ
- Emmymade - United States - Emmymadeinjapan - Youtube
    - https://www.youtube.com/user/emmymadeinjapan
- Markiplier - United States - Markiplier - Youtube
    - https://www.youtube.com/channel/UC7_YxT-KID8kRbqZo7MyscQ
- Rybonator - United States - Rybonator - YouTube
    - https://www.youtube.com/user/TheFluffinAmazin

4.3 Procedure

The research was conducted through the analysis of spectrograms from audio sample chunks retrieved from influencer's media with the help of the webpage *Loader* and the audio editing program *Audacity*. In the first place, *Loader* was used to download the entire audio of an influencer's video in WAV format. The reason for using WAV is because this type of audio format is uncompressed; therefore, it facilitates the analysis as it does not eliminate characteristics of the audio nor does it simplify the image the spectrogram shows when analyzing it.

Once the full audio was downloaded, it was imported to the audio software *Audacity*, where the chunks containing the phonemes were isolated; then, the isolated chunks were exported as a WAV file. The reason *Audacity* was chosen for this project is its status as a free and open-source digital audio editor software, which makes it accessible to everyone. Furthermore, it is easy enough to use, even for people who are unaccustomed to working with audio editing software.

Having the isolated chunk in WAV format, it was imported to the speech analysis program *Praat,* created by Paul Boersma and David Weenink. The program shows the sound wave and spectrogram of the audio imported. For the purpose of this study, the sound wave was discarded as it did not provide useful information to determine neither the phonemes nor the manner of articulation of the sounds. On the other hand, the spectrograms contain an image created between the frequencies of 0 and 8,000 Hz. The reason why this range was chosen is that some phonemes are recognizable in the 8,000 Hz range, as it is the case of fricatives and plosives. Moreover, the dynamic range was set to 50 dB as it is the standard value for analysis of speech recordings according to Boersma and Weenink, creators of *Praat* (2003). In relation to the formants, the configuration chosen was to have a formant ceiling of 5,000 Hz for male speakers and 5,500 for female speakers as well as order the program to display a total of 5 formants on the spectrogram based on Boersma and Weenink (2003).

The function “extract visible spectrogram” in the spectrum options was used, as it shows a different type of spectrogram than the default one. This new spectrogram (henceforth “the spectrogram”) is divided into intervals of 1000 Hz, allowing for a more thorough analysis of formants. Only the first (F1), second (F2), and third (F3) formants will be considered, as these three are enough to determine the segmental features detected in the first instance; however, the fourth formant (F4) will be considered if necessary. A screenshot of each spectrogram was taken and imported to the image editing software *Paint* to manually highlight the formants using the white pencil with the second width option; these highlights were based on the automatic option of *Praat* called “show formants.” Also, the presence of the sound analyzed was highlighted with a red square. The program *Paint* was used because of availability, as it is a free-to-access program pre-installed in any relatively modern computer that allows the user to freely edit images.

Furthermore, the influencers’ videos were downloaded as a complement, as they provide useful information on the manner of articulation of sounds. Using the website *y2mate*, the full videos were downloaded in their highest quality possible and then imported to the video editing software *Vegas Pro*. There, the chunk of interest, the portion of the audio containing the required phoneme, was cut out and exported in the same format and quality as the original file.

This research consists of determining the quality of the segmental features selected through the examination of their formants. The videos and spectrograms are available as a free corpus on a website created for educators to use in the teaching of pronunciation and listening skills of English as a Foreign Language.

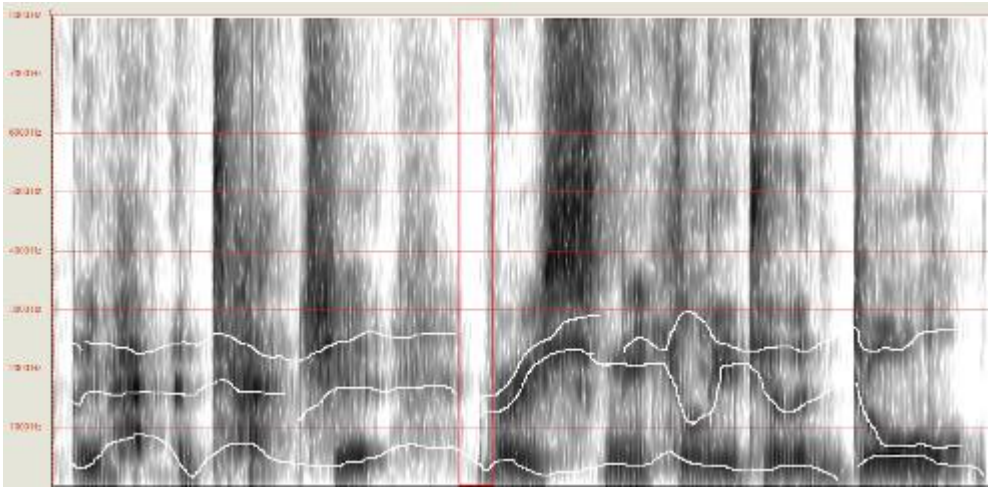
Finally, the analysis is shown both in the thesis and on the website created to show the sound analysis and host the audio samples used for the research. After this, the results will be discussed on how they can be applied for the teaching of the English sounds in the Chilean classroom.

4.4 Data Analysis

In this section, all of the data obtained was analyzed basing it on what is shown by Rob Hagiwara on his webpage “Monthly Mystery Spectrogram,” where the author showed and explained how segmental features look on spectrograms, how to recognize them, and what the formants indicate. Rob Hagiwara was the first source of information when analyzing the spectrograms obtained and two secondary authors were consulted: Ogden (2017) and Ladefoged (2010). The sounds to analyze were organized according to their place of articulation, starting from the lips to the glottis.

4.4.1 Voiced labiodental fricative /v/

Figure 2. CdawgVA - “they… attach it very sneakily to the door.”



Source: CdawgVA (2021, 7:49-7:51).

Voiced fricatives — v, ð, z, ʒ — are very similar to their voiceless counterparts — f, θ, s, ʃ (Ladefoged, 2010). Moreover, fricatives might get confused for background noise, as by definition they "involve an occlusion or obstruction in the vocal tract great enough to produce noise (frication)" (Hagiwara, 2009). Although the sound is short, the voicing can be seen by the striation (the vertical lines, which constitute formant structure and voicing) and the F1 has more

energy than any other place, thus excluding the possibility that this might be a /f/. Additionally, the speaker said /ð/ at the beginning; even though it has a similar F1 height, the F2 is higher in /ð/ than /v/, a key characteristic when differentiating these two sounds in spectrograms.

Another important aspect to mention is that the place of articulation is not visible by watching the video.

Figure 3. Daniel Howell - "that I've never shared before."

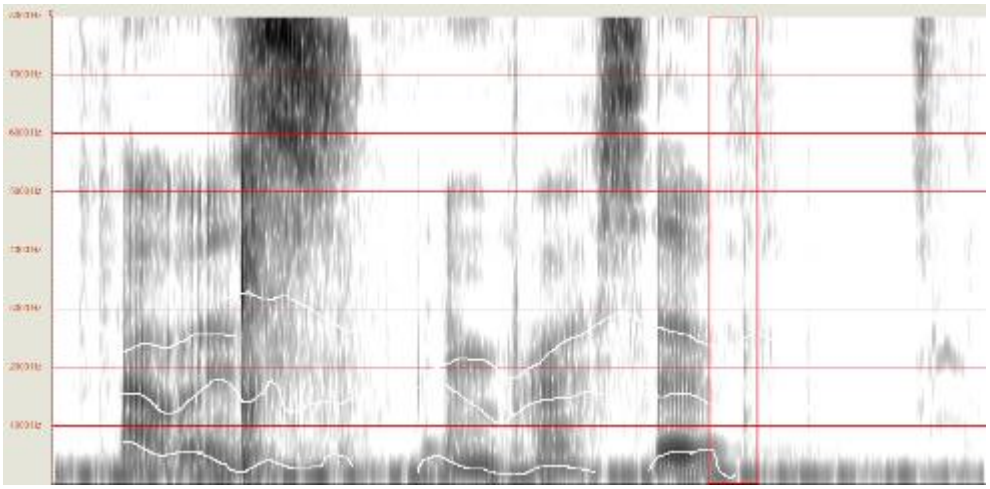
Source: Daniel Howell (2017, 0:03-0:05).

For the first /v/, it has an unusually high F2, but a normal F1. Also, it has an unusual concentration of energy near 8,000 Hz, which indicates that the sound has more frication. Additionally, the energy in F1 is not as dark as the second one, indicating that this /v/ is slightly devoiced. The sound /ð/ is discarded in this case, as it is present at the beginning, having a higher F2 and a similar F1, the same scenario as seen in Figure 2. The /f/ here is discarded due to the presence of , despite the fact that there is energy near 8,000 Hz. The place of articulation is clear by watching the video.

For the second /v/, the striation indicates clear voicing, F1 is a bit higher than usual, and F2 seems normal. Also, note that the energy near 8,000 Hz is not as darker as in the first one.

As the sound only has vowels around it, it is fully voiced and cannot be mistaken for a /f/. The place of articulation for this /v/ is also clear.

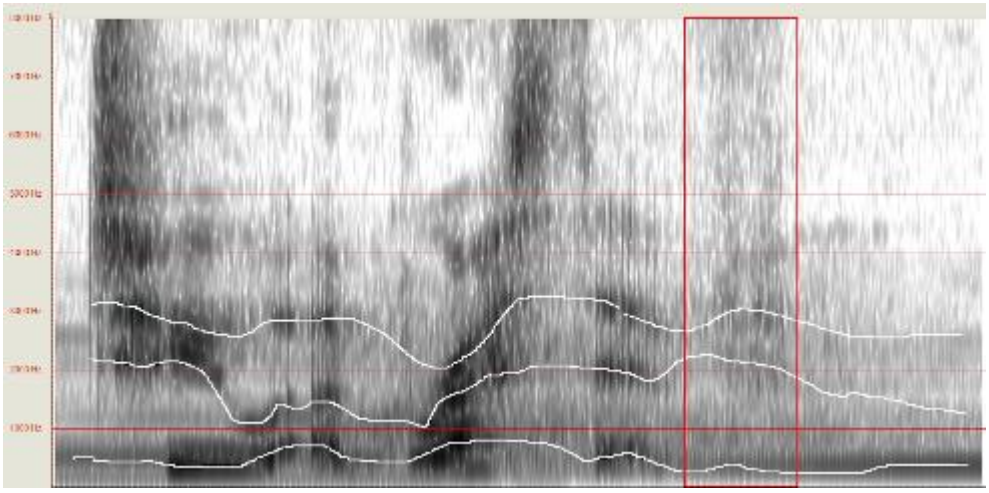
Figure 4. Emilia Clarke - "that was impressive."



Source: Vanity Fair (2019, 0:21-0:23).

Little energy is present due to the phoneme being in final position, added to the fact that the speaker talks softly here. The sound has a low F1 and F2, and striation yet not too much, indicating devoicing. Striation here excludes the possibility of this being a /f/ and the low F2 here discards the /ð/. Also, there is some energy near the 8,000 Hz range, so the /v/ has more frication, and the place of articulation is not at all clear in the video.

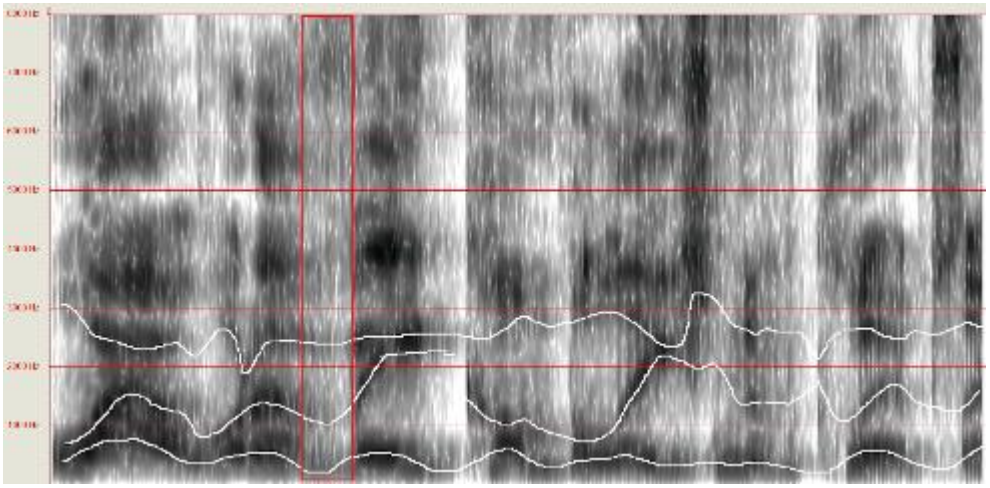
Figure 5. Emma Watson - "too aggressive."



Source: English Speeches (2017, 2:16-2:18).

Again, the /v/ is in final position, losing voicing. The background music and the type of microphone the speaker used most likely affected the spectrogram. Despite that, striation and transitions can be seen, having an unusual F2, most surely influenced by the vowel preceding it. As the /v/ is in final position, there is more frication than usual, that is why there is energy in the 8,000 Hz range, and there is little energy near F1. By watching the video, it can be seen that the pronunciation of this /v/ is similar to a /f/. Despite that the sound is in final position, the place of articulation can be clearly seen.

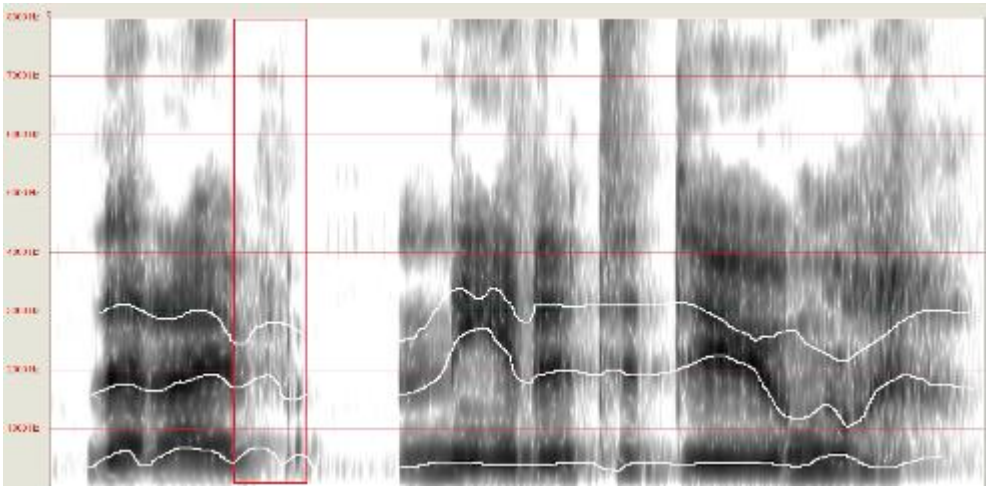
Figure 6. Tomska - "wow, what a vague compilation of clips."



Source: TomSka & Friends (2021, 2:13-2:16).

Both audibly and visually, this /v/ looks good. Both F1 and F2 are low and both striation and transition are clear. There is some energy near 8,000 Hz, probably because the speaker plays background music or that there is more frication. From 3,000 Hz downwards, F1 and F2 are low and confirm this is a /v/. Although it is a visually clear /v/, the place of articulation is not clear in the video.

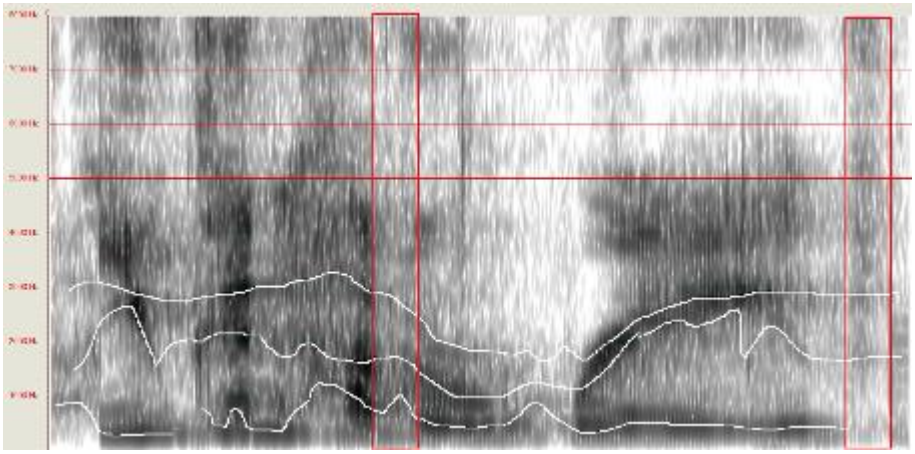
Figure 7. Artsy Madwoman - "that I've needed to do."



Source: Artsy Madwoman (2021, 0:11-0:13).

In this case, /v/ has little energy, especially near F1, indicating that the sound is slightly devoiced. However, /v/ has striation and less frication, so that discards the /f/. The sound /ð/ might get confused with /v/, at least visually, but /ð/ tends to have a lower F1 and a higher F2 than /v/. In this case, the speaker pronounces /ð/ at the beginning, and both /v/ and /ð/ have similar formant heights. Despite the visual difficulties, the video offers clarification, as it can be seen that the phoneme is a labiodental one, confirming this is a /v/.

Figure 8. Elise Ecklund - "You can have a rave."



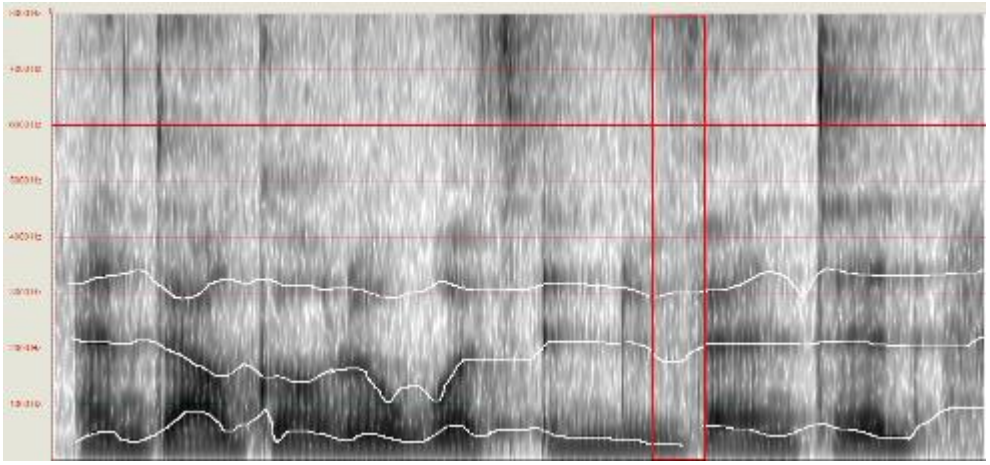
Source: Elise Ecklund (2021, 7:15-7:17).

A lot of energy and noise surrounding the first phoneme here makes it hard to determine if it is really a /v/. The surrounding phonemes near this /v/ are all voiced, thus it excludes the possibility of this being a /f/, apart from the fact that there is striation. The high F1 could discard the possibility of this being a /ð/. As there is background music, that possibly affected the spectrogram, making it hard to visually determine if this is a /v/. The place of articulation is clear in the video, thus confirming this is a /v/.

The second /v/ has little energy in F1 and throughout all frequencies in general, except from 3,000 Hz upwards. This means the /v/ is devoiced and has characteristics of frication similar to a /f/ due to the /v/ being in final position. It is not clear the place of articulation of this

/v/. As she pronounced the sound, it looked like the lips were closed together, like a bilabial sound. This is most likely an influence of a proceeding sound.

Figure 9. Emmymade - "...and dalgona was invented."

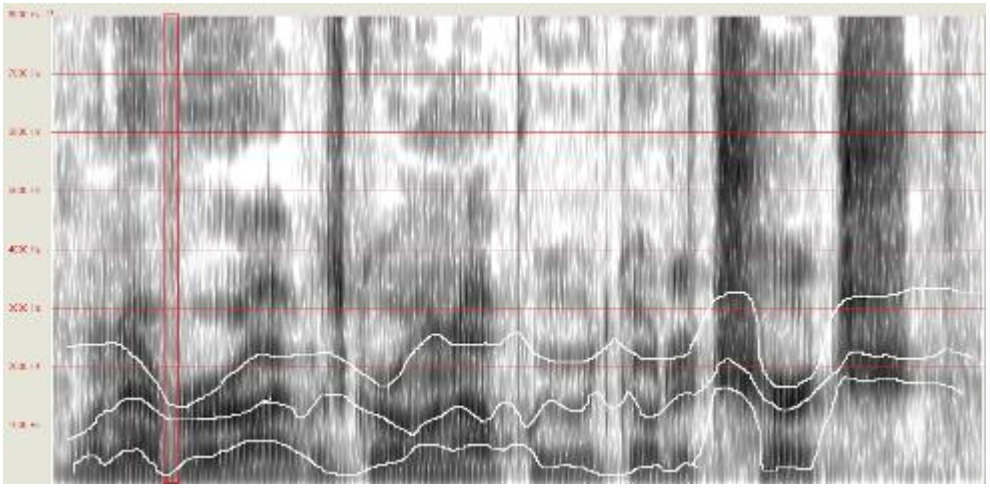


Source: emmymade (2021, 1:31-1:32).

This spectrogram is not clear regarding their formants. After the /d/ in <dalgona>, it can be seen another burst of energy corresponding to the plosive /g/; however, the F2/F3 pinch is not at all clear. In this regard, this spectrogram in particular does not provide accurate information. For the /v/ though, F1 looks normal and the F2 seems unusually high, same as the F2 of the following /e/ sound, which should be lower. There is striation and two voiced sounds, so it cannot be mistaken for a /f/, considering that there is some energy in the 8,000 Hz range.

The place of articulation is not clear, as the speaker raised her lips when pronouncing the sound, blocking the visual aid.

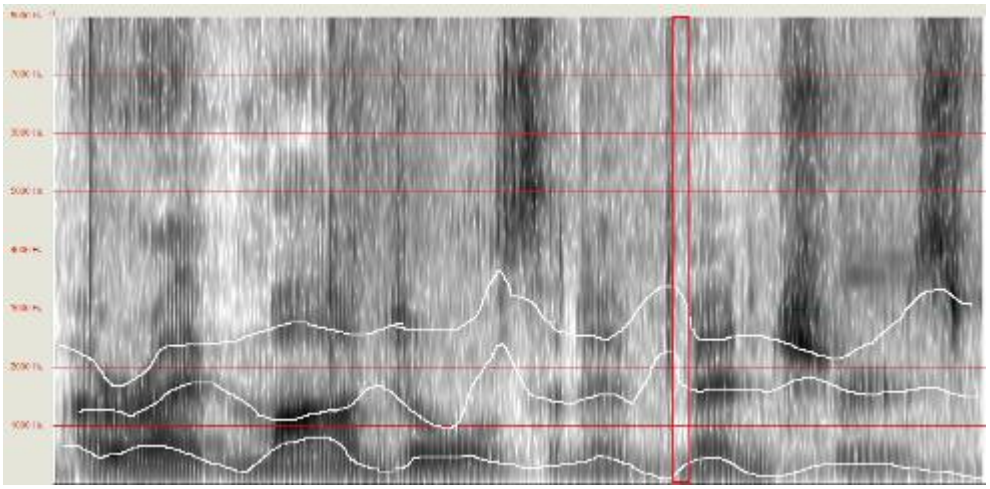
Figure 10. Markiplier - "Whenever I was growing up we would go to church."



Source: WIRED (2019, 1:26-1:28).

This might get confused by background noise, as it has energy in all frequencies. However, there is striation and transition, and an F2 that is rather low, both being characteristics of /v/. The place of articulation is also clear in the video.

Figure 11. Rybonator - "I write modules and adventures."



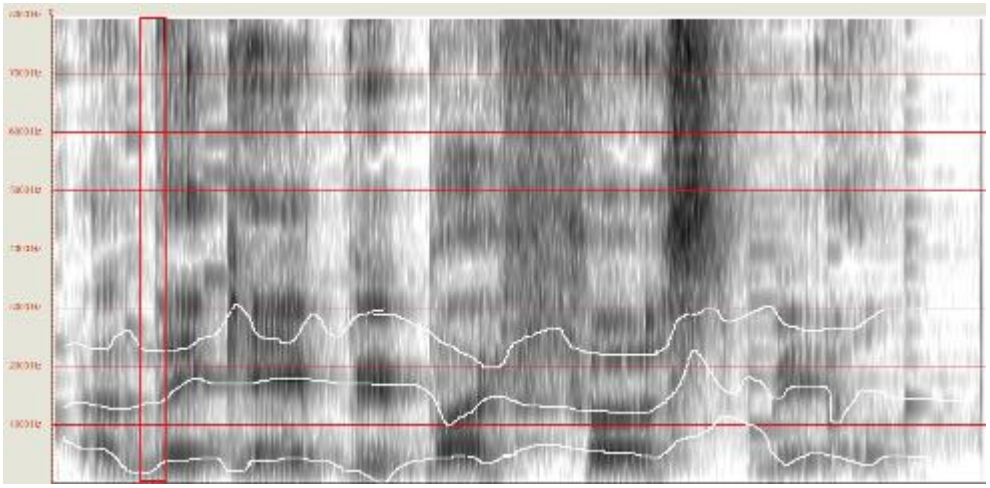
Source: Rybonator (2020, 0:39-0:40).

This is a rather short /v/, but it gives a clear examination of the /v/ because of the presence of the preceding /d/. As /d/ is a plosive, there is cessation of air, and then there is a transition

between that and the following vowel /e/. Striation and two voiced sounds surrounding the phoneme exclude the possibility of this being a /f/. The phoneme has a low F1 and a high F2, both being characteristics of /ð/. However, the place of articulation is clear in the video, confirming this is a /v/.

4.4.2 Voiced dental fricative /ð/

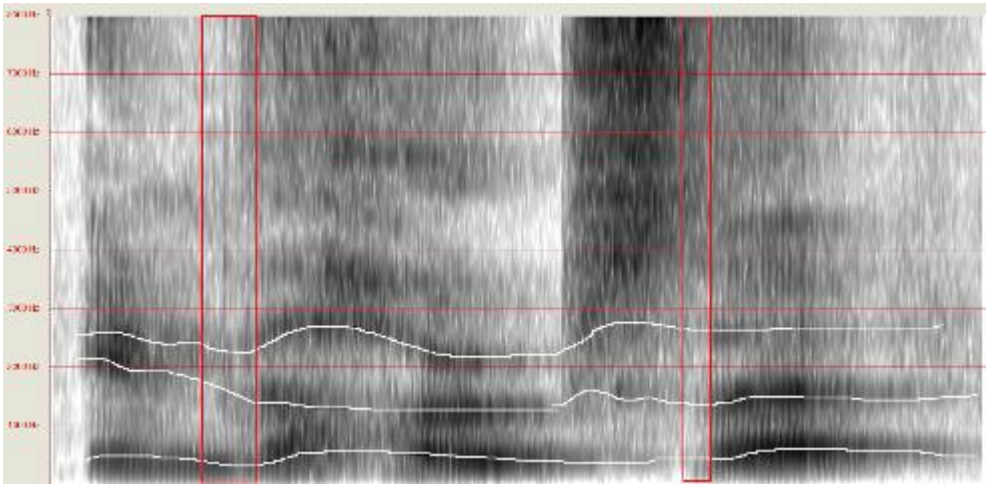
Figure 12. CDawgVA - "And they gave me my first meal."



Source: CdawgVA (2021, 0:08-0:10).

As was the case for /v/ and /f/, /ð/ and /θ/ can be distinguished mainly by striation, and adding the fact that /θ/ tends to have dark energy above 3,000 Hz whereas /ð/ does not. Additionally, /v/ and /ð/ are visually similar; /ð/ tends to have a higher F2 than /v/, and /v/ tends to have some energy near 8,000 Hz due to frication. /θ/ can be discarded due to the lack of dark energy above 3,000 Hz, but F1 and F2 are very similar to /v/. The background music could have affected the spectrogram, making it difficult to analyze this phoneme. The audiovisual material, this time, does not provide clear information about the sound, as the place of articulation is not visible; however, /v/ can be discarded by watching the video, as the speaker did not pronounce the sound as a labiodental one. It is not rare to pronounce the sound at the back of the teeth, which is the case for this speaker.

Figure 13. Daniel Howell - "In the hopes that."

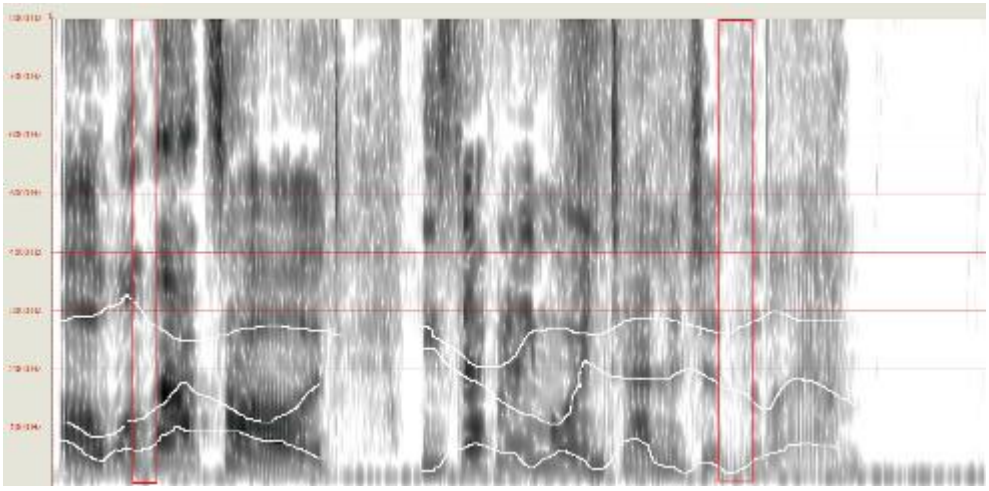


Source: Daniel Howell (2017, 0:44-0:45).

The first /ð/ looks normal in all the formants; specifically, the lack of energy above 5,000 Hz immediately excludes the possibility of this being a /θ/ or /v/. This first /ð/ is a great example of pronunciation of the sound, as the place of articulation can be clearly seen in the video of the speaker.

The second /ð/ is influenced by the /s/ sound that is before it, thus making the /ð/ have frication similar to /θ/ at the beginning, and finishing with a regular /ð/. At the beginning of the sound, there is a lot of energy concentrated near the 8,000 Hz range. Conversely, in the end, energy dissipates near that range, and the energy in all three formants (especially F1 and F2) becomes darker. Also, the tongue of the speaker cannot be seen, so he placed his tongue at the back of his teeth, but that cannot be seen in the video. This is because the /s/ is an alveolar sound, and the tongue is closer to the back of the teeth rather than in between them.

Figure 14. Emilia Clarke - "Oh well that's hard. He's definitely more intelligent than I am."

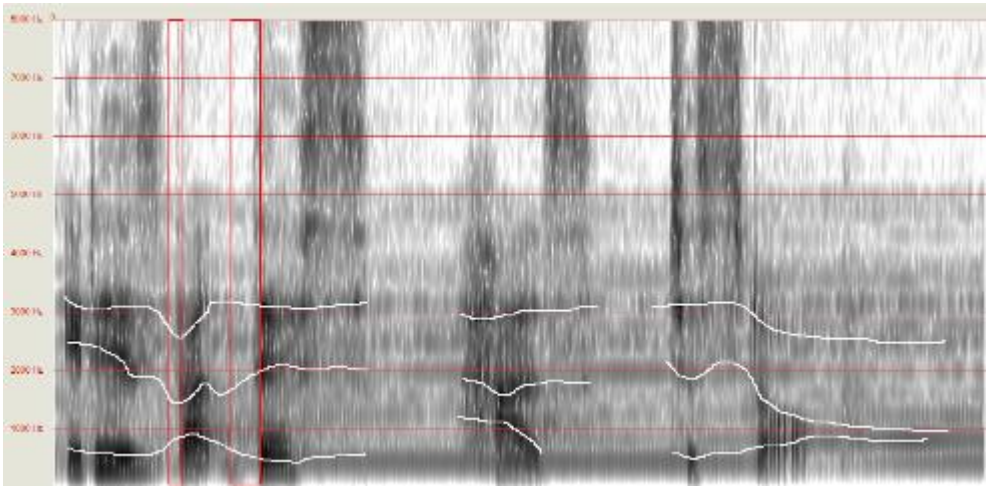


Source: Vanity Fair (2019, 1:03-1:06).

For the first /ð/, F1 and F3 are a bit higher than usual, but it could be influenced by the two sounds near it. First, there is an /l/, which is likely to be a dark one, before it and an /æ/ after it. It can be seen that F2 starts low and then rises during pronunciation. By watching the video, the speaker does place her tongue on her teeth, but it is rather fast and hard to notice.

The second /ð/ has low energy in general because the speaker is lowering her pitch and speaking more softly. All formants look normal. In the video, just as the speaker is pronouncing the sound, there is an editing cut, so the place of articulation cannot be seen.

Figure 15. Emma Watson - "It is that this has to stop."



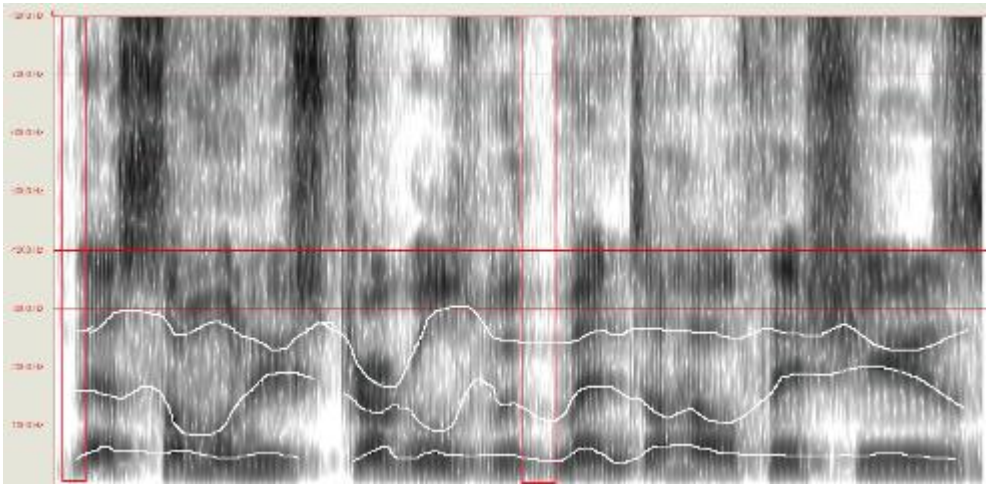
Source: English Speeches (2017, 0:36-0:38).

The background music affected the spectrogram, so there are a lot of gray areas under 5,000 Hz. Despite this, both /ð/ are clear, although the first sound has a higher F1 than the second one. They share similar formant height and are not exactly in the same place because of the influence of the sounds near them.

For the first /ð/, F1 and F2 rise because of the /æ/, which is a front-open sound. It is not possible to notice the place of articulation by watching the video.

For the second /ð/, it is influenced by the /ɪ/ after it. F1 lowers for it being a closed sound and F2 does not change that much, but it rises to around 2,000 Hz, which is normal for an /ɪ/. The place of articulation for this sound is clear.

Figure 16. Tomska - "There's someone who struggles with mental health issues."

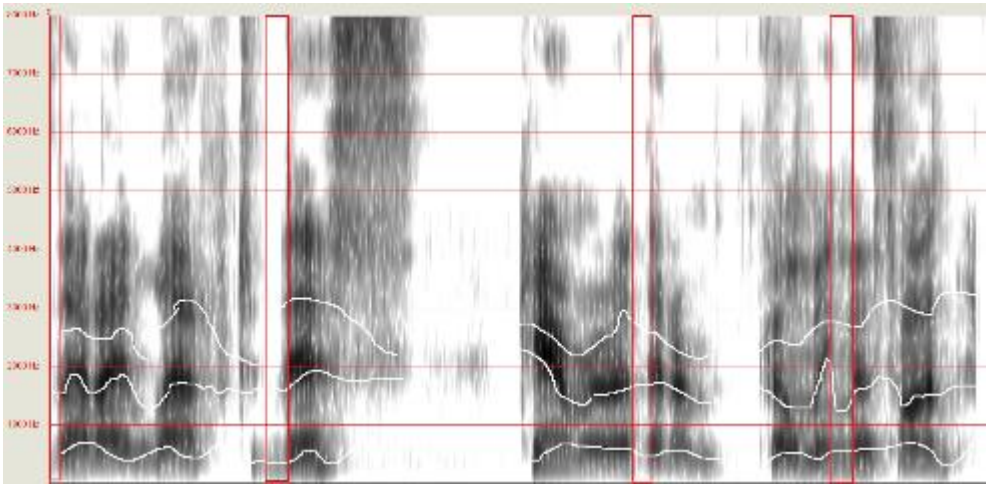


Source: TomSka & Friends (2021, 1:02-1:04).

For the first /ð/, all formants look normal. It does not have that much energy near 8,000 Hz for it to be confused with a /v/. As the sound is part of the Onset, a /v/ would have a much lower F2, thus confirming this is a /ð/. By watching the video, the speaker moves when pronouncing the sound, so it is not reliable information.

The second /ð/ looks normal, except the F2 that is lower than usual, but it is just an influence of the /m/ after it. Again, the place of articulation of the /ð/ is not visible.

Figure 17. Artsy Madwoman - "That I left this camera that I'm filming this on."



Source: Artsy Madwoman (2021, 0:21-0:24).

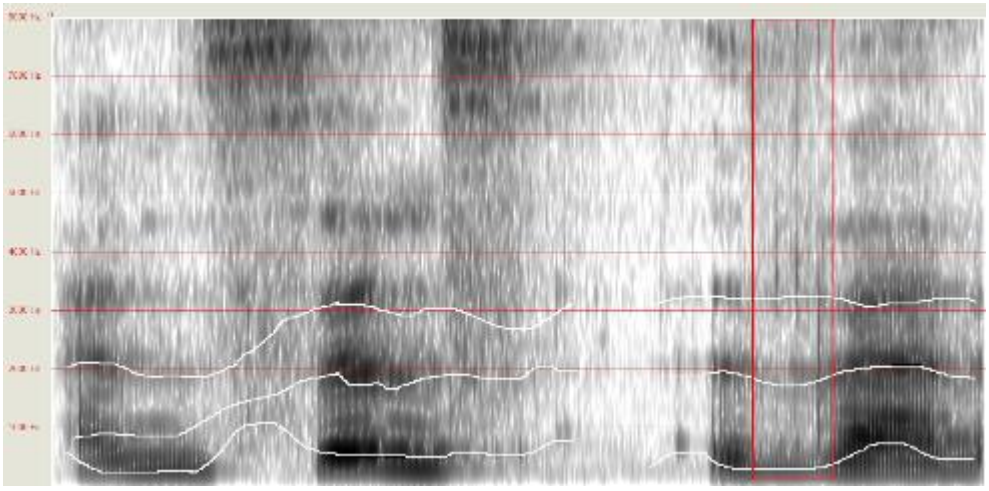
All phonemes here do not have energy from 5,000 Hz, except the last one for a little bit, which is a characteristic of some /v/ pronunciations. Also, every phoneme here is part of the Onset; if these were /v/ sounds, the F2 would be lower. It is safe to say that every phoneme here is well pronounced, with particularity in the place of articulation. The speaker clearly showed her tongue for the first /ð/; for the rest, it can be seen that she placed her tongue at the back of her teeth, which is not uncommon.

Figure 18. Elise Ecklund - "First... this one looks intriguing."

Source: Elise Ecklund (2021, 0:47-0:49).

F1 seems normal while F2 and F3 are just a bit higher than usual. Due to the lack of energy near the 8,000 Hz range, the /v/ is discarded. From the video, the speaker can be seen placing her tongue for the most part at the back of her teeth. She made an editing cut the moment she pronounced the sound.

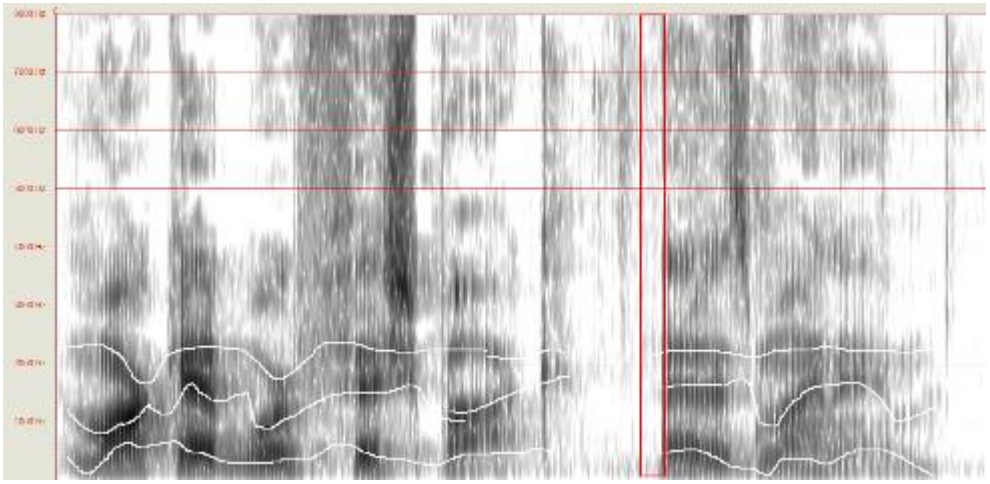
Figure 19. Emmymade - "And since then."



Source: emmymade (2021, 0:41-0:43).

This speaker, in general, had high F2's, but this /ð/ looks normal. It is a great example of pronunciation of a /ð/ as it is longer than usual and it can be clearly seen how she positioned her tongue in between her teeth.

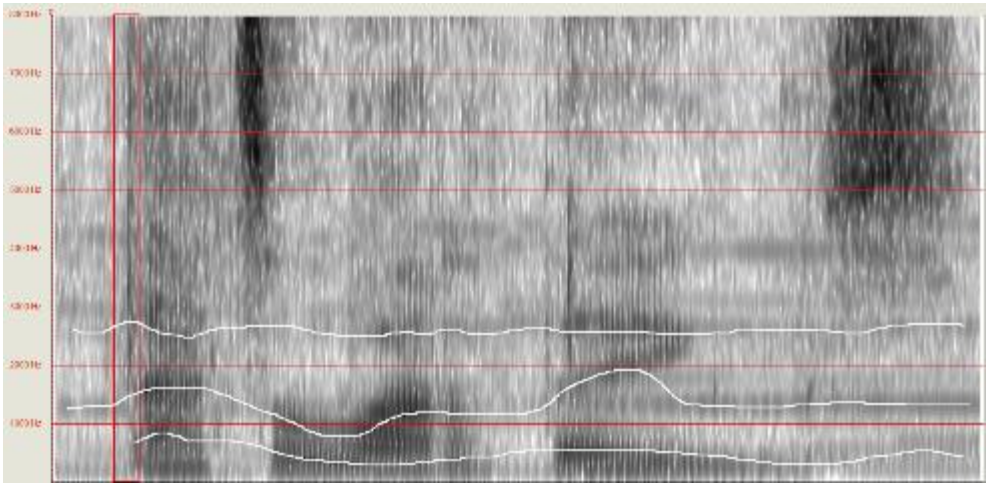
Figure 20. Markiplier - "Mark Edward Fischbach, that is my name."



Source: WIRED (2019, 0:18-0:19).

Although the sound is short, all formants seem normal for a /ð/, except F1 which is a bit higher. There is little energy near the 8,000 Hz range, thus discarding the possibility of this being a /v/. By watching the video, the place of articulation cannot be seen, so the speaker placed his tongue at the back of his teeth.

Figure 21. Rybonator - "That's a lot of things."



Source: Rybonator (2020, 0:58-0:59).

All formants are normal, except for F1 that is higher than usual, influenced by the openness of /æ/ after it. There is also energy in all frequencies, maybe because of the background music or because the sound was pronounced with more frication than usual. In the video, it is clearly seen the manner of articulation of the sound, placing the tongue in between the teeth. This is a great example of the pronunciation of /ð/.

4.4.3 Voiceless dental fricative /θ/

Figure 22. CDawgVA - "quarantine hotel for three days."

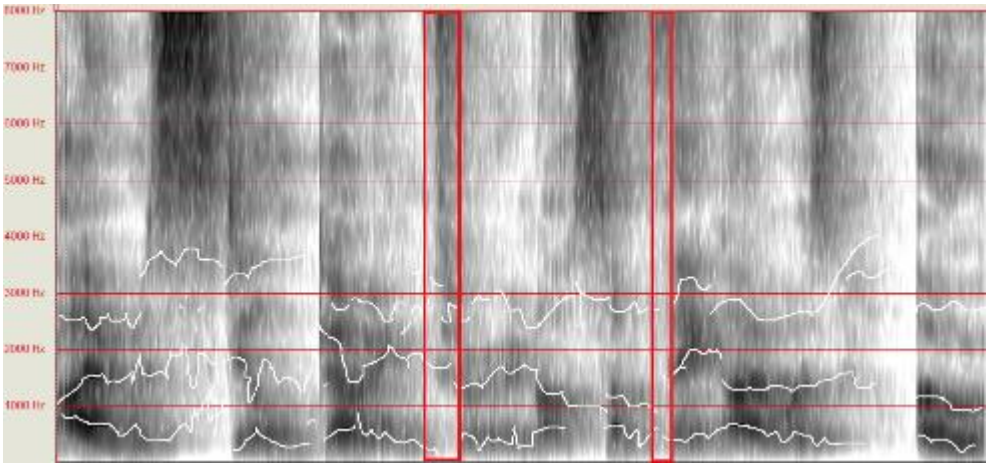
Source: CDawgVA (2021, 0:07-0:09).

Fricatives are characterized by the expulsion of big amounts of energy, which is represented in the high frequencies of the spectrogram. Usually, dental and interdental fricatives are voiceless and present more energy above 3,000 Hz.

In this case, it is possible to observe a concentration of energy near the range of 6,000 Hz to 8.000 Hz. There is also a decrease of energy below 2,500 Hz. Finally, there is an absence of a voicing bar in the F1 which indicates that the sound is voiceless. In the video, it is possible

to find some lips movement due to the vowel sounds near the analyzed fragment. Besides, it is not possible to see the tongue while the sound is produced, therefore it is located in the back of the upper teeth meaning it is a dental fricative sound.

Figure 23. Daniel Howell - "I was still going through you know something I was."

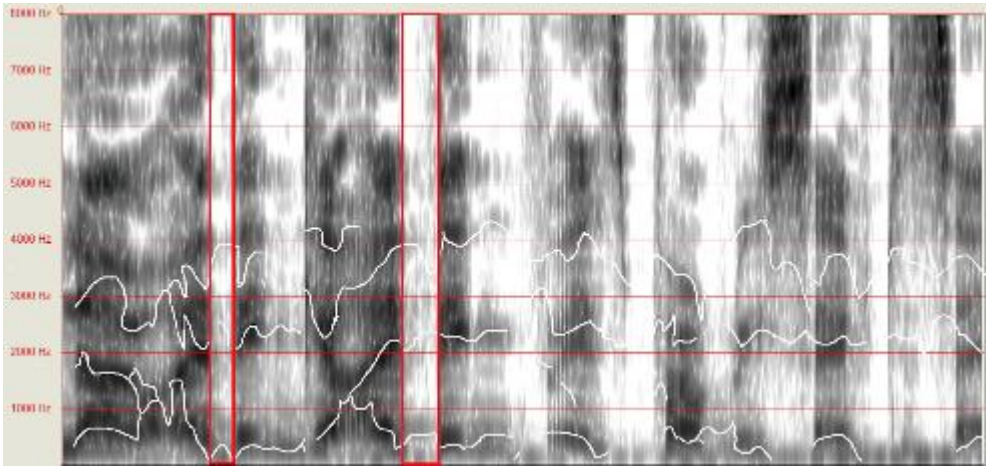


Source: Daniel Howell (2017, 0:23-0:25).

In the first case, it is possible to observe a constant amount of energy above 1,500 Hz; while the concentration of most energy is located above 7,000 Hz. The noise also decreases noticeably near the F1 where there should be the absence of a voicing bar. In the video, it is possible to find that there is no lips movement and the tongue is not observable, which means it is not an interdental but a dental fricative sound produced with the tongue against the back of the upper teeth.

For the second case, the energy can be better noticed above 4,500 Hz, while the concentration of most energy is located above 7,500 Hz. There is also a low amount of energy near the F1. In the video, it is possible to find that there is no lips movement but the tongue is placed between the teeth, which indicates it is an interdental fricative sound.

Figure 24. Emilia Clarke - "um… no, I think, I think it is probably safe."

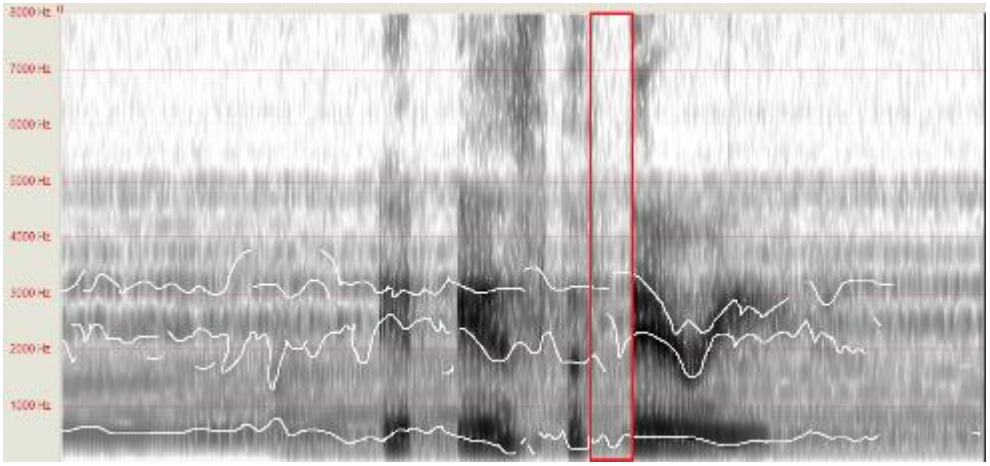


Source: Vanity Fair (2019, 1:12-0:15).

In both cases, it is possible to observe a decrease of energy above 3,000 Hz. This may signify the presence of a plosive sound (since plosives present such characteristics). Besides, the zone of most energy is located below 1,000 Hz, which means that it does not comply with two of the main characteristics of the sound: having more energy in high frequencies above 3,000 Hz., and the lack of a voicing bar. Due to these reasons, it would be possible to conclude that the spectrogram is not representative of the interdental fricative sound /θ/.

In the video, it is possible to find some lips movement due to the vowels near the sounds, and it is not possible to see the tongue between the lips which means that the sound was produced by putting the tongue against the back of the upper teeth, therefore, it is a dental fricative sound.

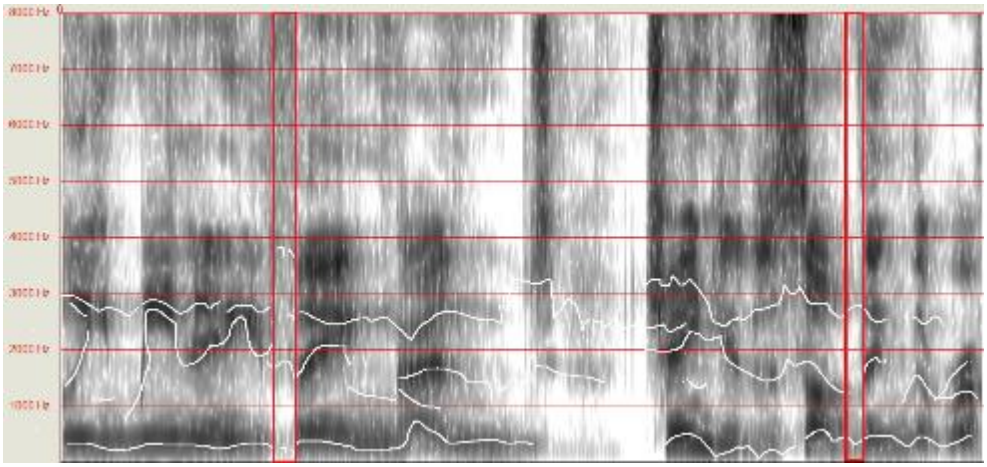
Figure 25. Emma Watson - "it is the theory."



Source: English Speeches (2019, 0:51-0:53).

In this case, a low and constant amount of energy can be observed from 0 to approximately 5,200 Hz. Also, the energy starts decreasing above 3,000 Hz and it is almost undetectable from 5,200 Hz until 8,000 Hz. In the video, it is possible to find that there is no lips movement but the tongue is between the teeth while producing the sound, thus it is an interdental fricative sound.

Figure 26. Tomska - "mean anything and change it to something..."

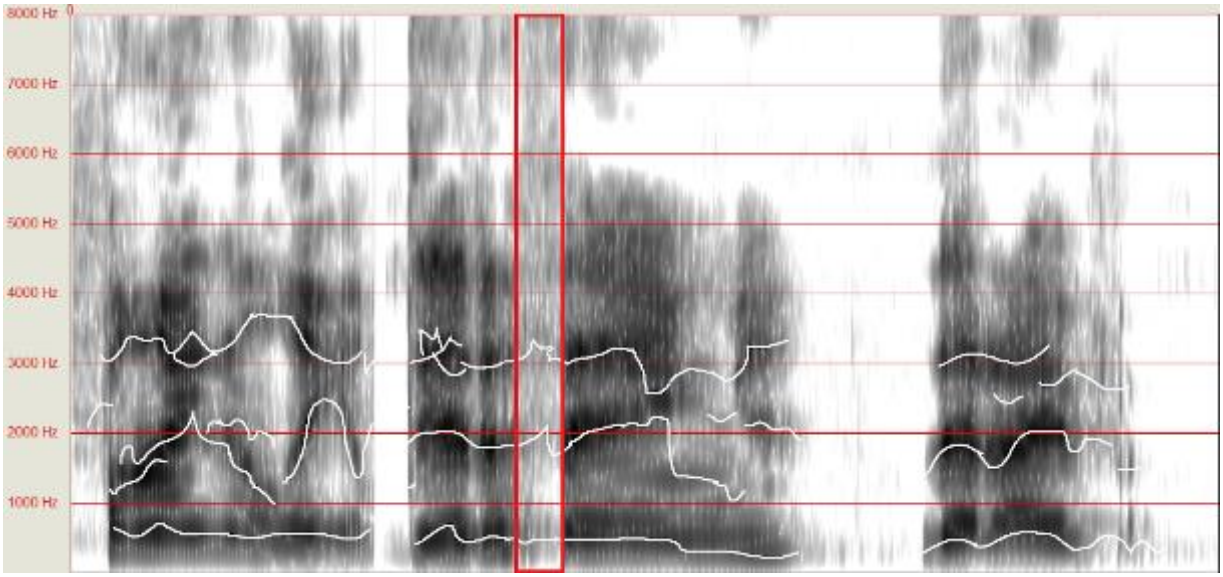


Source: TomSka & Friends (2021, 0:12-0:15).

In the first case, it is observable to have a low amount of energy which decreases slowly from 8,000 Hz where it is more concentrated, until 1,300 Hz where it becomes almost undetectable. Some energy is also possible to find near the F1 which might be due to background noise. In the video, it is not possible to see lips movement but the tongue is placed between the teeth, therefore it is an interdental fricative sound.

For the second case, just three low amounts of concentrated energy are observable from 0 Hz to 1,500 Hz, from 2,500 Hz to 3,500 Hz, and above 7,000 Hz. In the video, it is not possible to see lips movement and the tongue is also placed between the teeth meaning it is an interdental fricative sound.

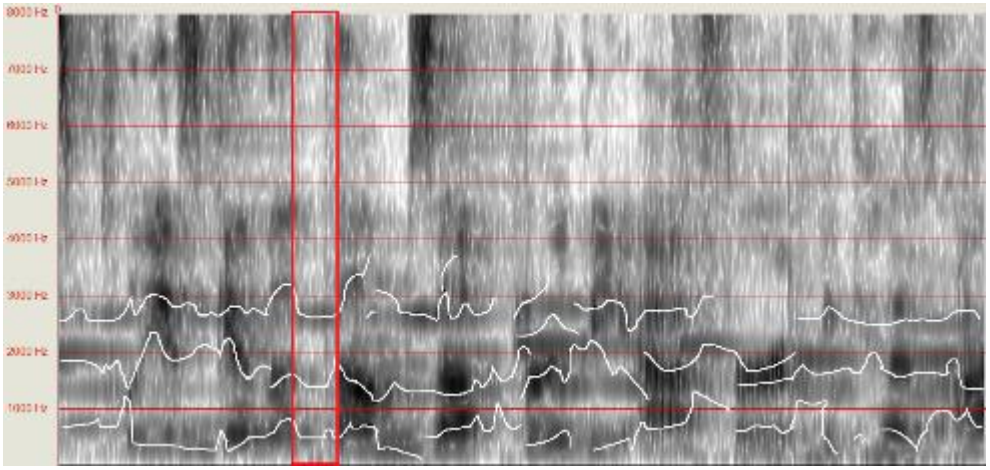
Figure 27. Artsy Madwoman - "I finally did a thing that I've."



Source: Artsy Madwoman (2021, 0:08-0:12).

It is possible to find a low amount of energy spread constantly around the whole fragment analyzed, with a decrease of almost the totality of energy between 900 Hz to 1,600 Hz approximately. The sound might be affected by the vowels surrounding it. In the video, it is possible to observe that there is no lips movement but the sound is articulated with the tongue between the teeth, which confirms it is an interdental fricative sound.

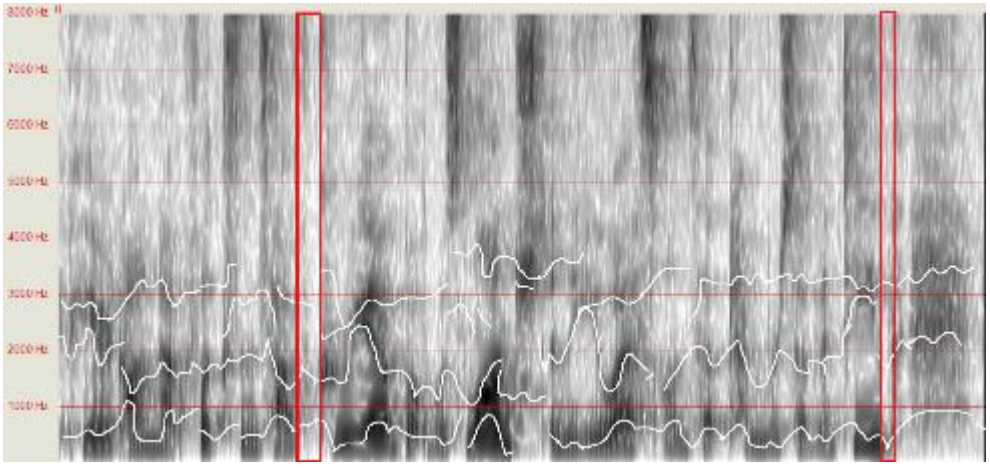
Figure 28. Elise Ecklund - "please give a thumbs up for the effort."



Source: Elise Ecklund (2021, 0:14-0:16).

In this case, it is possible to observe energy around almost the whole occurrence of the sound. The darkest area where most energy is concentrated is in the F3, but it is also possible to find another concentration of energy in the F1. The spectrogram might be affected by the background noise. It is possible to find in the video that there is no lips movement while the sound is produced.

Figure 29. Emmymade - "appeared on Netflix in south korea but was launched in the US on september seventeenth."

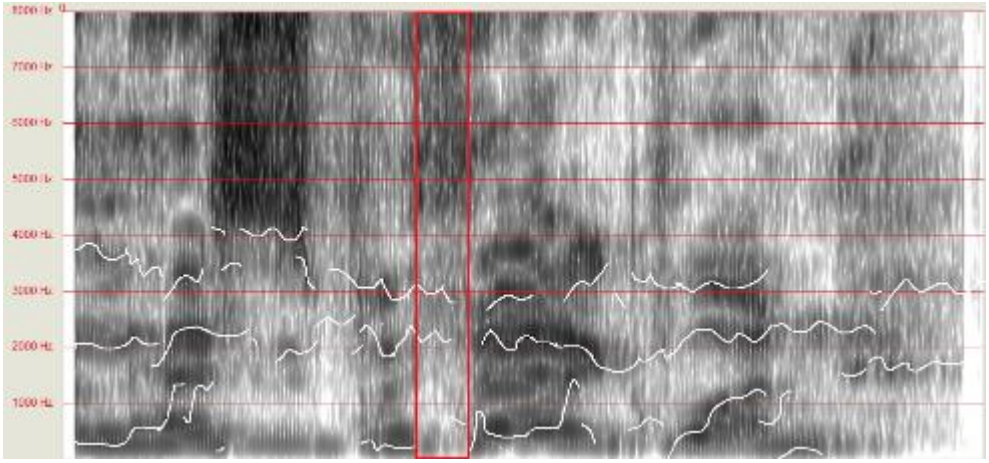


Source: Emmymade (2021, 0:36-0:41).

In the first case, it is possible to find a very low amount of energy along the whole occurrence. The concentration of energy is located in the F1 to F2 range. The energy slightly decreases above 2,500 Hz until 6,200 Hz where it starts increasing again. In the video, it is possible to appreciate some lips movement due to the previous vowel sound.

For the second case, it is possible to observe two areas of concentrated energy, located in the F1 and in the F3, being the first of them the one with the most energy. The amount of energy in the whole fragment analyzed is very slow and is almost constant. In the video, it is possible to find that the lips are wide open due to the previous vowel sound and the tongue is between the teeth while the sound is produced which indicates it is an interdental fricative sound.

Figure 30. Markiplier - "has this pathetic one."

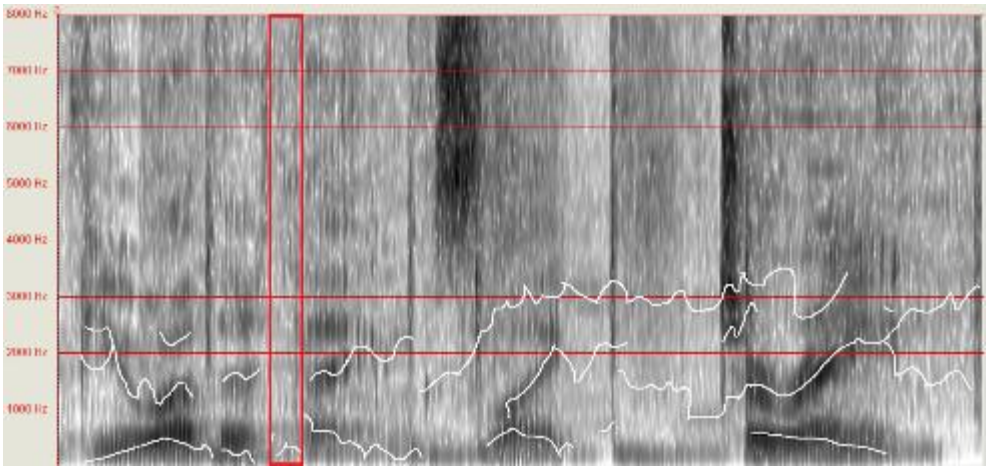


Source: Markiplier (2021, 9:26-0:27).

In this fragment, it is possible to appreciate a concentration of energy between 4,000 Hz to 6,100 Hz approximately. Below 4,000 Hz, the noise starts decreasing slowly until a section around 1,400 Hz to 800 Hz in which the amount of energy is very low. Besides, there is an increase of energy around the F1 which might be considered a voicing bar. This sound might be influenced by the vowel sounds around it. In the video, it is possible to find some lips movement

due to the previous plosive sound, and the tongue is between the teeth while the sound is produced which indicates it is an interdental fricative sound.

Figure 31. Rybonator - "few other things like terrain."

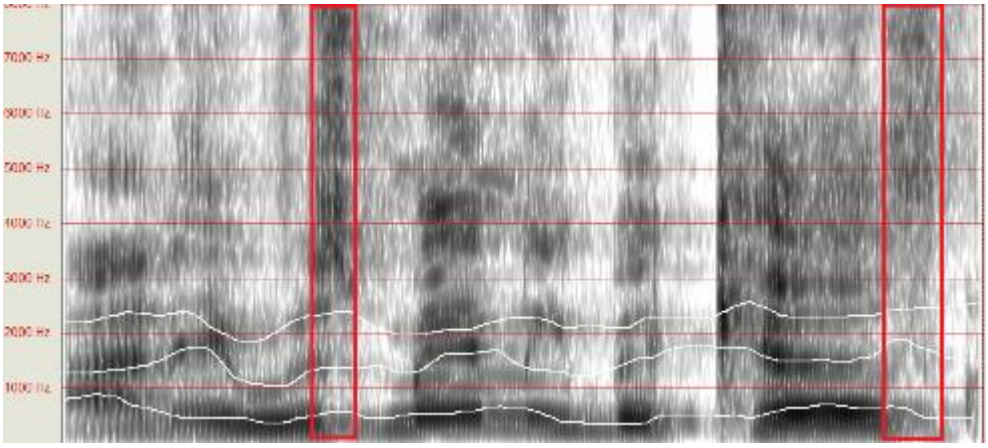


Source: Rybonator (2020, 0:28-0:31)

In this occurrence of the sound, the energy is spread almost evenly through the whole fragment. It is possible to find a slight increase of energy around F1 as well as around 5,000 Hz. Notwithstanding, the analysis of this fragment has been hindered by the presence of background noise. In the video, it is possible to observe that there is no lips movement and the tongue is between the teeth, which confirms it is an interdental fricative sound.

4.4.4 Voiced alveolar fricative /z/

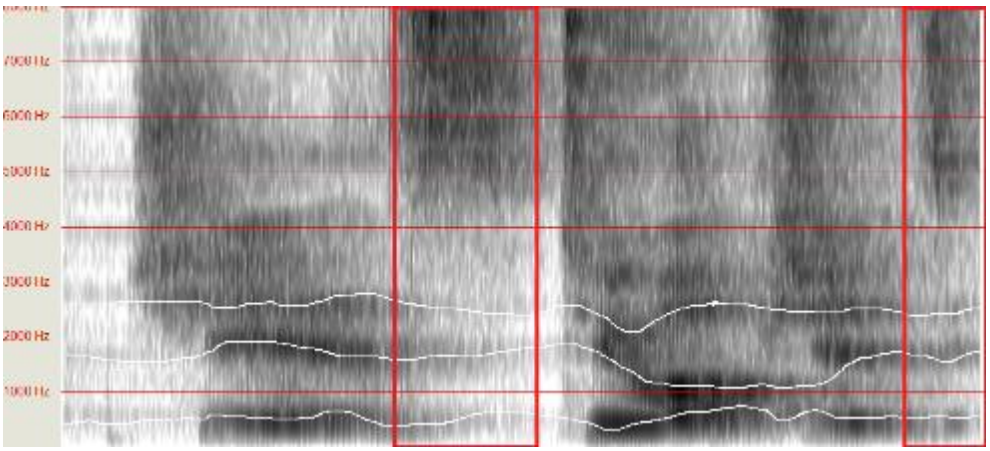
Figure 32. CDawgVA - "and it was mainly because."



Source: CdawgVA (2020, 0:49-0:50).

In both occasions (represented by the highlighted sections of the spectrogram) there is noise in the higher frequencies while disappearing said energy between the frequencies of 1,000 Hz and 3,000 Hz, making them both alveolar fricatives. Both sounds are voiced /z/ as there is a constant voicing bar at the bottom.

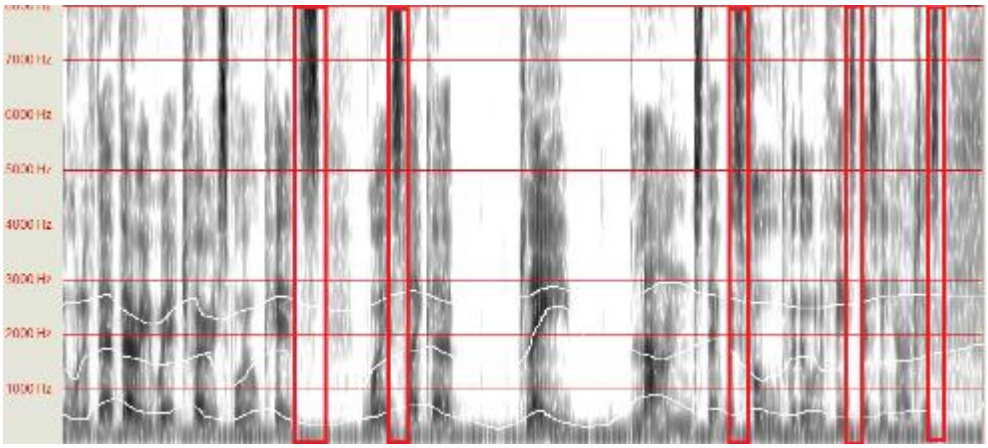
Figure 33. Daniel Howell - "cheers, two options, either."



Source: Daniel Howell (2017, 1:52-1:54).

On both occasions there is noise in the higher frequencies while disappearing said energy between the frequencies of 1,000 Hz and 3,000 Hz, making them both alveolar fricatives. However, only the second one is a traditional /z/ with a constant voicing bar at the bottom, as the first one is an allophone of /z/ as its voicing bar vanishes midway through its pronunciation.

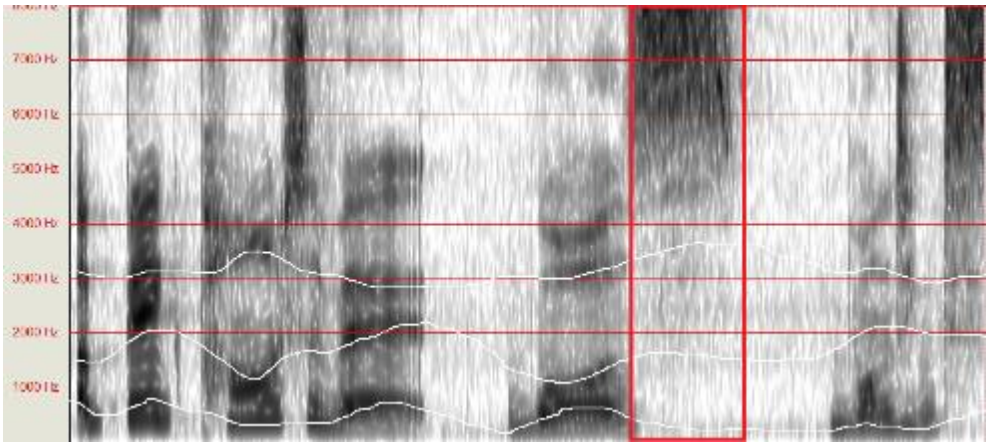
Figure 34. Emilia Clarke - "all of their very rich, very spoiled kids ... was like a playland ... that was full of all these Disney characters."



Source: Vanity Fair (2019, 7:49-7:57).

The five occasions present /z/ as there is an accumulation of energy in frequencies around and higher than 4,000 Hz accompanied by a voicing bar in the frequencies between 0 and 1,000 Hz. However, the last two occasions in which /z/ is present are not easily perceived by a listener.

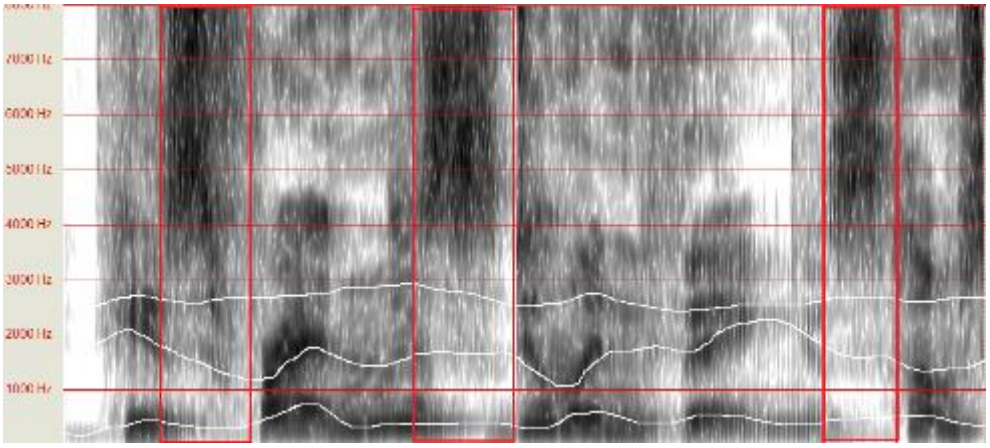
Figure 35. Emma Watson - "a big part of it was wanting to."



Source: Emma W. Thailand (2017, 15:06-15:09).

There is an alveolar fricative sound pronounced as there is a higher release of energy in higher frequencies that disappears in 4,000 Hz and below. However, it cannot be considered a /z/ as it does not possess a voicing bar, making the sound more akin to /s/.

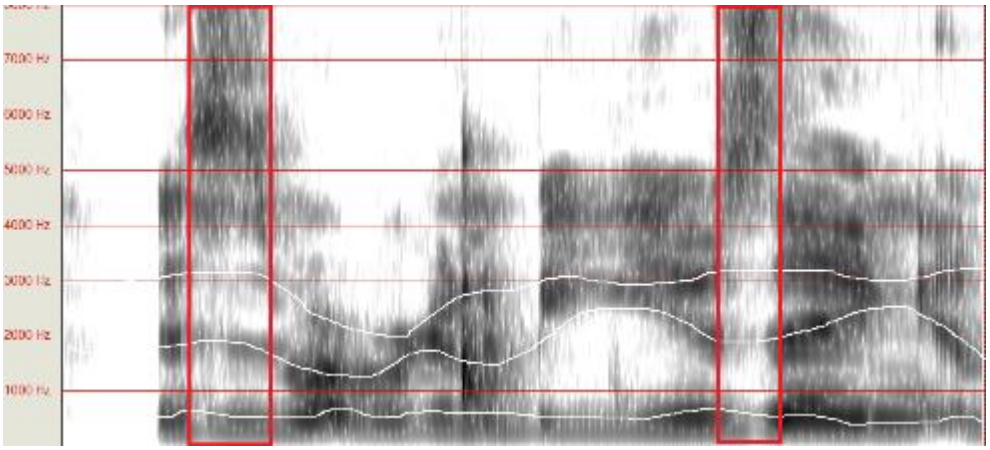
Figure 36. Tomska - "and his friends doing things."



Source: TomSka & Friends (2021, 0:21-0:23).

The two first occasions are situations in which /z/ is present as there is an accumulation of energy in frequencies around and higher than 4,000 Hz accompanied by a voicing bar in the frequencies between 0 and 1,000 Hz. However, the first and third ones are allophones of /z/ as the voicing bar is not constant through the spectrogram.

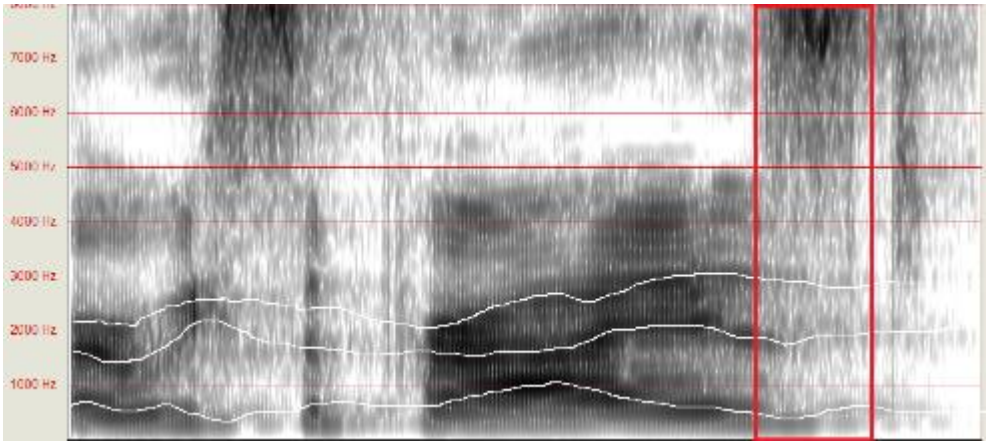
Figure 37. Artsy Madwoman - "this is what we are going to be using."



Source: Artsy Madwoman (2021, 3:40-3:41).

The two occasions present /z/ as there is an accumulation of energy in frequencies around and higher than 4,000 Hz accompanied by a voicing bar in the frequencies between 0 and 1,000 Hz. However, it is worth noting that in the second highlighted section, the energy disappears from the voicing bar in the middle.

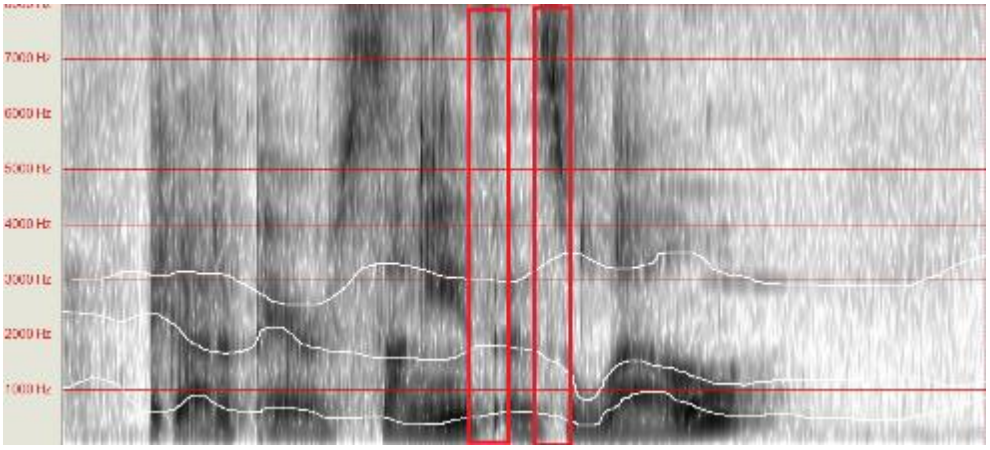
Figure 38. Elise Ecklund - "are we surprised?."



Source: Elise Ecklund (2021, 0:06-0:07).

It is an allophone of /z/ as while there is a concentration of energy in the higher frequencies, its voicing bar vanishes midway through its pronunciation.

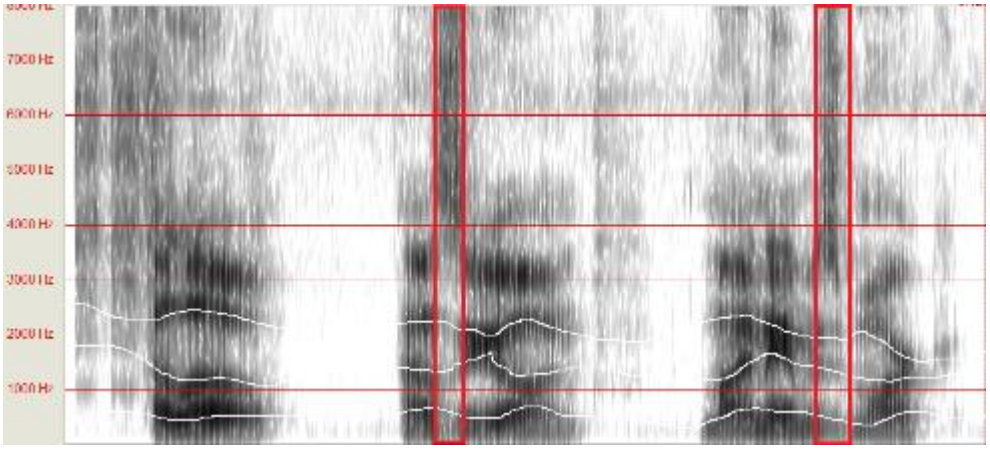
Figure 39. Emmymade - "I got a pair of sunnies as well."



Source: emmymade (2021, 0:56-0:58).

The two occasions present /z/ as there is an accumulation of energy in frequencies around and higher than 4,000Hz accompanied by a voicing bar in the frequencies between 0 and 1,000Hz. However, it is worth noting that in the first highlighted section, the accumulation of energy is not that much higher than in the rest of her speech while in the second one, the energy disappears from the voicing bar in the middle.

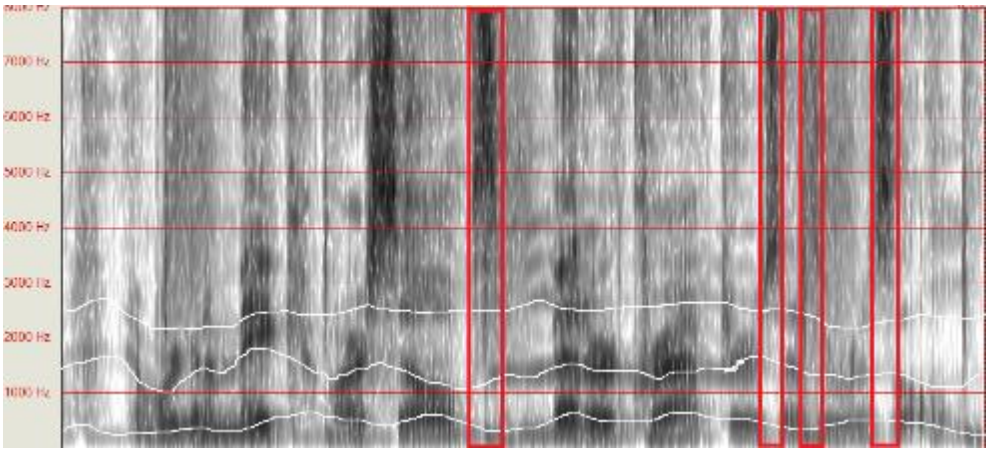
Figure 40. Markiplier - "Hello, how is it going? My name is Mark."



Source: Markiplier (2021, 0:00-0:03).

The two occasions present /z/ as there is an accumulation of energy in frequencies around and higher than 4,000 Hz accompanied by a voicing bar in the frequencies between 0 and 1,000 Hz.

Figure 41. Rybonator - "and we are going to make ourselves better crafters as a result."

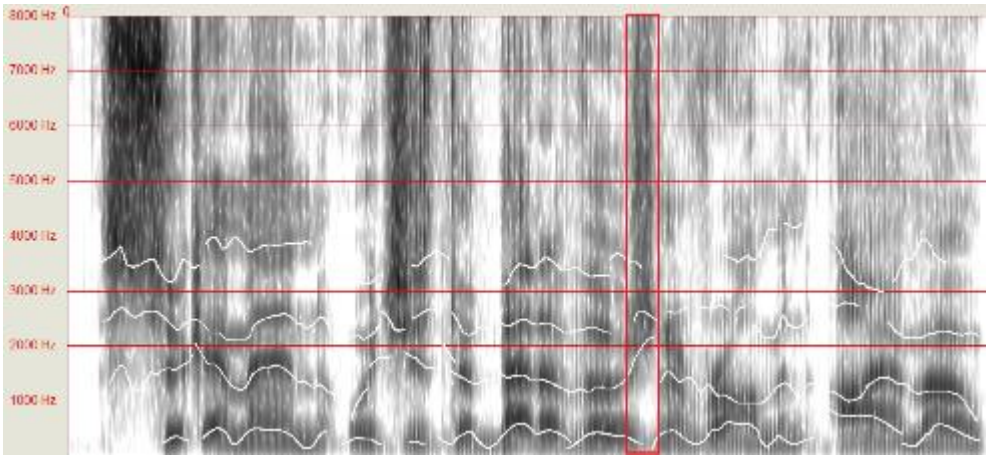


Source: Rybonator (2020, 1:13-1:15).

In the highlighted sections we appreciate fricative alveolar sounds as there is a concentration of energy frequencies higher than 3,000Hz. However only 3 present a constant voicing bar making them /z/; the other one does not present a constant voicing bar, which makes it a devoiced allophone of /z/

4.4.5 Voiced postalveolar fricative /ʒ/

Figure 42. CDawgVA - "sukemen which is the better version of ramen"

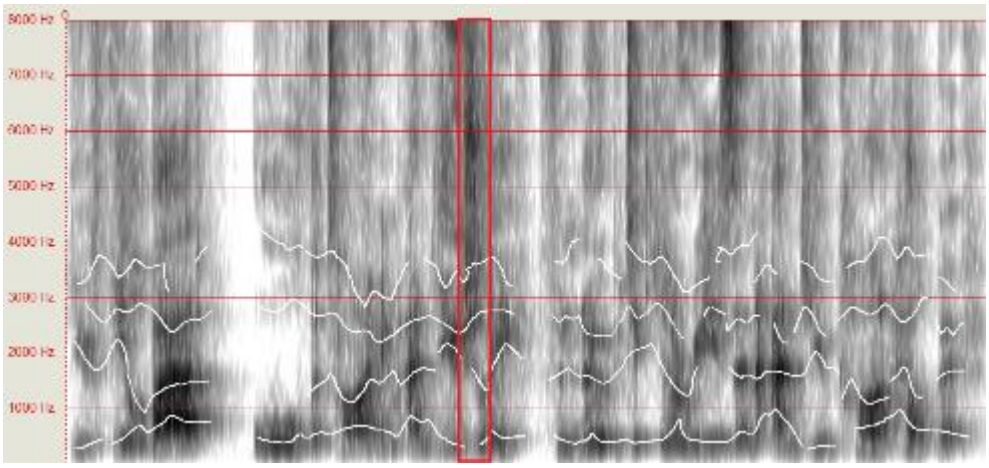


Source: CDawgVA (2021, 20:38-6:41).

The postalveolar fricative /ʒ/ is produced behind the alveolar ridge (Ogden, 2017). For this sound to occur there needs to be some degree of occlusion of the vocal tract (enough to produce frication), and it is characterized as having a lesser amplitude and a shorter duration of frication, as well as a voicing bar, which sets it apart from its voiceless counterpart /ʃ/.

In this sample you can see how the darkest part of the noise is above 2.000 Hz, as well as the presence of a voicing bar, which indicates vibration, making this a clear example of the voiced fricative /ʒ/.

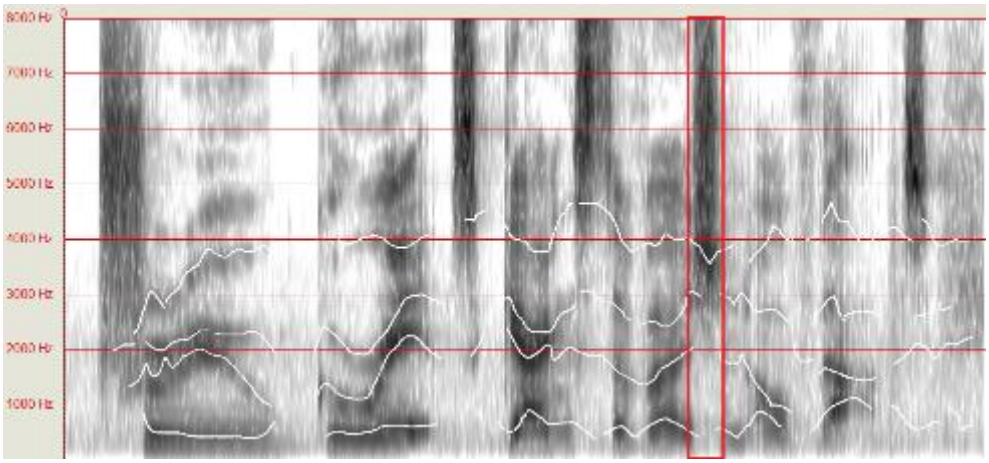
Figure 43. Daniel Howell - "you feel bad all the time, and usually, for no particular reason."



Source: Daniel Howell (2017, 1:31-1:37).

In this sample, you can see how the energy is concentrated above the F2 range, and the frication is very short. You can also distinguish a voicing bar, which indicates vibration of the vocal folds, marking this sound as a voiced fricative.

Figure 44. Emilia Clarke - "a true grit gangster version of that."

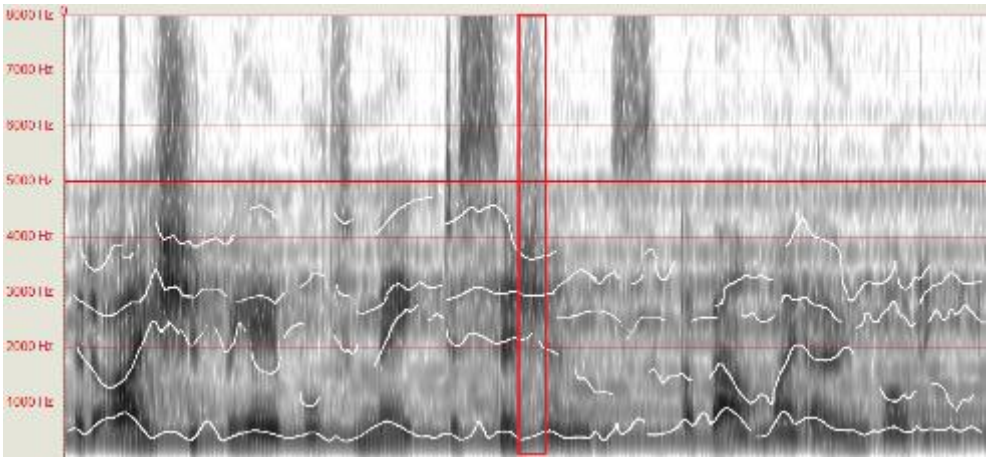


Source: The Upcoming (2018, 6:08-6:11).

In this sample, you can see how the energy is concentrated above the F3 and F4 range, or above 3.000 Hz, and the frication is short. While this mostly corresponds to the characteristics

of the voiced fricative /ʒ/, there is no voicing bar below the concentrated energy, which indicates this sound is devoiced, making it more similar to a /ʃ/ sound.

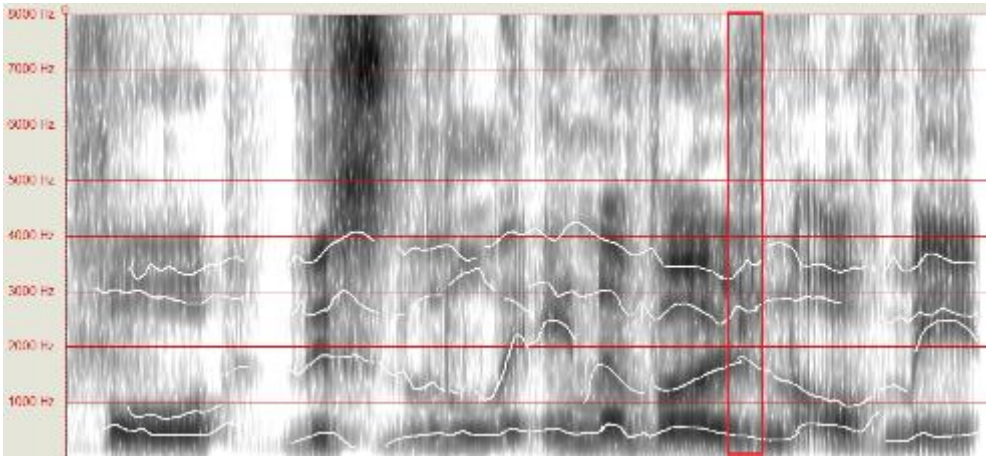
Figure 45. Emma Watson - "that I should be able to make decisions about my own body."



Source: English Speeches (2017, 2:47-2:50).

Although the frication here is short, the energy is mostly concentrated between the F2-F4 range instead of going higher. The voicing bar is weak, which means the sound is partially devoiced, making the /ʒ/ sound more like a /ʃ/. It is worth noting that there is noise in this sample, and that could be affecting the spectrogram readings.

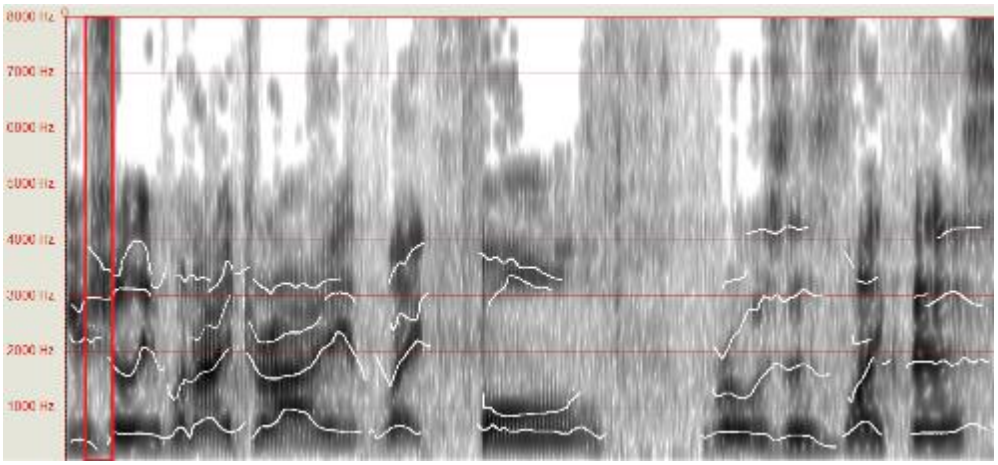
Figure 46. Tomska - "for this little inner version of me."



Source: TomSka & Friends (2021, 6:31-6:34).

In this sample, the /ʒ/ sound is centered between F2 and F4, while the visible frication is short, and the voicing bar is strong, making the vibration clear. This corresponds with the characteristics of a voiced fricative consonant.

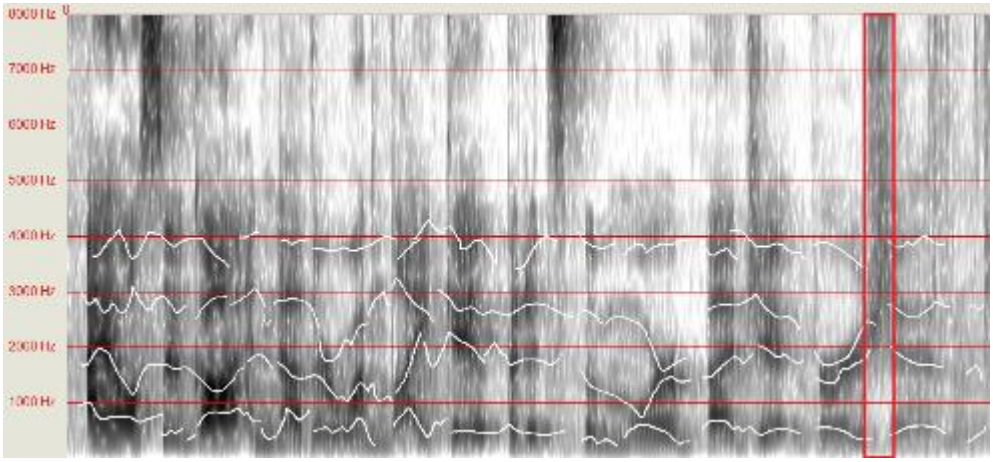
Figure 47. Artsy Madwoman - "usually when I dry full roses like this."



Source: Artsy Madwoman (2020, 18:40-18:43).

In this sample the energy is centered well above F4, going past 8.000 Hz. Although the frication is not short, the voicing bar becomes weak partway through, which makes it sound less like a /ʒ/ and more like its voiceless counterpart /ʃ/, indicating that this sound is partially devoiced.

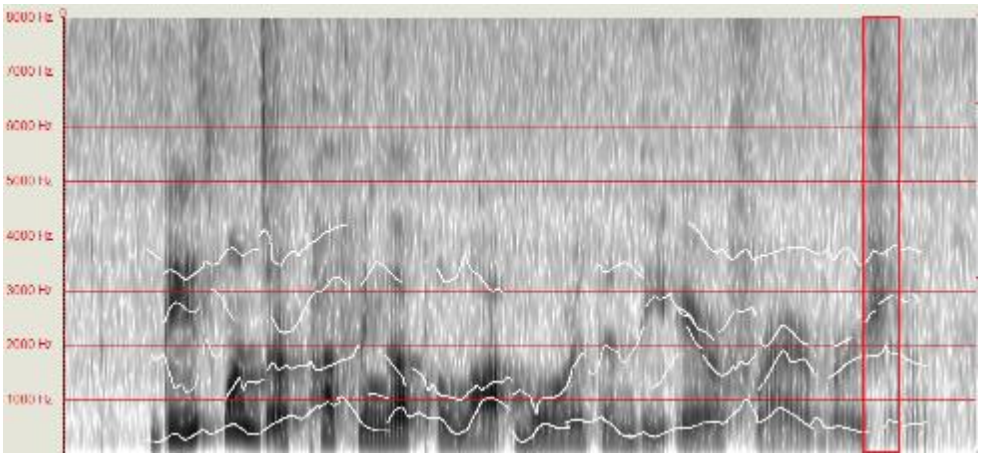
Figure 48. Elise Ecklund - "it's the more budget version."



Source: Elise Ecklund (2021, 2:41-2:43).

In this sample, the energy is centered above F2, and the frication is short. While the voicing bar appears somewhat weak, there is still vibration, which makes this a /ʒ/.

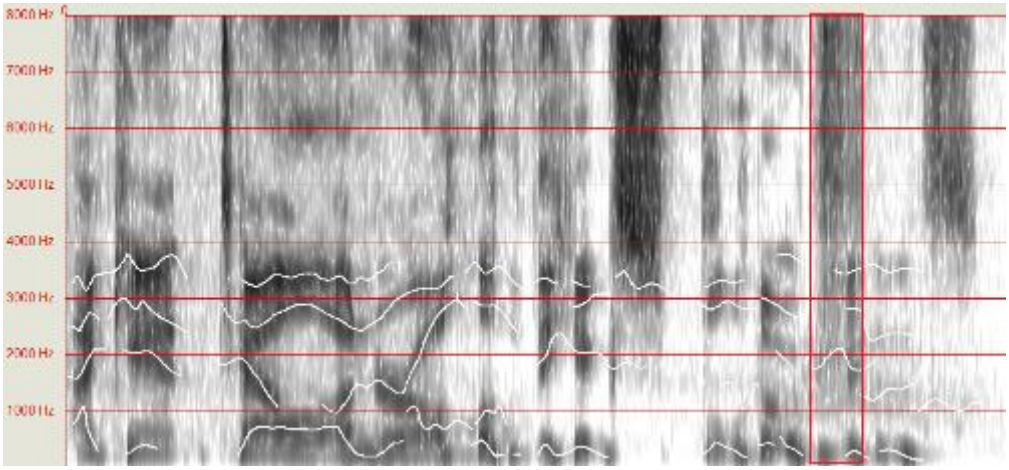
Figure 49. Emmymade - "these ones are puffing up a lot more than the air-fryer version."



Source: emmymade (2021, 9:32-6:34).

In this sample, you can see the energy is centered between F2 and F4, and the duration of the frication is short. If you look below the concentrated energy, you can see that there is no voicing bar. This means the sound is devoiced, making it more similar to a voiceless fricative rather than a voiced one.

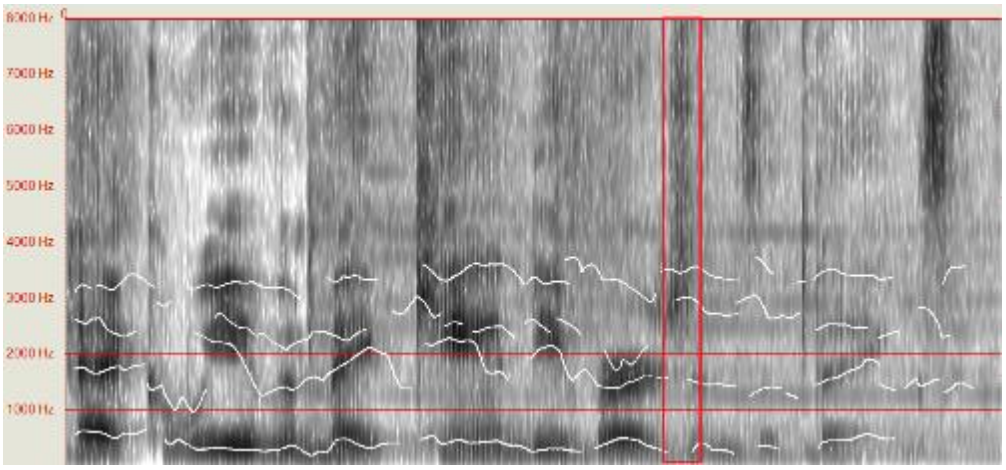
Figure 50. Markiplier - "I can't tell under all the explosions."



Source: Markiplier (2021, 20:44-20:47).

In this sample, you can see a voicing bar below the concentration of energy past the F4 range, around 5.000 Hz. The friction is longer than usual, but the voicing bar indicates vibration, making this a slightly longer voiced fricative consonant.

Figure 51. Rybonator - "heck, we even make candy versions of things."



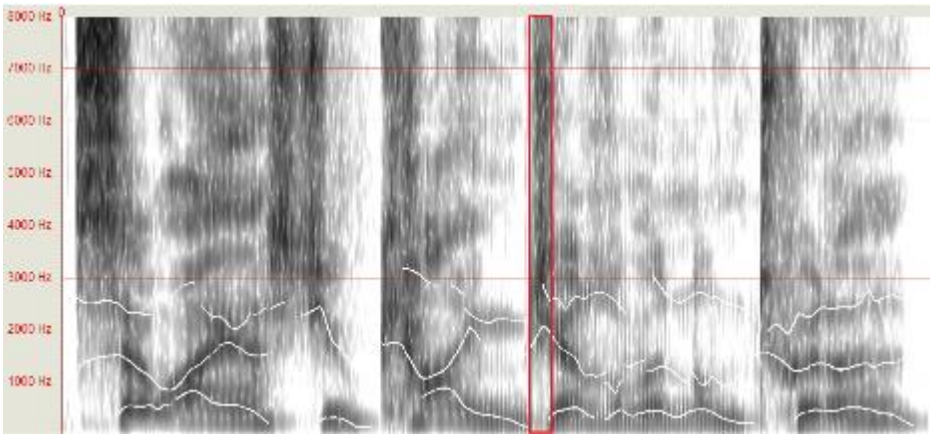
Source: Rybonator (2020, 0:56-0:58).

In this sample, the energy is centered between F3 and F4, and you can see that the frication is short. The voicing bar is weak, but there is still some vibration there. You can

compare the marked section with the /s/ sound you can see next to it, where the energy is focused above 5.000 Hz, and there is no voicing bar at all.

4.4.6 Voiced postalveolar affricate /ʤ/

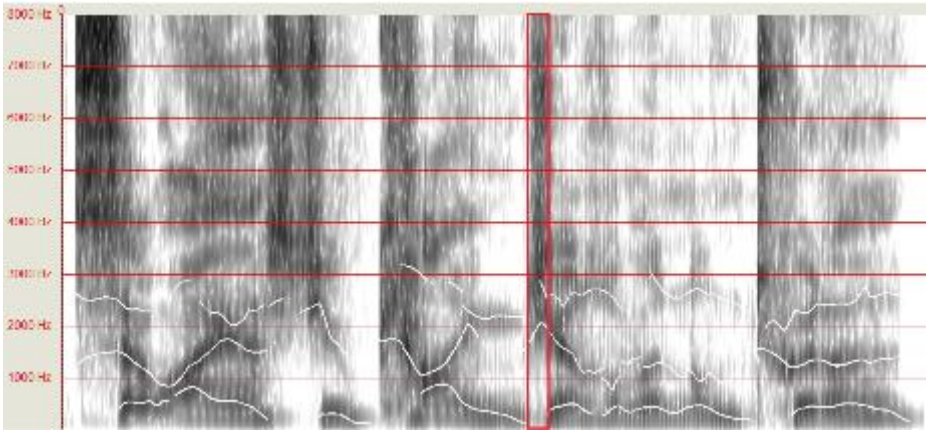
Figure 52. CDawgVA - "so I asked you, kind gentlemen on Twitter."



Source: CDawgVA (2021, 0:10-0:12).

Affricates are defined as plosives that release in fricatives (Ogden, 2017). As such, similarly to plosives, they present a gap due to occlusion of the vocal tract at first, but tend to be released as fricatives, meaning that they are usually expressed as noise in higher frequencies. Another key component is that a voicing bar appears not at the beginning of the sound, but halfway through its production. In this example, the occurrence of the sound fulfills all the characteristics aforementioned with the exception of lacking a voicing bar at the start of the sound, as in this case there is a weak voicing bar through the entirety of the fragment containing the sound.

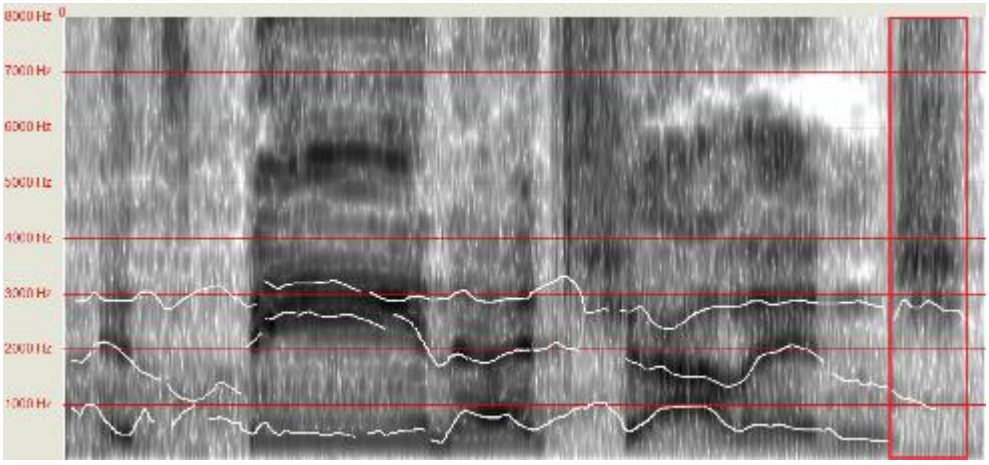
Figure 53. Daniel Howell - "it could be genetic."



Source: Daniel Howell (2017, 4:23-4:24).

In this sample, the sound extracted does not present the characteristics of a plosive, because it lacks a stop of the noise at the beginning and has an active voicing bar. It resembles more a /ʒ/ in the sense that it possesses the characteristics of a fricative sound including the constant voicing bar.

Figure 54. Emilia Clarke - "it's been a challenge."

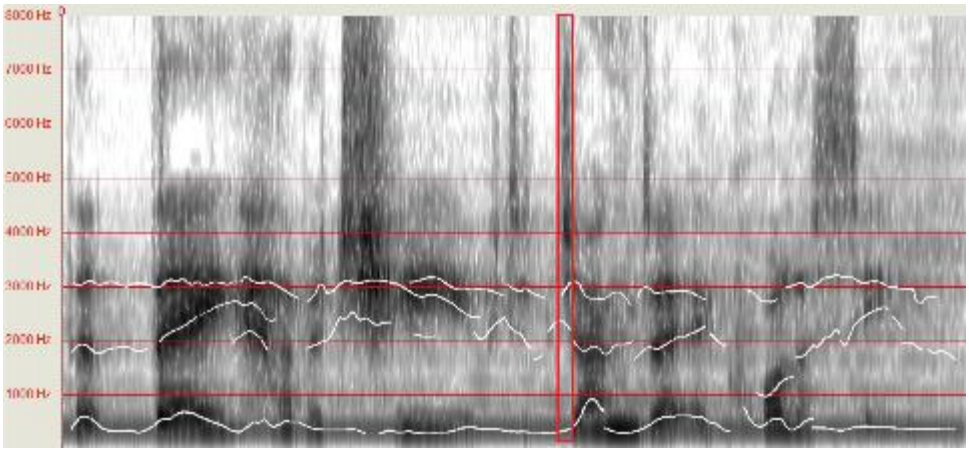


Source: Different Choice (2019, 2:34-2:37).

In this sample, there is a visible gap between the /n/ and /dʒ/ sounds, a common characteristic of a plosive, as well as the lack of a soundbar after the consonant starts (where most of the energy is concentrated in the 3,000 Hz - 5,000 Hz range), making the release similar

to that of a voiceless fricative. It is also worth pointing out that there is an amount of external noise through the entire spectrogram.

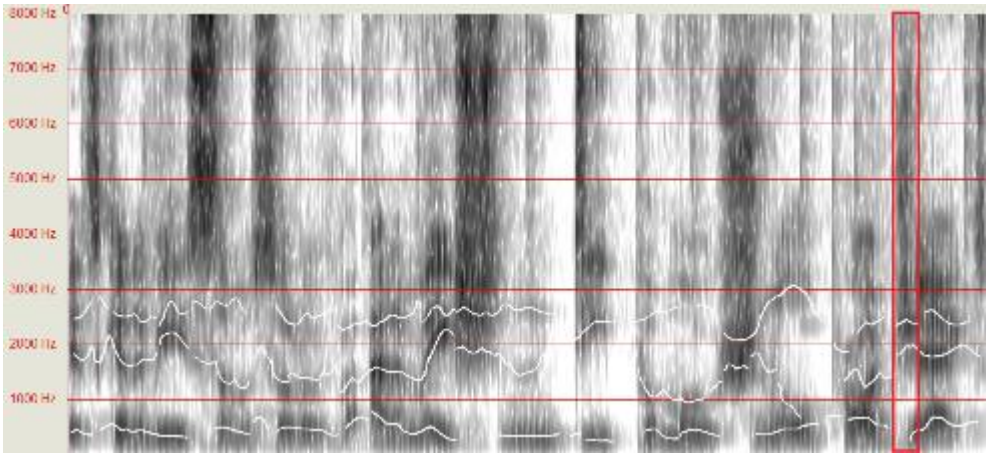
Figure 55. Emma Watson - "that they have achieved gender equality."



Source: English Speeches (2017, 3:27-3:30).

In this example, there is not a clear gap between the consonant and the vowel (which is clearly seen when analyzing plosives). You can see, however, that the energy is concentrated between the 4,000 Hz and 7,000 Hz range, which makes it similar to voiceless fricatives. However, the presence of a voicing bar through the entire sound makes it a sound more closely related to /ʒ/.

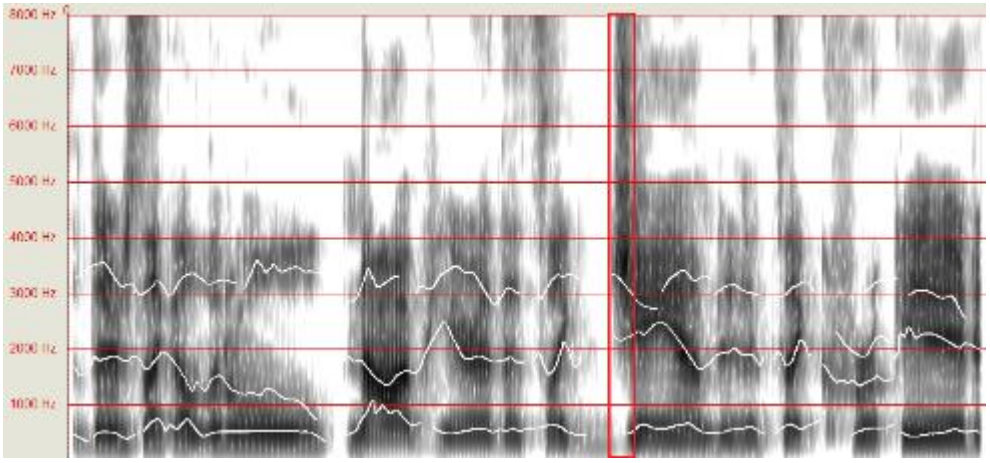
Figure 56. Tomska - "a deep emotional project."



Source: TomSka & Friends (2021, 12:45-12:48).

In this sample, you can observe a gap between the vowel and the /dʒ/ sound, as well as the concentrated energy above 2.000 Hz with no voice bar, which makes this a clear example of an affricate consonant.

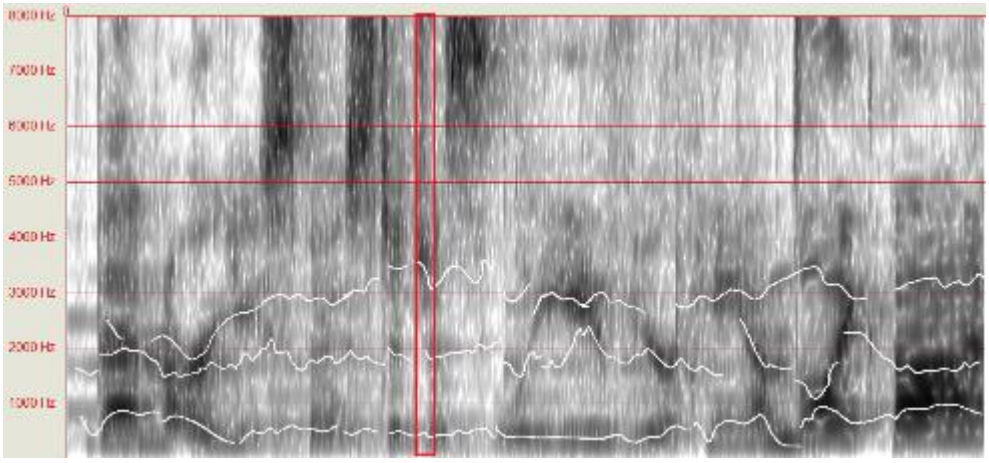
Figure 57. Artsy Madwoman - "there's like a little knob here that's like jammed."



Source: Artsy Madwoman (2021, 0:45-0:48).

Here the sound possesses all the characteristics of a voiced affricate sound, as it has a stop in all the energy before producing the sound and releases like a fricative as seen in the energy concentrated around the F2 and F3, and higher. Additionally, it also has a voicing bar that starts halfway through the segment containing the sound.

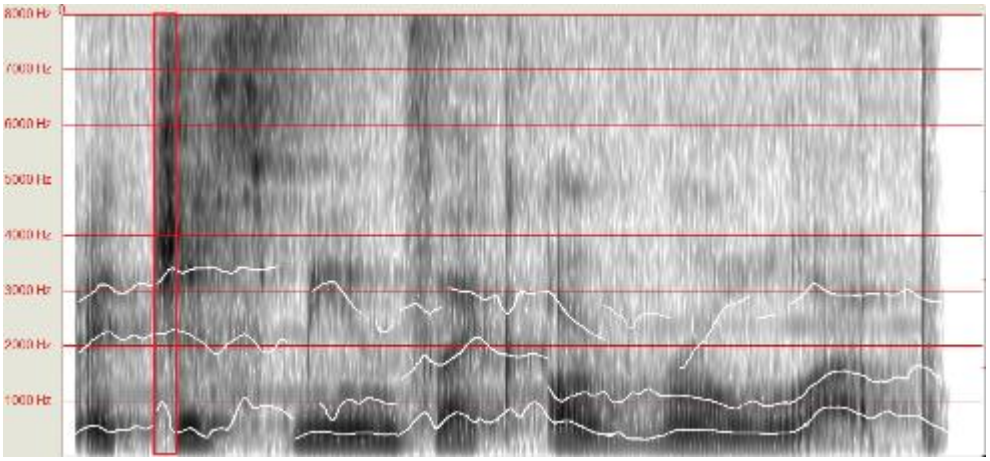
Figure 58. Elise Ecklund - "are you supposed to just leave your hand like that."



Source: Elise Ecklund (2021, 1:34-1:37).

On this occasion, the sound present does not have a stop in the energy produced at the beginning of the sound, which means that it is not a /ʤ/ sound. The sound also has energy above 3,000 Hz and also possesses a fluctuant voicing bar. It is worth considering the amount of noise present throughout the whole spectrogram.

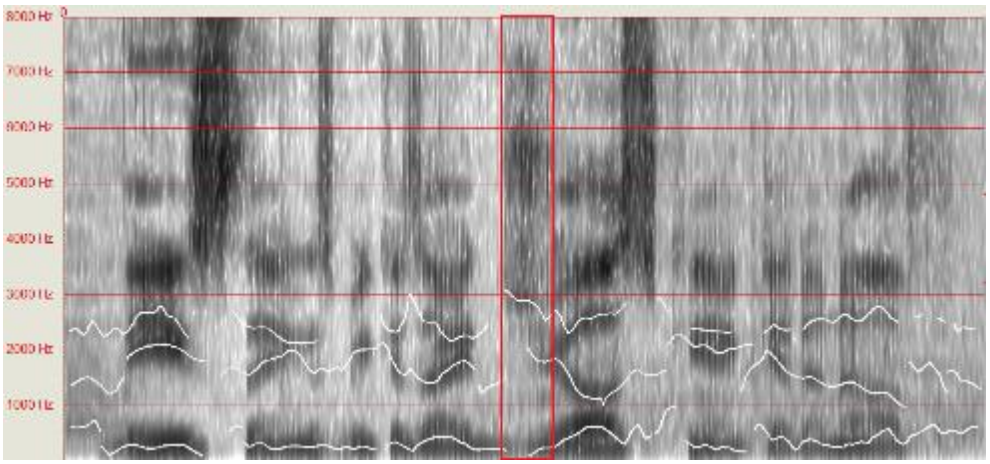
Figure 59. Emmymade - "it just means like a roll-up."



Source: emmymade (2021, 1:11-1:12).

The occurrence of the sound does not have a stop in the energy produced at the beginning of the sound, which means that it is not a /ʤ/ sound. However, the sound has energy above 3,000 Hz and does not have a voicing bar until the end of the sound giving the impression that the amount of noise present throughout the whole spectrogram may influence that there is no pause present.

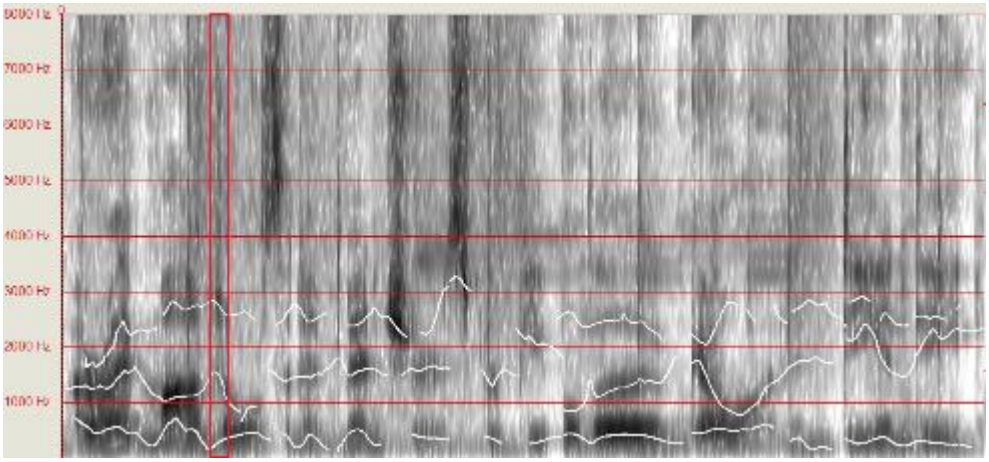
Figure 60. Markiplier - "these things because I just can't get enough."



Source: Markiplier (2021, 5:51-5:53).

The sound present fulfills all the characteristics of a /ʤ/ sound. It has a pause in the energy provided by the speaker before the sound, the energy is released as a fricative sound with the noise above 3,000 Hz and the voicing bar reappears when the sound is finishing.

Figure 61. Rybonator - "I write modules and adventures for people to go on in D&D."

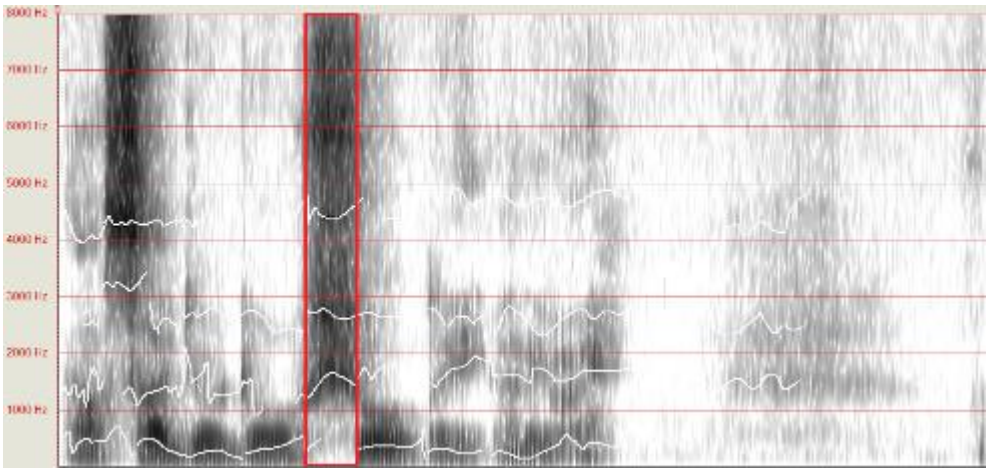


Source: Rybonator (2020, 0:39-0:42).

In this sample, there is a slight gap between the vowel and the consonant, but while there is some energy above the 2000 Hz range, it does not seem to be fully concentrated there. Furthermore, although weak, a voicing bar can be observed, making this not a proper/ʤ/ sound example.

4.4.7 Voiceless postalveolar fricative /ʃ/

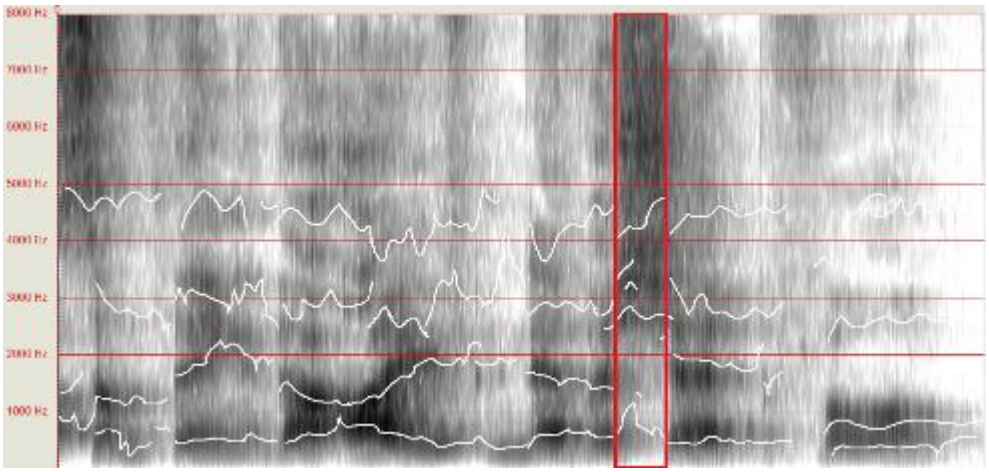
Figure 62. CDawgVA - "outside your door shortly anyway."



Source: CdawgVA (2021, 1:01-1:03).

In this case, despite more than one area with a concentration of energy can be found (namely near the ones the F2, F3, and 6,000 Hz), it is still possible to observe the concentration of energy near the F3-F4 range that usually characterizes the /ʃ/ sound. There is also the absence of a voicing bar which indicates that it is a voiceless sound. Finally, the decrease of energy below the F2 is also a recognizable characteristic of this postalveolar fricative sound.

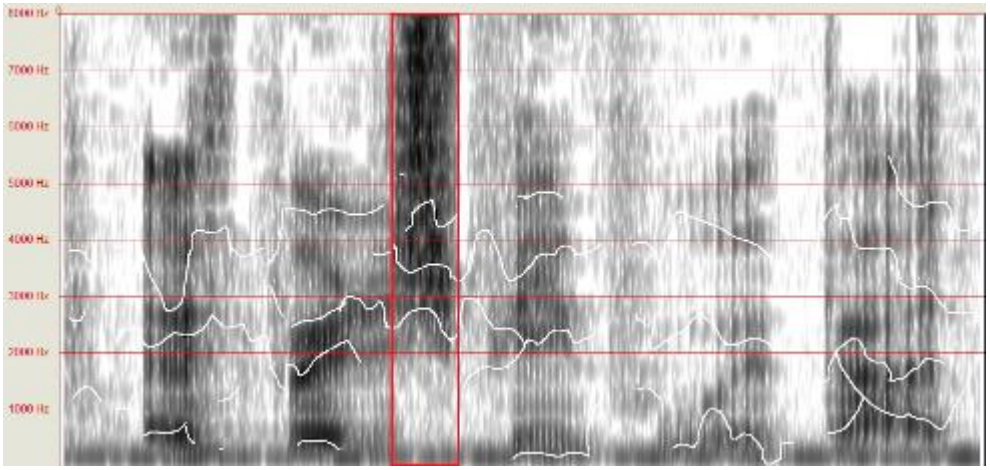
Figure 63. Daniel Howell - "something that I've never shared before."



Source: Daniel Howell (2017, 0:03-0:05).

The concentration of energy near the F3-F4 range, along with the noticeable decrease of energy below 2,000 Hz, confirms the presence of a postalveolar fricative. In the F1, it is possible to observe a low amount of energy where there should be the absence of a voicing bar, this may signify that the sound was produced partially voiced, being then an allophone of a voiceless /ʃ/ sound.

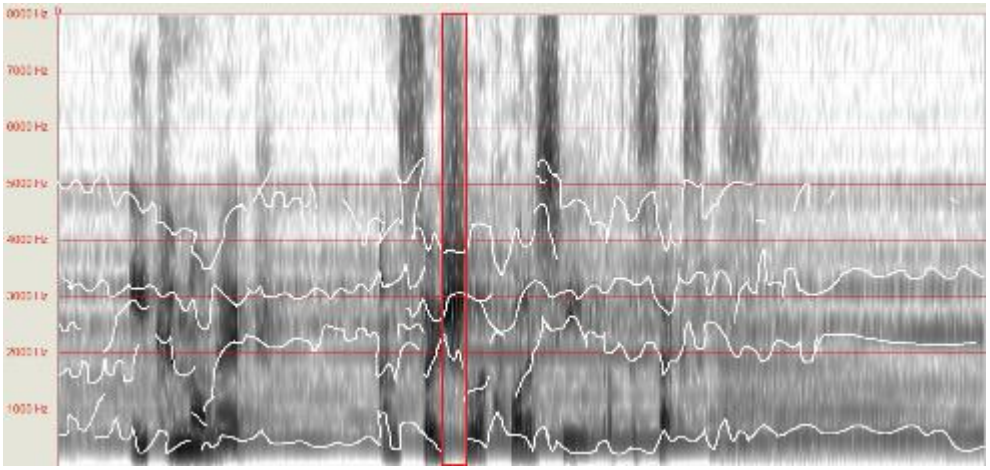
Figure 64. Emilia Clarke - "of British people with."



Source: Vanity Fair (2019, 1:58-2:00).

In this case, the presence of energy in the area where there should be the absence of a voicing bar, as well as the high concentration of energy from near the F2 to the 8,000 Hz show that the sound may have been influenced by the presence of background noise. Likewise, the great concentration of noise may occur because the sound was produced with more friction than usual. Another possible scenario is that the considerable amount of energy is due to the sound being actually an allophone of the voiceless postalveolar fricative. Notwithstanding, the energy drops off drastically below 2,000 Hz which is a common characteristic of the /ʃ/ sound.

Figure 65. Emma Watson - "economic and social equality."



Source: English Speeches (2017, 0:55-0:58).

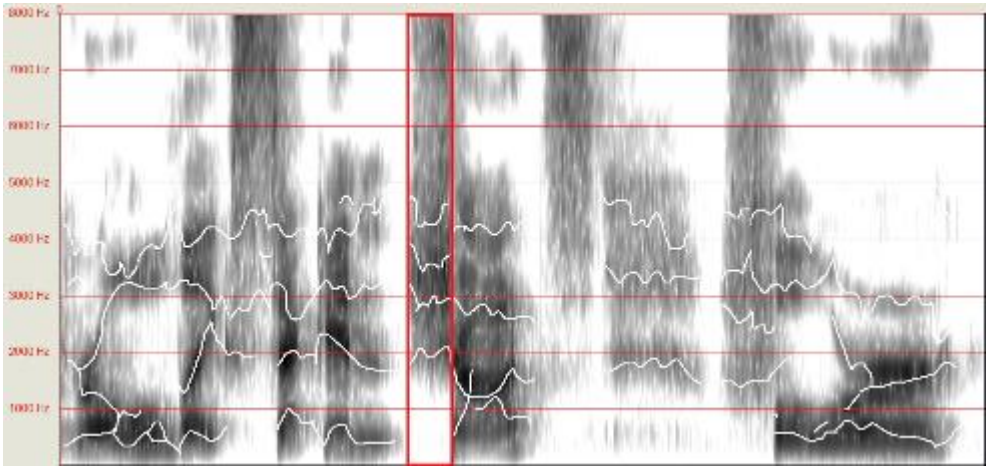
In this case, most of the energy is concentrated between the formants F3 and F4 range. Besides, the noise drops off drastically below the peak of energy. These characteristics confirm it is a postalveolar fricative /ʃ/. Although it is a voiceless sound, there is energy in the voicing bar likely due to background noise.

Figure 66. Tomska "just hang out parasocially."

Source: TomSka & Friends (2021, 0:37-0:38).

In this case, there is possible to find a low amount of energy in the F1 zone, this may occur due to the fact that the energy was maintained from the previous vowel sound, or otherwise it may signify that this sound was produced as a partially voiced allophone of the /ʃ/ sound. Besides, there exists concentration of energy in three zones: near the F3, near the F4, and in the 6,000Hz to 8,000Hz range. The majority of energy being above 6,000 Hz could be also a consequence of the sound being an allophone. Otherwise, it might happen due to excessive frication of the sound. It is also possible to observe a noticeable decrease of energy below 1,500 Hz.

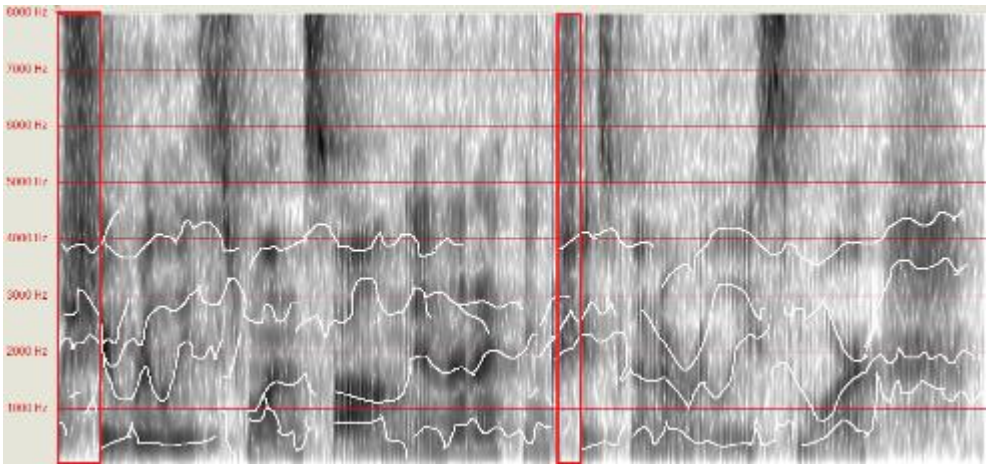
Figure 67. Artsy Madwoman - "like second shots for."



Source: Artsy Madwoman (2021, 0:59-1:51).

It is possible to observe that most of the energy is concentrated in the F3 to F4 range. Also, the energy decreases completely below 1,500 Hz and the absence of energy in the F1 means the absence of a voicing bar. Such characteristics indicate that this is a voiceless postalveolar fricative sound /ʃ/.

Figure 78. Elise Ecklund - "she releases one song and then I think she's gonna release more."



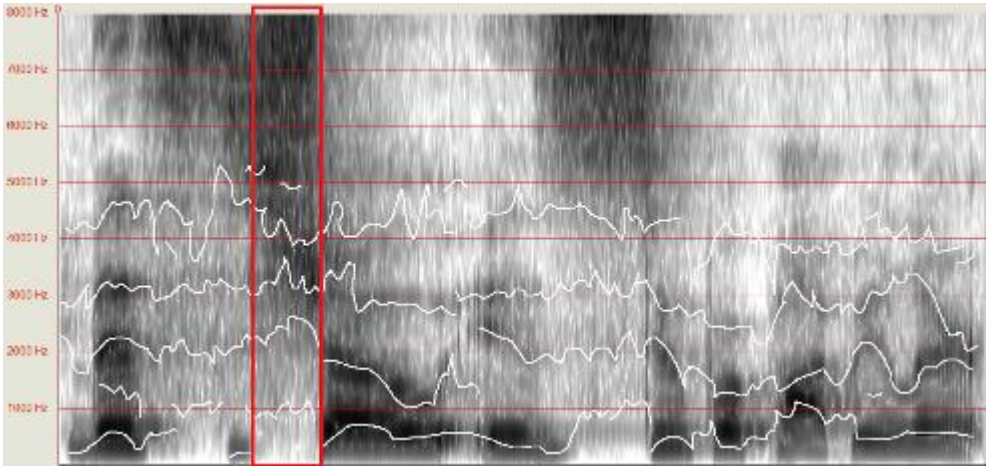
Source: Elise Ecklund (2021, 4:45-4:48).

For the first case, it is possible to observe that the darkest zone i.e. with most energy is placed near the F3. Likewise, there are concentrations of energy near the F4 and above the

6,500Hz which could signify that the sound was produced with more frication than usual. There is also a decrease of energy below 1,500Hz but the low amount of energy near the F1 (instead of the absence of a voicing bar) may indicate that this is a semi-voiced allophone of the voiceless postalveolar fricative /ʃ/.

For the second case, it is possible to find two zones with concentrated energy, namely near the F2 and near the F4, being the last the darkest one (i.e. the zone with the most energy). There is also a decrease of energy below 1,500Hz approximately. The presence of low energy near the F1 instead of the absence of a voicing bar may be due to background noise.

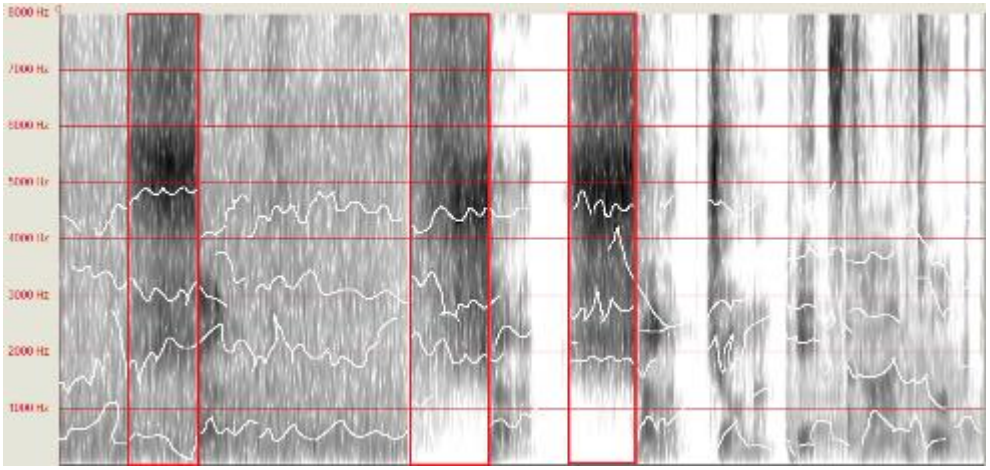
Figure 69. Emmymade - "this show is super popular."



Source: emmymade (2021, 0:33-0:36).

In this case, there is a decrease of noise below 2,000 Hz approximately. There is also the absence of a voicing bar even considering that it is possible to observe a low amount of energy near the F1. Besides, there is a concentration of energy in the F3 to F4 range. Such characteristics indicate it is a voiceless postalveolar fricative sound. There is also a concentration of energy above 5,000 Hz and near 7,000 Hz and 8,000 Hz. This may occur because of excessive frication of the sound.

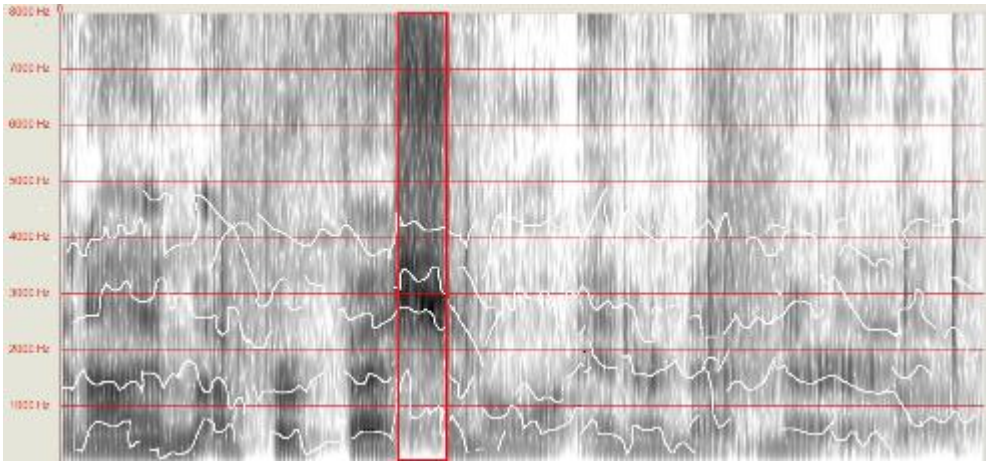
Figure 70. Markiplier - "shut up shut up shut up Jonah."



Source: Markiplier (2021, 4:32-4:34).

Although in the three cases there is a concentration of energy near the F4, in the first case, it is not possible to find the absence of a voicing bar probably because of the presence of background noise. In cases two and three (from left to right) there is possible to observe a decrease of energy below 1500Hz and the absence of a voicing bar which would confirm it is a voiceless postalveolar fricative sound.

Figure 71. Rybonator - "and I am a professional internet nerd."

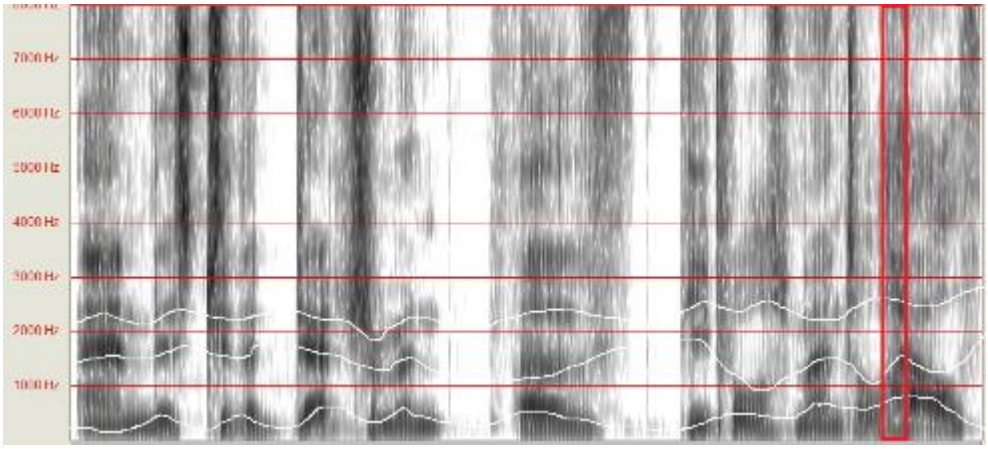


Source: Rybonator (2021, 0:06-0:07).

There is a postalveolar fricative /ʃ/ sound as most of the energy is concentrated between the F3 and F4 range. Also, noise decreases noticeably below 2,000Hz. The absence of the voicing bar indicates it is a voiceless sound.

4.4.8 Voiceless glottal fricative /h/

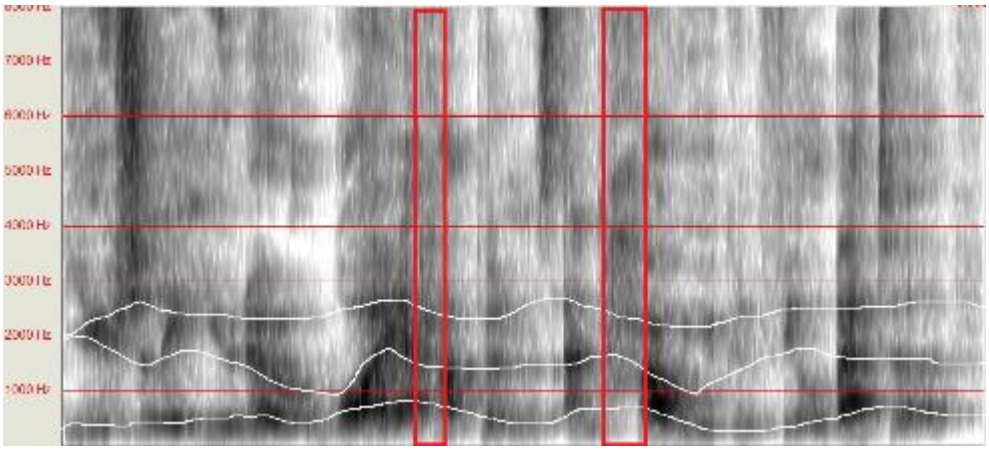
Figure 72. CDawgVA - "to the detriment of people's mental health."



Source: CdawgVA (2021, 1:33-1:35).

The fricative sound /h/ is particular in the sense that according to Ogden (2017), it adopts the characteristics of the vowel sound that follows it, making it hard to analyze and pinpoint which characteristics are inherent to the voiceless glottal fricative sound. Because of this, the only constant characteristics that can be appreciated on a spectrogram are the formants F1, F2, and F3 in the transition of a vowel sound plus the lack of a voicing bar in frequencies below 1,000 Hz. This sound is not /h/ as there is a constant voicing bar in the frequencies below 1,000, making it a fully voiced sound.

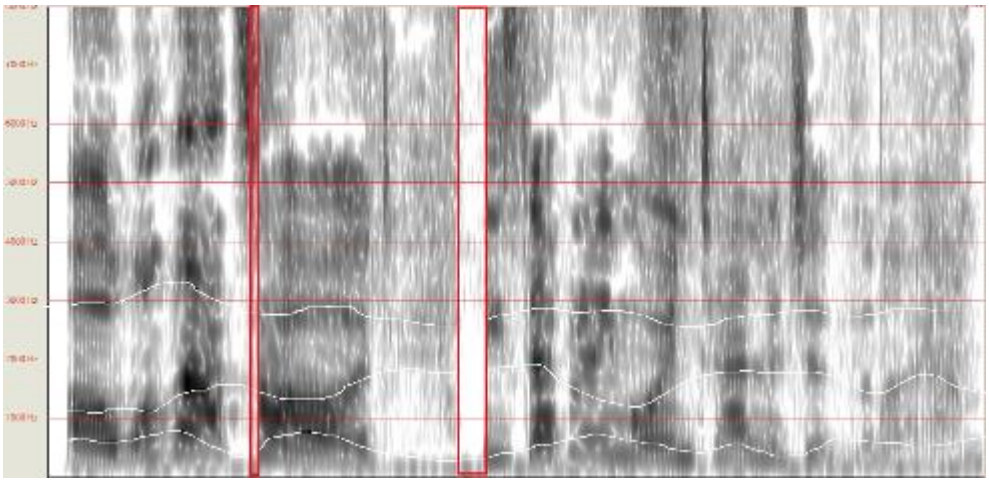
Figure 73. Daniel Howell - "usually you know how it happens and how you can feel better."



Source: Daniel Howell (2017, 1:26-1:30).

In this case, both occurrences of /h/ fulfill both characteristics; however, the first one is being partially voiced and therefore it is an allophone of /h/. The second one has energy near the voicing bar in the lower frequencies, which gives it a certain grade of voicing but is much lower than the sounds around it.

Figure 74. Emilia Clarke - "well that's hard, he is definitely more intelligent than I am."

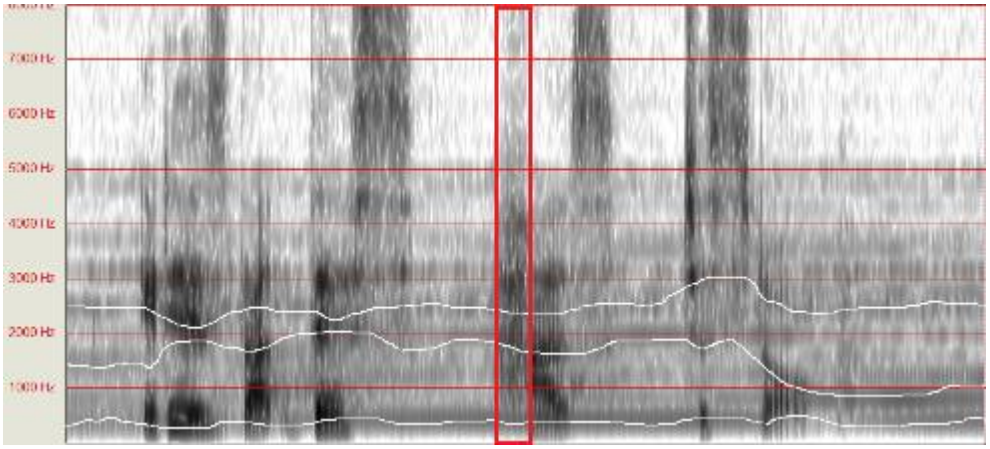


Source: Vanity fair (2019, 1:03-1:06).

There are two instances where a /h/ sound should be produced. The first one of them. In the first instance, there is a total elision of the sound; the spectrogram highlights a section where

the sound should be, but it instead is how the previous consonant starts to articulate the following vowel sound. In the second instance, there is an /h/ sound but has a voicing bar which makes it an allophone of the sound.

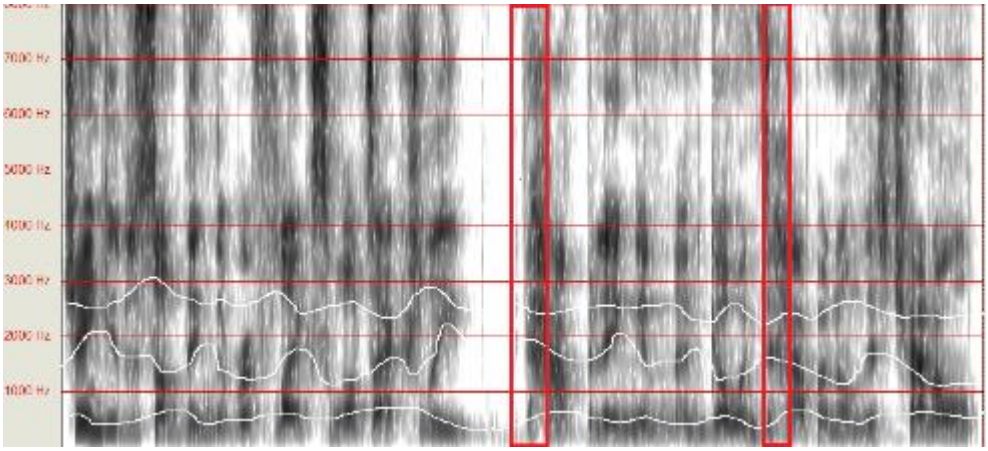
Figure 75. Emma Watson - "it is that this has to stop."



Source: English Speeches (2017, 0:35-0:38).

This sound is a /h/ sound as it possesses the characteristics of a vowel sound. It has energy in frequencies below 1,000 Hz but is low that should not involve the voicing of the sound. The spectrogram absorbed too much background noise that some characteristics are not expressed on the spectrogram like the amount of energy in some sections.

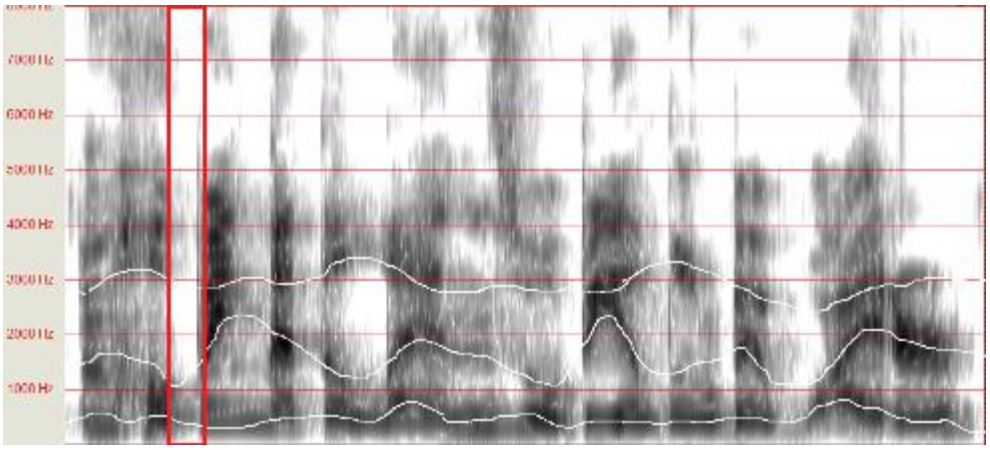
Figure 76. Tomska - "happening and going to happen on this channel."



Source: TomSka & Friends (2021, 0:16-0:18).

The first occurrence of /h/ is inconclusive as the sound could be considered an allophone due to it being influenced by the vowel sound next to it or a short /h/ sound. The second occurrence is an allophone of /h/ as it has energy in the position of a voicing bar although is lower than the segments that surround it.

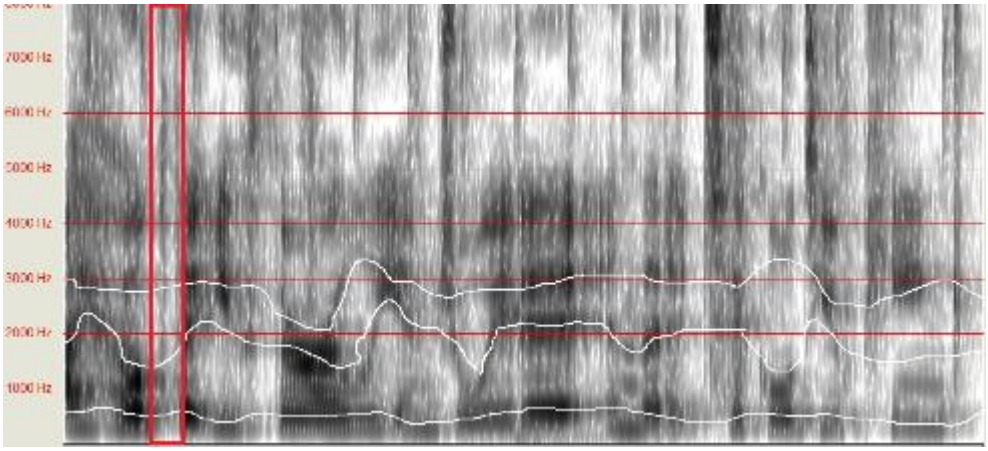
Figure 77. Artsy Madwoman - "I'm hoping that the lens will be a little bit wider."



Source: Artsy Madwoman (2021, 1:24-1:26).

The sound produced is an allophone of /h/ as although the energy is lower than the segments that surround it, the energy is part of the voicing bar. A characteristic that stands out is the lack of energy in higher frequencies, just some noise between 3,000 and 5,000 Hz.

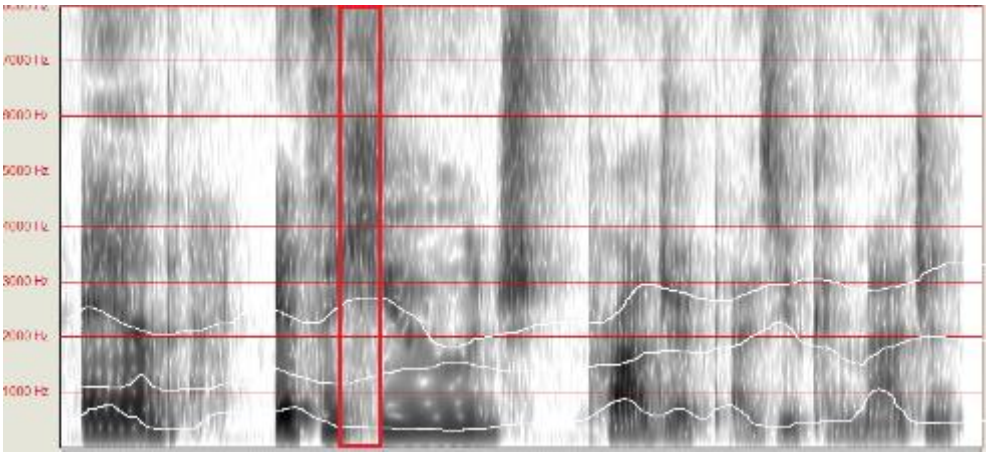
Figure 78. Elise Ecklund - "I hope that you thoroughly enjoyed it."



Source: Elise Ecklund (2021, 1:47-1:49).

The sound produced is an allophone of /h/ as, although lower than the segments that surround it and is inconsistent, the occurrence still has energy in the voicing bar. The voicing bar is proof of vibration in the vocal folds in a sound that when produced in isolation should not have any vibration.

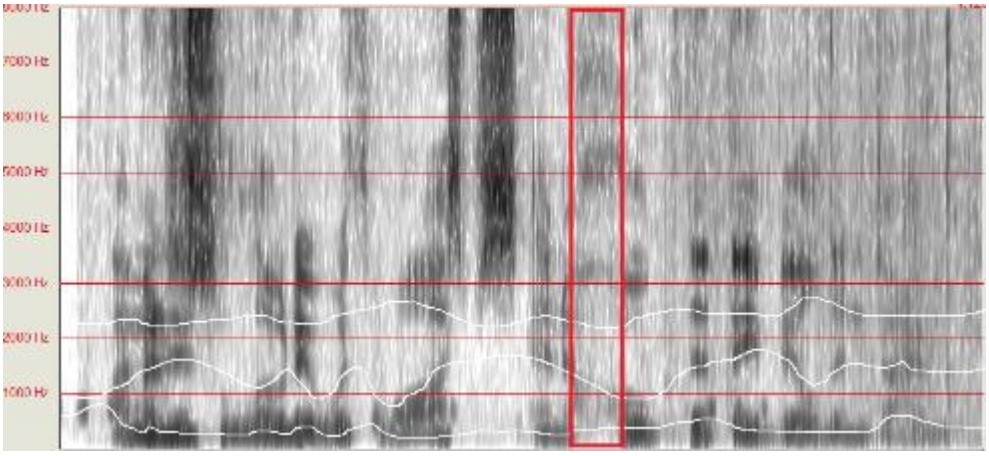
Figure 79. Emmymade - "and it's a huge project."



Source: emmymade (2021, 0:52-0:54).

The sound produced is an allophone of /h/ as, although lower than the segments that surround it and is inconsistent, the occurrence still has energy in the voicing bar. The vibration of the vocal folds may not be present in reality and it could be shown in the image due to the background noise of the audio.

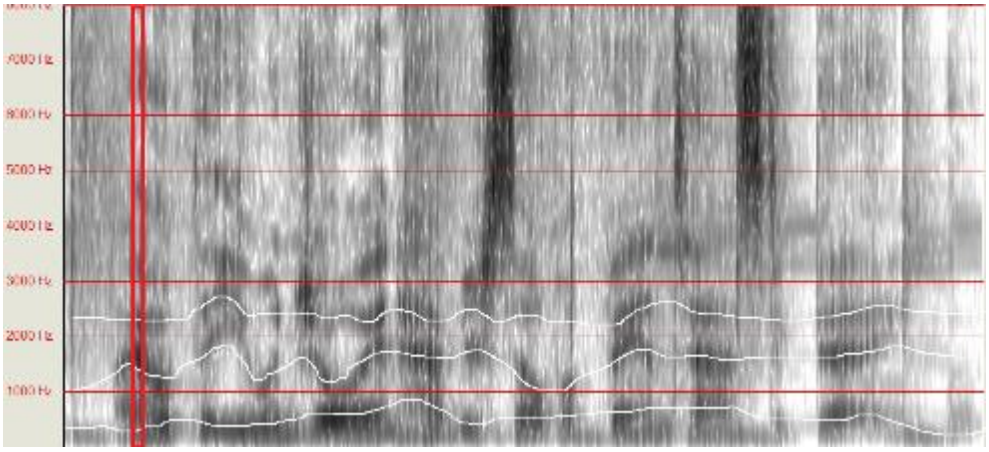
Figure 80. Markiplier - "I'm just gonna load it with wasps and hope that they build in time."



Source: Markiplier (2021, 19:31-19:35).

This is a /h/ sound as it possesses the characteristics of a vowel sound but it lacks a voicing bar frequencies below 1,000.

Figure 81. Rybonator - "I hope that you enjoy."



Source: Rybonator (2020, 1:26-1:27).

This is a really short occurrence of the sound, but it contains a full voicing bar during its duration, making a different sound. It is worth considering the amount of noise in the spectrogram due to the background music present on the audio.

5. Results

In the following section, an overarching analysis of the spectrograms will be displayed, with the determination of generating material and pedagogical suggestions for the teaching of English sounds in the Chilean classroom. This determination is based on the sounds fulfilling certain characteristics that the sound and material used to teach them has, like consistency for the teaching of the sound according to the position of the sound in the syllable (Onset or the consonant cluster before the vowel and Coda the cluster after the vowel sound on a syllable), intelligibility, and comprehensibility.

5.1 Voiced labiodental fricative /v/

Number of occurrences extracted	12	Allophonic variations present	3

When /v/ was part of the Onset, which were 6 occurrences (Figures 2, 3, 6, 9, 10, 11), no variations were present. Of these 6, on 2 occasions there was unusual activity from 5,000 Hz to 8,000 Hz (Figures 6 and 9). This means that there was more frication; however, in one audio there was background music (Figure 6), so the spectrogram could have been influenced by that. Nevertheless, all these occurrences had striation, so they do not pose any difficulties.

When /v/ was part of the Coda, which were 6 occurrences (Figures 3, 4, 5, 7, and 8), it varied, either losing voicing, having more frication than normal, or both. Of these, 3 occurrences happened in Coda in the final position (Figures 4, 5, and 8). These 3 [v̥] are even more devoiced, have little energy, and are close to being a /f/. For pedagogical purposes in the beginning stages, it is advisable to use samples that have the /v/ as part of the Onset. However, teachers must be

careful when selecting samples, and especially careful about the surrounding sounds. For example, if /v/ were to be close to a voiceless consonant, part of the /v/ will accommodate the voiceless sound, making it a more fricative /v/. Unfortunately, such context did not occur for this sound, but a similar phenomenon happened for /ð/, which will be explained in detail in section 5.2.

5.2 Voiced dental fricative /ð/

Number of occurrences extracted	17	Allophonic variations present	1

Of all the 17 occurrences extracted, 16 were part of the Onset and only 1 was part of the Coda (Figure 16), and there were no big differences between these two. It was found that not all speakers showed their tongue while pronouncing the sounds; some could be seen placing it at the back of their teeth, some did not show it at all, and some were unclear; the last two count as not showing it, as they do not provide a visual aid. No relevant differences were found between speakers who showed their tongue — with a total of 10 occurrences (Figures 13, 14, 15, 17, 18, and 21), with 4 of these placing the tongue at the back of the teeth (Figures 17 & 18) — to those who did not — which were 7 (Figures 12, 13, 14, 15, 16, and 20) — in the analysis of the spectrogram. It is relevant to note that some speakers had background music (Figures 12, 15, 16, 18, 20, 21), so this affected any possible comparison.

One /ð/ was present after the voiceless sound /s/ (see Figure 13), making the /ð/ more fricative with a great concentration of energy near the 8,000 Hz range, a phenomenon that is not uncommon for /ð/, but not at that scale. The sound, in the beginning, was more akin to a /θ/, and then it slowly became a fully voiced /ð/; also, in this case, the speaker did not show his tongue.

This could be explained because /s/ is an alveolar sound and the tongue is closer to the back of the teeth rather than in between them.

For pedagogical purposes, when selecting audiovisual material, /ð/ should not be close to a voiceless sound to fully appreciate the voicing of /ð/, as well as the place of articulation. Additionally, the place of articulation should be clearly seen in the video, with the tongue in between the teeth at first. In later stages, it could be explained or shown that it is not rare to pronounce /ð/ with the tongue at the back of the teeth.

5.3 Voiceless dental fricative /θ/

Number of occurrences extracted	14	Allophonic variations present	6

As previously mentioned and concerning acoustics, /θ/ is highly similar to other fricative sounds, making its analysis difficult. However, some characteristics that this sound has are being longer in duration than other fricatives, possessing energy above 3,000 Hz that is not as strong nor concentrated like /s/ and /ʃ/, and lacking a voicing bar.

When the sound was analyzed, there were various instances in which the sound was presented with some degree of voicing, which could be due to the influence of strongly voiced sounds around /θ/. Due to this situation, we cannot advise using this type of material with this sound, yet we cannot discourage the use as well; what we do advise is the use of this material selectively from part of teachers as it is not a common situation the occurrence of allophonic variations yet not as uncommon to blindly choose a sample.

5.4 Voiced alveolar fricative /z/

Number of occurrences extracted	24	Allophonic variations present	9

The sound /z/ is a voiced alveolar fricative sound that can occur on Onset or Coda. When /z/ is part of the Onset, /z/ can only be pronounced as a fully voiced sound but when it is part of the Coda, /z/ can be pronounced as a fully voiced sound or as a partially devoiced one, [z̥]. On a spectrogram, these two possible scenarios could be seen through the voicing bar formed by Formant 1 as the energy fluctuated and vanished on the image when the sound was partially devoiced (go to section 4.4.4, pages 59, 61, and 62 for reference). The energy in the higher frequencies was always consistent, but the devoicing of the sound was a constant variation in the samples analyzed.

The allophone [z̥] was appreciated when the sound was the coda of a syllable and the following sound was a voiceless sound or stop on the speech, pause, or end of the speech. The allophone tends to start as a fully voiced sound that finishes on a voiceless or partially devoiced sound, similar to /s/. The allophone does not present itself as onset as there is no devoicing when the next sound is fully voiced.

In relation to the pedagogical suggestions, when teaching the sound it is advisable to show a sample of the sound from a native speaker in onset position and ask the learners to mimic the sound. This is done to help the students understand the characteristics of the sounds such as the voicing and the vibration present. When teaching the sound as a coda, it is advised to work first on producing the sound and then show the variations present when a native speaker produces the phoneme in connected speech.

5.5 Voiced postalveolar fricative /ʒ/

Number of occurrences extracted	10	Allophonic variations present	4

The sound /ʒ/ is a voiced postalveolar fricative that is distinguished from its voiceless counterpart /ʃ/, by its shorter and less ample frication (or the period in which noise is produced), and most importantly, by the presence of a voicing bar, which indicates vibration. Because it is a fricative, when looking at it in a spectrogram it typically has most of its energy concentrated in the higher frequencies.

In regards to the analysis, 6 out of the 10 speakers selected pronounced this sound correctly, while devoicing or partial devoicing were common occurrences among the rest. This may indicate that the devoicing of this sound is somewhat common in general, but the scope of this analysis is limited, so we cannot say for sure. In addition, we noticed that speakers Emma Watson and Emmymade shortened their vowels right before pronouncing the phoneme /ʒ/.

One way to help students learn how to produce this sound correctly is to let them know what to do physically to pronounce it, as well as drawing comparisons to similar sounds like /ʃ/.

5.6 Voiced postalveolar affricate /ʤ/

Number of occurrences extracted	10	Allophonic variations present	6

Affricates are defined as plosives that release in fricatives (Ogden, 2017). As such, similarly to plosives, they present a gap due to occlusion of the vocal tract at first but tend to be released as fricatives, meaning that they are usually expressed as noise in higher frequencies. Another key component of /ʤ/ is that a voicing bar appears not at the beginning of the sound, but halfway through its production. Due to these characteristics, /ʤ/ shows a particular image on a spectrogram, an image that, curiously enough, did not appear most of the time on the spectrograms analyzed.

A possible reason that our spectrograms of the phoneme /ʤ/ are different from the norm could be because the original material does not always match the prescribed norm, as real life often goes beyond the scopes of theory.

While it might be difficult to teach this to students of elementary levels, it may be possible to explain the circumstances that led to these differences to more advanced and/or older students as well as teach them how to properly produce this sound via vocal tract occlusion.

5.7 Voiceless postalveolar fricative /ʃ/

Number of occurrences extracted	13	Allophonic variations present	3

The /ʃ/ sound is a voiceless postalveolar fricative sound, making a direct counterpart to /ʒ/. The sound is characterized for being a strident sound that is visualized in a spectrogram as noise in the higher frequencies and lacking a voicing bar, just like /s/. A core difference between the two sounds is that the noise is more localized around F3 and F4 in /ʃ/ and the noise starts in lower frequencies than /s/.

Concerning the analysis, /ʃ/ had a tendency to follow the pattern with a few exceptions in which the sound was partially voiced. This partial voicing is due to the vowel sounds around the consonant influencing the sound, causing the vocal folds to be used. Nevertheless, due to the consistency of the sound when produced in a real context, it is advisable to use this material for the teaching of differentiation and production of /ʃ/. Likewise, another suggestion is that the aforementioned variations could also be considered in the teaching of the sound and its allophones to students with enough level of knowledge.

5.8 Voiceless glottal fricative /h/

Number of occurrences extracted	13	Allophonic variations present	10

All the occurrences extracted and analyzed of /h/, a voiceless glottal fricative sound, were on Onset position. It was shown that the speakers most of the time did not produce a pure sound; instead, they tended to produce some allophonic variation of /h/ with various degrees of vibration on the vocal folds. These allophones were expressed on the spectrograms with some accumulation of energy in the lower frequencies, between 0 and 1.000 Hz, where the voicing bar is formed; a situation that was not supposed to occur when producing a voiceless sound. This situation can be explained by seeing that /h/ does not possess its own characteristics, instead, it expresses characteristics of the vowel sound present in the nucleus of the syllable which can explain why it is so common for speakers to also produce these fricative consonants with a degree of voicing or fully voiced. This allophone, /ɦ/, is a middle that can be considered similar to /x/ from words like gente y jalea in Spanish.

Due to the aforementioned characteristics of /h/, it is advised to teach this sound by comparing it to the Spanish sound /x/ as it is known by the learners and could be a helpful tool to use it before being able to fully understand and produce the English sound and its allophonic variation.

6. Discussion

This section considered the data analysis and the results section to make pedagogical suggestions to teach the consonants selected. The material and suggestions mentioned in this section are contained in the following link:

https://sites.google.com/umce.cl/pedagogicalsuggestionnmaterial/inicio

Voiced labiodental fricative /v/

One of the biggest problems with this sound is that sometimes it tends to sound like its voiceless counterpart /f/. However, when /v/ was part of the Onset, it had full voicing; in contrast, the /v/ sounds that were part of the Coda were devoiced (Figures 3, 4, 5, 7, and 8), and sometimes the place of articulation was not at all clear. When selecting a /v/ sample, two things need to be considered: the manner and place of articulation of speakers and the adjacent sounds. On some occasions, the contact between the teeth and lips of speakers was not clear, which is natural in a conversation. As the materials were not created to teach English nor sounds of English, the samples have more complications in contrast with teaching material, which is specifically designed and carefully selected to teach. Moreover, if a voiceless sound were to be close to /v/, /v/ will accommodate to that, thus becoming a more fricative /v/ than in isolation.

Once the most basic form of /v/ is acquired, more advanced occurrences can be explained. At first, isolated occurrences should be selected, so students can see how the sound is pronounced and mimic it. At later stages, other occurrences of /v/, like the devoiced one [v̥] which is very common when in Coda and final position, can be shown in class.

Voiced dental fricative /ð/

Only two problems arise when encountering this sound: A clear place of articulation from speakers and adjacent sounds. It was not uncommon that some speakers did not put their tongues in between the teeth; rather, they placed them at the back of the teeth (which is the only place left to produce this sound). Both places are acceptable, but for teaching purposes, it is convenient to show how speakers place the tongue in between the teeth at first. At later stages, teachers can show students that speakers also place their tongues at the back of their teeth. Having stated this, the next aspect to consider when selecting samples is the surrounding sounds. On one (and the only) occasion, there was a /s/ before a /ð/ (See figure 13), and this context made this /ð/ fricative and devoiced, almost similar to a /θ/. Additionally, in this very context, the speaker did not show his tongue in between the teeth; as /s/ is an alveolar sound, the tongue is closer to the back of the teeth than in between them.

For teaching purposes, any audiovisual material needs to show how the speaker clearly places the tongue in between the teeth and no voiceless sound should be near it at first. At later stages, once the sound has been dealt with in isolation, similar contexts as the one mentioned earlier could be shown to students, thus exposing them to natural occurrences of the language and sounds. These *rare* or *atypical* pronunciations are actually common. Some sounds influence others, thus the pronunciation, place, and manner of articulation also change from the standard and isolated pronunciation of sounds. Using audiovisual material will likely contain plenty of variations, and that is how language is; neglecting these occurrences would be neglecting real-life pronunciation.

Voiceless dental fricative /θ/

Even considering that /θ/ sound does have a few specific characteristics that make it recognizable (such as its big amount of energy and the absence of a voicing bar), it tends to vary a lot regarding the place and manner of articulation in which it occurs. For instance, while it

should be found as a voiceless fricative sound in all of its occurrences (as in figure 22 on page 55); there were several occasions in the analysis in which /θ/ was produced with some degree of voicing (see figure 24 on page 56). The reason for this might be that in most cases it was affected by the energy of the strongly voiced sounds surrounding /θ/ (as is the case of the diphthong /aɪ/ in figure 24 on page 56). Another important characteristic of this sound is that its energy increases from 3,000 Hz and above, but in some cases of the study it was possible to observe that the energy was spread almost all over the occurrence (see figures 28 and 31 on pages 59 and 61 respectively).

Due to the several variations in the examples analyzed, we consider that the use of this study for the teaching of /θ/ is not very recommendable; this is because the frequent occurrence of allophonic variations makes it particularly difficult to define how the sound tends to behave. Notwithstanding, considering that in most cases the sound does present some of its proper characteristics, we cannot discourage completely the use of social media material as a tool for teaching this sound. Instead, we advise that teachers use this material selectively according to the context and the level of proficiency of their students. A good suggestion for this sound would be to use the audiovisual material to show students how the sound can be produced and how it varies according to each of both possible places of articulation: Dental (see figure 23, case 1 on page 56), and interdental (see figure 23, case 2 on page 56).

Voiced alveolar fricative /z/

The main conflict that /z/ faces is the inconsistencies when it is part of the Coda, as it can be partially devoiced, [z̥], being more closely perceived as /s/. This is mostly perceived when /z/ is followed by a voiceless sound or a stop in the utterance due to a pause, or the end of the utterance as a whole. In consequence, it is suggested that teachers explain the differences present in the traditional English taught in classrooms and the English spoken by native speakers when teaching using the material provided. Additionally, the material provided can help with the pronunciation of /z/ in the final position in a syllable and help with the production of the next sound when producing utterances.

Concerning /z/ as Onset, it does not present any allophonic variations that affect the learners' perception of the sound, as it is usually influenced mainly by the nucleus of a syllable and, therefore, it should always be voiced. In relation to this, material that integrates /z/ as Onset does not generate any problem in a classroom due to the previously mentioned characteristics. Suggestions regarding /z/ as Onset, are to show the learners instances in which the phoneme is used in said position and ask them to recreate the same sound and then incorporate it into their speech when it corresponds to use it. If the acquisition of the results of the sound is problematic, the use of tongue twisters in which the sound is used can be used in the classroom to facilitate the reproduction of the sound.

Voiced postalveolar fricative /ʒ/

This sound shares some characteristics of its voiceless counterpart since both have most of their energy concentrated in the higher frequencies of the spectrum. The main difference between these sounds is the presence of vibration when producing the voiced postalveolar fricative /ʒ/, but this also means that the biggest issue encountered when analyzing samples of this phoneme was devoicing, so the /ʒ/ sound had a tendency to sound more like a /ʃ/. For example, as you can see in figure 44, the speaker Emilia Clarke shows no voicing bar when saying "version", while in figure 50, the speaker Markiplier clearly displays a voicing bar when saying "explosion".

The samples we gathered were mostly consistent when it came to the correct pronunciation of the sound. It is important to tell any potential learners of English that devoicing is something that can happen with /ʒ/ when you shorten vowels right before producing this sound, as in figure 45, for example, when the speaker Emma Watson says "decision", you can see how short her vowel is in the spectrogram. The same happens in figure 49 when the speaker Emmymade says "version", the vowel right before the phoneme /ʒ/ is shortened, and it becomes

especially apparent when you compare it to figure 51, when the speaker Rybonator says the same word, you can see his vowel is longer.

A simple way to teach this sound to learners of English would be to make the connection between /ʒ/ and /ʃ/, telling students that the first sound is what you get when you pronounce /ʃ/ while adding vibration, but specifying that the phoneme /ʒ/ needs more pulmonic air strength, so learners need to use their thoracic muscles in order to produce it clearly.

The sound /ʒ/ is also present in some varieties of Chilean Spanish, specifically in words like “llave” or “lluvia”, and it could be used to show students how to pronounce it.

Voiced postalveolar affricate /ʤ/

This sound, being an affricate, has characteristics of both plosives and fricatives when produced. When viewing it on a spectrogram, it should exhibit a slight gap between the initial part of the sound and whatever comes before it, as well as having most of its energy concentrated in the higher frequencies of the spectrum, like a fricative sound. The samples collected did not all follow this pattern, however. Some speakers had a constant voicing bar throughout the sound (when the voicing bar for this sound should start halfway through its production), while some of them did not have a voicing bar at all, turning their /ʤ/ into something more akin to a voiceless fricative. For example, in figure 54, the speaker Emilia Clarke shows a gap before the /ʤ/ sound, but there is no voicing bar present, which makes her release similar to a voiceless fricative. On the other hand, in figure 56, the speaker Tomska shows the gap, while also showing a voicing bar that shows up halfway through the sound, like it is supposed to.

While the amount of noise present in the samples gathered could have affected the spectrogram readings, the results showed that many speakers did not pronounce this sound correctly, which could indicate that people in real life are changing the way they pronounce this sound.

A good way to teach it in the classroom would be to compare it to /ʒ/ while adding occlusion of the vocal tract, or a stop, to illustrate what /ʤ/ sounds like. It is also important to mention that some dialects of Spanish include the /ʤ/ sound in words like “yo”.

Voiceless postalveolar fricative /ʃ/

The main characteristics of this sound are that it contains a high amount of energy and that the majority of it can be found concentrated in a specific frequency range (F3 to F4). This consonant is also recognizable for its lack of a voicing bar. Such characteristics are present in most of the cases analyzed and just 3 allophonic variations were found; which translates into consistency in the behavior of the sound that includes being a voiceless consonant (with a few exceptions in which the sound was partially voiced due to the influence of its surrounding voiced sounds), and a considerable expulsion of air.

Due to the aforementioned consistencies, we consider that the use of social media material in the teaching of the /ʃ/ sound is highly recommendable. Some suggestions for the teaching of this sound are to benefit from the big amount of energy that is expelled when the sound is pronounced: teachers can for instance ask their students to try pronouncing the Spanish onomatopoeia to ask for silence: “Shhh” but with a considerable increase of energy/air expulsion (thus getting duration and volume variations) along with posterior decreases. This type of exercise could also be helpful when comparing /ʃ/ and /ʤ/ sounds; adding some voicing (vocal cords vibration) to the /ʃ/ sound would help students understand the difference between two

sounds as well as voiced versus voiceless sounds. This comparison could also be helpful for the students to recognize the place where /ʃ/ and /ʤ/ postalveolar fricative sounds are articulated.

Voiceless glottal fricative /h/

The sound /h/ has the distinction that it does not have many distinctive characteristics, acoustically speaking, that help to identify it among other sounds; even further, it usually adopts the characteristics of the nucleus in a syllable. Due to this distinction, it is advisable to use influencers or model speakers as referential material to teach the sound in relation to how the sound adapts to each specific context, as well as use the students' L1 to teach this sound using a contrastive approach. A suggestion is to use the sound /x/ that is found in Spanish words like "gente", "jirafa", or "jalea" and ask students to produce the sound in words in English to then ask them to start producing the sound without vibration of the vocal folds to steadily transform the sound into /h/. Another suggestion is to ask the students to heavily breathe or sigh at the beginning of words that begin with the sound /h/ to help them make the sound.

7. Limitations

In this section, we will discuss the factors, shortcomings, and limitations that hinder the reliability of the investigation. The limitations will be presented according to the order in which they were encountered, those being the lack of previous investigations on the topic at hand, the scope of the investigation, and the analysis of samples.

The first aspect that limits the research is the lack of studies and agreements by the scientific community on the matter. There is a lack of research on Influencers from a linguistic standpoint, as most of the research done on them focuses mainly on the economic and social effect they have on the masses and how brands and companies deal and associate with them. Additionally, concerning linguistic and phonemic discussions, there is a lack of consensus about what are the problematic sounds for Spanish speakers learners of English plus little to no research specifically done on Chilean learners of English.

The second group of limitations found is related to the scope of the research. On one hand, selecting only problematic consonants for Spanish speakers learners of English limits the scope in which the material and pedagogical suggestions are applicable in the classroom, as it does not incorporate all the sounds that are necessary to produce intelligible utterances in the English language. On the other hand, the format used in the study, the number of researchers, and the time constraints limited the length of the dissertation.

The last aspect that limits the research is the samples utilized in the analysis, as they were not originally designed to be analyzed in any capacity. This is due to the samples being taken from audiovisual material destined to appeal to mass audiences, so it tends to have background noise, music, and edited fragments; in other words, elements that are usually not present in the speech samples used in this type of study, to make sure that the spectrograms displayed for the acoustic analyzes are not negatively affected.

8. Conclusion

This research was done to help teachers deal with a big weakness of TEFL in Chile at the school level, which is that students are not achieving the expected level of proficiency, mainly if they come from public or private-subsidized schools, which constitute around 90% of the total student population, at least from 10th grade. The reasons for the low level of proficiency may vary, but it was found that three main aspects were present: the insufficient English level of proficiency of Chilean teachers, the lack of exposure to the language to students, and the lack of weekly English lessons. Consequently, EFL learners do not acquire certain knowledge like pronunciation, vocabulary, and grammatical structures, and have suboptimal performance when using the language. To tackle one of these shortcomings, specifically pronunciation and the production of sounds, this study proposed pedagogical suggestions and free access material based on public access references and videos from Influencers, actors that are relatable and have a degree of credibility in the eyes of the students and learners.

The research retrieved Influencers' audiovisual samples from the internet and acoustically analyzed the audio from said samples to determine the segmental features of consonants of casual English, and which characteristics they needed in order to integrate them into TEFL in Chilean classrooms from 7th to 12th grade. The analysis consisted of identifying the characteristics of sounds on a spectrogram as produced by the speakers while explaining which phenomena occur when the sounds are produced in connected speech in a casual context. After the analysis, the pedagogical suggestions were designed based on the audiovisual material and the spectrograms, presenting guidelines when selecting material and how to teach the consonants selected.

The pedagogical suggestions and the material designed aim to incorporate the variety of casual English into the classroom as it is the register more commonly used by speakers of any dialect of English. This is also an attempt to incorporate the register and this form of English into the mainstream forms of TEFL, as English is usually taught using material strictly designed to be used in a classroom, which ends up not reflecting the real use of the language.

9. **References**

Abidin, C. (2016). Please Subscribe! Influencers, Social Media, and the Commodation of Everyday Life.(PhD Thesis) The University of Western Australia, Perth, Australia.

Adela (2017). The influence of using audiovisual media towards students' pronunciation mastery of eighth grade at the second semester of SMPN01 in the academic year of 2015/2016 .(Bachelor's degree Thesis) STATE INSTITUTE OF ISLAMIC STUDIES RADEN INTAN LAMPUNG, Aceh, Indonesia.

Akmajian, A., Demers, R. A., Farmer, A. K., & Harnish, R. M. (2001). Linguistics: An Introduction to Language and Communication (Fifth Edition,fifth edition). The MIT Press.

Alarcón, M. (2013, June 6). Simce de Inglés: de los 100 colegios con mejor puntaje 99 son particulares pagados. BioBioChile - La Red de Prensa Más Grande de Chile. https://www.biobiochile.cl/noticias/2013/06/06/simce-de-ingles-de-los-100-colegios-con-mejor-puntaje-99-son-particulares-pagados.shtml

Arellano, R. Cisterna, S. (2009). Phonetics and Pronunciation in the Teaching of English in Secondary School Programmes.(Bachelor's degree Thesis) Universidad Austral de Chile, Valdivia, Chile.

Artsy Madwoman. (2020, July 29). *Can You Dry Flowers In A Food Dehydrator?* [Video]. YouTube. https://www.youtube.com/watch?v=F6O-7mtEFrM&ab_channel=ArtsyMadwoman

Artsy Madwoman. (2021a, September 9). *DIY HALLOWEEN SOLA WOOD FLOWERS & MAKING A SPOOKY FLORAL PIECE!* □ ❤[Video]. YouTube. https://www.youtube.com/watch?v=Zwu2l9CpUtc&ab_channel=ArtsyMadwoman

Artsy Madwoman. (2021b, October 19). *THE GIANT PUMPKIN REGATTA & GETTING A NEW VLOG CAMERA!* [Video]. YouTube. https://www.youtube.com/watch?v=8S5IcuNuCAo&t=25s&ab_channel=ArtsyMadwoman

Bardone, A. & Gargiulo, C. (2014). Aprendizaje en las escuelas del siglo XXI: Nota 6: Normas y costos de la infraestructura escolar. Banco Interamericano de Desarrollo. https://publications.iadb.org/publications/spanish/document/Aprendizaje-en-las-escuelas-del-siglo-XXI-Nota-6-Normas-y-costos-de-la-infraestructura-escolar.pdf

Bellei C, C. (2013). El estudio de la segregación socioeconómica y académica de la educación chilena. Estudios Pedagógicos (Valdivia), 39(1), 325–345. https://doi.org/10.4067/s0718-07052013000100019

Berryman, S. E., & Region, W. B. E. C. A. (2000). Hidden Challenges to Education Systems in Transition Economies. World Bank, Europe and Central Asia Region.

Boersma, P., & Weenink, D. (n.d.). Praat: doing Phonetics by Computer. Praat. Retrieved December 8, 2021, from https://www.fon.hum.uva.nl/praat/

Borst, J. M. (1956). The Use of Spectrograms for Speech: Vol. Volume 4 Issue 1 pp. 14–23 (Journal of the Audio Engineering Society ed.). AES.

Boser, U. (2017, May 31). Isolated and Segregated. Center for American Progress. https://www.americanprogress.org/issues/education-k-12/reports/2017/05/31/433014/isolated-and-segregated/

Bustos, P. (2019, February 7). Estudio sobre Inglés: 7 de cada 10 alumnos de 3° medio no logran conocimiento de 8° básico. https://www.facebook.com/teletrece. https://www.t13.cl/noticia/nacional/estudio-sobre-ingles-7-de-cada-10-alumnos-de-3-medio-no-logran-conocimiento-basicos-de-8-basico

Burris, C., Vorperian, H., Fourakis, M., & Kent, R. (2011). Acoustic analysis software: A quantitative and qualitative comparison of four systems. ASHA Convention. Session 8855, Poster 373 . San Diego, California. Published.

Cámara Chilena de la Construcción. (2016). Apartado III INFRAESTRUCTURA QUE NOS INVOLUCRA DE USO SOCIAL EDUCACIÓN. http://infraestructuraparachile.cl/wp-content/uploads/2016/03/10involucra_educacion.pdf

CdawgVA. (2020, January 17). *Voice Actor Reviews MORE GOOD Voice Acting* [Video]. YouTube. https://www.youtube.com/watch?v=KiE9cknScOg&ab_channel=CDawgVA

CdawgVA. (2021a, October 6). *I spent 72 hours locked in a Japanese hotel room* [Video]. YouTube. https://www.youtube.com/watch?v=0_EVvtaq95Y&t=471s&ab_channel=CDawgVA

CdawgVA. (2021b, October 20). *When Will I Leave Japan?* [Video]. YouTube. https://www.youtube.com/watch?v=Viy79GKbLDY&ab_channel=CDawgVA

Coe, N. (2001). Speakers of Spanish and Catalan. Learner English, 90–112. https://doi.org/10.1017/cbo9780511667121.008

Correa, T., Hinsley, A. W., & de Zúñiga, H. G. (2010). Who interacts on the Web?: The intersection of users' personality and social media use. Computers in Human Behavior, 26(2), 247–253. https://doi.org/10.1016/j.chb.2009.09.003

Daniel Howell. (2017a, October 11). *Daniel and depression* [Video]. YouTube. https://www.youtube.com/watch?v=Wp2TUPo5W0c&t=48s&ab_channel=DanielHowell

Daniel Howell. (2017b, November 29). *Internet Support Group 10* [Video]. YouTube. https://www.youtube.com/watch?v=Em_vDFxlnI8&ab_channel=DanielHowell

Darville, P., & Rodríguez, J. (2007). Institucionalidad, Financiamiento, y Rendición de Cuentas de Educación. Ministerio de Hacienda. https://www.dipres.gob.cl/598/articles-21658_doc_pdf.pdf

De Améstica, C. (2013). ¡Oh my God! El fracaso de los planes para enseñar Inglés en los colegios y convertir a Chile en un país bilingüe. LaSegunda.Com. http://www.lasegunda.com/Noticias/Nacional/2013/08/872074/oh-my-god-el-fracaso-de-los-planes-para-ensenar-ingles-en-los-colegios-y-convertir-a-chile-en-un-pais-bilingue

Dean, B. (2021, April 26). Social Network Usage & Growth Statistics: How Many People Use Social Media in 2021? BLACKLINKO. https://backlinko.com/social-media-users

DECRETO 100, Diario Oficial de la República de Chile, 22 of September 2005. https://www.bcn.cl/leychile/navegar?idNorma=242302

DFL 2, Diario Oficial de la República de Chile, 02 of July 2010. https://www.bcn.cl/leychile/navegar?idNorma=1014974&idParte=0

Dimos, J. Groves, S. Powell, G. (2015). Influencer Persona in the Media Engagement Framework. ROI of Social Media, 95–115. https://doi.org/10.1002/9781119199403.ch4

Djafarova, E., & Rushworth, C. (2017). Exploring the credibility of online celebrities' Instagram profiles in influencing the purchase decisions of young female users. Computers in Human Behavior, 68, 1–7. https://doi.org/10.1016/j.chb.2016.11.009

Educación Pública (2017, October 27) La nueva Educación Pública es más que un cambio de administración. Ministerio de Educación. https://educacionpublica.cl/la-nueva-educacion-publica-mas-cambio-administracion/

EF. (2020). EF EPI 2020 - EF English Proficiency Index. https://www.ef.com/cl/epi/#

Different Choice. (2019, May 20). *Game of Thrones | Last Interview for "The Cast Remembers" by Emilia Clark | Daenerys Targaryen-2019* [Video]. YouTube. https://www.youtube.com/watch?v=xWjKFNDhpQk

emmymade. (2021a, August 26). *BUBBLE Potatoes Magically Self-Inflate 🌬Into a Crispy Snack* [Video]. YouTube. https://www.youtube.com/watch?v=rzo7gP7hraY&ab_channel=emmymade

emmymade. (2021b, October 9). *Make Squid Game Dalgona Candy in 5 Minutes | Homemade Ppopgi* [Video]. YouTube. https://www.youtube.com/watch?v=qjUXJ8XN3F4&t=18s&ab_channel=emmymade

emmymade. (2021c, October 21). *Turducken - A Duck Inside Of A Chicken Inside Of A Turkey - Roulade* □🐓🦃🍥[Video]. YouTube. https://www.youtube.com/watch?v=qlnE6o82tCw&ab_channel=emmymade

emmymade. (2021d, October 24). *Air-Fried Doughnuts Take 3 Minutes To Cook But Are They GOOD?* [Video]. YouTube. https://www.youtube.com/watch?v=paF2ahi_s08&ab_channel=emmymade

Espinoza, M. G. T., Cárdenas, Y. I. C., Martinez, C. V. P., & Saavedra, F. I. B. (2021). The Use of Audiovisual Materials to Teach Pronunciation in the ESL/EFL Classroom. South Florida Journal of Development, 2(5), 7345–7358. https://doi.org/10.46932/sfjdv2n5-074

Farrel, T., & Martin, S. (2009). To Teach Standard English or World Englishes? A Balanced Approach to Instruction. English Teaching Forum, 2–7.

Fernandez, J. H., Rojas, J., & British Council (Mexico) Staff. (2018). English Public Policies in Latin America. British Council.

Finch, D. F., & Lira, H. O. (1982). A Course in English Phonetics for Spanish Speakers. Adfo Books.

Gavaldà, N. (2021, September 28). Espectrograma. Núria Gavaldà. Retrieved February 21, 2022, from https://nuriagavalda.com/espectrograma/

Gen, R. (n.d.). Language Registers. Genconnection. Retrieved December 12, 2021, from http://genconnection.com/English/ap/LanguageRegisters.htm

Gong, W.; Lim, E.; Zhu, F.(2015). Characterizing silent users in social media communities. Proceedings of the Ninth International AAAI Conference on Web and Social Media. 140-149.

Gonzalez, R. (2018). Segregación Educativa en el Sistema Chileno desde una Perspectiva Comparada. Ministerio de Educación. https://centroestudios.mineduc.cl/wp-content/uploads/sites/100/2018/03/Capítulo_-Segregación-Educativa-en-el-Sistema-Chileno-desde-una-perspectiva-comparada.pdf

Gut, U. (2014). Introduction to English Phonetics and Phonology. Peter Lang Gmbh, Internationaler Verlag Der Wissenschaften.

Elise Ecklund. (2021, July 15). *taking myself on a date to McDonald's* ❤[Video]. YouTube. https://www.youtube.com/watch?v=S8eYmxWis1I&feature=youtu.be

Elise Ecklund. (2021a, March 5). *Testing Awful Guitar Products That Shouldn't Exist!* [Video]. YouTube. https://www.youtube.com/watch?v=UiTK3Hd4AGo&t=50s&ab_channel=EliseEcklund

Elise Ecklund. (2021b, October 5). *I dyed my hair PINK *i actually messed up this time** [Video]. YouTube. https://www.youtube.com/watch?v=Yr77xKkArxQ&ab_channel=EliseEcklund

English Speeches. (2017, June 22). *ENGLISH SPEECH | EMMA WATSON: Gender Equality (English Subtitles)* [Video]. YouTube. https://www.youtube.com/watch?v=nIwU-9ZTTJc&t=39s&ab_channel=EnglishSpeeches

Hagiwara, R. (2009). How to read a spectrogram - Rob Hagiwara. Monthly Mystery Spectrogram Webzone. Retrieved December 8, 2021, from https://home.cc.umanitoba.ca/%7Erobh/howto.html

Halliday, M. (2012). THE USERS AND USES OF LANGUAGE. Readings in the Sociology of Language, 139–169. https://doi.org/10.1515/9783110805376.139

Halliday, M., & Hassan, R. (1985). Language, Context, & Text: Aspects of Language in a Social Semiotic Perspective (Specialised curriculum. Language & learning). Hyperion Books.

Hayes, A. (2021, April 28). What Is the Gini Index? Investopedia. Retrieved December 12, 2021, from https://www.investopedia.com/terms/g/gini-index.asp

Herrada, M., Rojas, D., & Zapata, A. (2013). Enseñanza del Inglés en los Colegios Municipales de Chile ¿Dónde Estamos y Hacia Dónde Vamos? . Foro Educacional(22nd ed.). (pp. 95–108).

Howatt, A. P. R., & Smith, R. (2014). The History of Teaching English as a Foreign Language, from a British and European Perspective. Language & History, 57(1), 75–95. https://doi.org/10.1179/1759753614z.00000000028

Jones, M. (2021, April 8). Dialect vs. Accent: Definitions, Similarities, & Differences. SpeakUp Resources. Retrieved December 9, 2021, from https://magoosh.com/english-speaking/dialect-vs-accent-differences-and-examples/

Joos, M. (1967). *The Five Clocks: A Linguistic Excursion Into the Five Styles of English Usage* (5th Printing ed.). Harco

Joze, M. (2015). An Overview of Contrastive Analysis Hypothesis. Cumhuriyet Üniversitesi Fen Fakültesi, 36(3), 1106–1113.

Jung, M-Y. (2010). The Intelligibility and Comprehensibility of World Englishes to Non-Native Speakers. Journal of Pan-Pacific Association of Applied Linguistics, 14(2), 141-163.urt.

Kang, O., Thomson, R. I., & Moran, M. (2018). Which Features of Accent affect Understanding? Exploring the Intelligibility Threshold of Diverse Accent Varieties. Applied Linguistics, 41(4), 453–480. https://doi.org/10.1093/applin/amy053

Kang, O., Thomson, R. I., Murphy, J. M., Sonaat, S., & Levis, J. (2017). Pronunciation Teaching in the Early CLT Era. In The Routledge Handbook of Contemporary English Pronunciation (pp. 502–519). Taylor & Francis.

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. Business Horizons, 53(1), 59–68. https://doi.org/10.1016/j.bushor.2009.09.003

Karjo, C. H., & Wijaya, S. (2020). The Language Features of Male and Female Beauty Influencers in Youtube Videos. English Review: Journal of English Education, 8(2), 39. https://doi.org/10.25134/erjee.v8i2.2593

Kathirvel, K., & Hashim, H. (2020). The Use of Audio-Visual Materials as Strategies to Enhance Speaking Skills among ESL Young Learners. Creative Education, 11(12), 2599–2608. https://doi.org/10.4236/ce.2020.1112192

Kenneth, C. (2009). Minimum Classroom Size and Number of Students Per Classroom. The University of Georgia. https://www.scarsdaleschools.k12.ny.us/cms/lib/NY01001205/Centricity/Domain/1105/2014-11-19%20Meeting%20of%20Greenacres%20Building%20Committee%20Meeting%20Handout%203%20-%20Classroom%20Size%20Standards.pdf

King, R. D. (1967). Functional Load and Sound Change. Language, 43(4), 831. https://doi.org/10.2307/411969

Ladefoged, P. (2021). A Course in Phonetics 6th Edition (6th ed.). Cengage Learning.

Lambeth, G., Otero, C., & Vergara, D. (2020, March 10). Parte II: La desigualdad es una decisión política [Press release]. https://www.ciperchile.cl/2019/12/10/parte-ii-la-desigualdad-es-una-decision-politica/

Levis, J. (2018). Intelligibility, Comprehensibility, and Spoken Language. Intelligibility, Oral Communication, and the Teaching of Pronunciation, 11–32. https://doi.org/10.1017/9781108241564.004

Ley 20832, Diario Oficial de la República de Chile, 01 of May, 2016. https://www.bcn.cl/leychile/navegar?idNorma=1077040&buscar=20832

Ley 21040, Diario Oficial de la República de Chile, 04 of November, 2017.

Małinowski, B. (1923). *The problem of meaning in primitive languages. In The Meaning of Meaning: A Study of the Influence of Language upon Thought and of the Science of Symbolism*, ed. C Ogden, I Richards, pp. 296–336. London: Kegan Paul Trench Trubner

Marcel, M., & Tokman, C. (2005). ¿Cómo se financia la Educación en Chile? Ministerio de Hacienda.

McMahon, A. M. S. (2002). An Introduction to English Phonology. Amsterdam University Press.

Ministerio de Educación. (2016). Bases Curriculares 7º básico a 2º medio. Curriculum Nacional. Retrieved July 26, 2021, https://www.curriculumnacional.cl/614/articles-37136_bases.pdf

Ministerio de Educación. (2016). Idioma Extranjero: Inglés Programa de Estudio de Octavo básico (1st ed.). MINEDUC. https://bibliotecadigital.mineduc.cl/bitstream/handle/20.500.12365/630/MONO-136.pdf?sequence=1&isAllowed=y

Ministerio de Educación. (2018). UNIDAD DE CURRÍCULUM Y EVALUACIÓN PLAN DE ESTUDIO 2018. Curriculum Nacional. https://www.curriculumnacional.cl/614/articles-34970_recurso_plan.pdf

Montano-Harmon, M. R. (2014). "Developing English for Academic Purposes". Fullerton.: California State University.

Montero, A. (2017, September 26). ¿Cuánto invierten en educación los países de América Latina y el Caribe? Aika Educación. http://www.aikaeducacion.com/tendencias/cuanto-invierten-educacion-los-paises-america-latina-caribe/

Mróz-Gorgoń, B., & Peszko, K. (2016). Marketing analysis of social media – definition considerations. European Journal of Service Management, 20, 33–40. https://doi.org/10.18276/ejsm.2016.20-04

Munro, M. J., & Derwing, T. M. (1995). Foreign Accent, Comprehensibility, and Intelligibility in the Speech of Second Language Learners. Language Learning, 45(1), 73–97. https://doi.org/10.1111/j.1467-1770.1995.tb00963.x

Munro, M. J., & Derwing, T. M. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. System, 34(4), 520–531. https://doi.org/10.1016/j.system.2006.09.004

Nelson, C. L., Smith, L. E. (1985). International intelligibility of English: directions and resources. World Englishes, 4(3), 333–342. https://doi.org/10.1111/j.1467-971x.1985.tb00423.x

Newman, E. J., & Schwarz, N. (2018). Good Sound, Good Research: How Audio Quality Influences Perceptions of the Research and Researcher. Science Communication, 40(2), 246–257. https://doi.org/10.1177/1075547018759345

Nordquist, Richard. (2019). What Is Register in Linguistics? Retrieved from https://www.thoughtco.com/register-language-style-1692038

Norman, J. (2017). Student's Self-perceived English Accent and Its Impact on Their Communicative Competence and Speaking Confidence : An Empirical Study Among Students Taking English 6 in Upper-Secondary School (Dissertation).

http://urn.kb.se/resolve?urn=urn:nbn:se:hig:diva-25351

Nuraeni. (2018). Using Audio Visual Material to Enhance Students' Speaking Skills. TEFLIN International Conference, 65, 132–140.

OECD (2007), Participative Web and User-Created Content: Web 2.0, Wikis and Social Networking, OECD Publishing, Paris, https://doi.org/10.1787/9789264037472-en.

OECD (2020), PISA 2018 Results (Volume V): Effective Policies, Successful Schools, PISA, OECD Publishing, Paris, https://doi.org/10.1787/ca768d40-en.

Ogden, R. (2017). An Introduction to English Phonetics. Amsterdam University Press.

Olivares, I. (2020, February 13). El 88% de los chilenos termina la educación media. La Tercera. https://www.latercera.com/noticia/88-los-chilenos-termina-la-educacion-media/

Owens, A. (2016, July 29). Growing economic segregation among school districts and schools. Brookings. https://www.brookings.edu/blog/brown-center-chalkboard/2015/09/10/growing-economic-segregation-among-school-districts-and-schools/

Page, R., Barton, D., Unger, J. W., & Zappavigna, M. (2014). Researching Language and Social Media: A Student Guide (1st ed.). Routledge.

Pennington, M. C., & Richards, J. C. (1986). Pronunciation Revisited. TESOL Quarterly, 20(2), 207. https://doi.org/10.2307/3586541

Pérez, N. (2013, June 6). 82% de estudiantes de 3ro medio no logró acreditar conocimientos mínimos en Simce de Inglés 2012. BioBioChile - La Red de Prensa Más Grande de Chile. https://www.biobiochile.cl/noticias/2013/06/06/82-de-estudiantes-de-3ro-medio-no-logro-acreditar-conocimientos-minimos-en-simce-de-ingles-2012.shtml

Poulopoulos, V., Vassilakis, C., Antoniou, A., Lepouras, G., Theodoropoulos, A., & Wallace, M. (2018). The Personality of the Influencers, the Characteristics of Qualitative

Discussions and Their Analysis for Recommendations to Cultural Institutions. Heritage, 1(2), 239–253. https://doi.org/10.3390/heritage1020016

Raake, A. (2016). Views on Sound Quality. ICA.

https://www.researchgate.net/publication/311739211_Views_on_Sound_Quality_Communication_Acoustics_Paper_ICA2016-391

Rahmi, S. (2018). THE ROLE OF AUDIO VISUAL TO DEVELOP STUDENTS' PRONUNCIATION .(Bachelor's degree Thesis) AR-RANIRY STATE ISLAMIC UNIVERSITY, Aceh, Indonesia.

Ramírez, F. (2019). Education at a Glance 2019: Análisis de los resultados más relevantes para Chile. Ministerio de Educación. https://centroestudios.mineduc.cl/wp-content/uploads/sites/100/2019/09/EVIDENCIAS-45.pdf

Richards, J. C. (2006). Communicative Language Teaching Today (1st revised ed.). Cambridge University Press.

Rodriguez, P. (2016). La geografía de las oportunidades educativas: Determinando el acceso real de los estudiantes a establecimientos educacionales efectivos para generar políticas públicas que mejoren la provisión de educación de calidad. MINEDUC. https://centroestudios.mineduc.cl/wp-content/uploads/sites/100/2017/07/INFORME-FINAL-F911435.pdf

Rybonator. (2020, May 24). *Rybonator - Channel Trailer* [Video]. YouTube. https://www.youtube.com/watch?v=Pjl3kQXpRAc&ab_channel=Rybonator

Said, C. (2020, February 8). Chile tiene la mayor inversión por alumno entre 14 países de la región. La Tercera. https://www.latercera.com/nacional/noticia/chile-la-mayor-inversion-alumno-14-paises-la-region/410641/

Sakoda, J. M. (1981). A Generalized Index of Dissimilarity. Demography, 18(2), 245–250. https://doi.org/10.2307/2061096

Subsecretaría de Administración Parvularia. (2021, January 7). Historia. Educación Parvularia. https://parvularia.mineduc.cl/historia/

The Upcoming. (2018, May 18). *Paul Bettany and Emilia Clarke ultimate Solo: A Star Wars Story interview* [Video]. YouTube. https://www.youtube.com/watch?v=n0uutQYqKMA&ab_channel=TheUpcoming

TomSka & Friends. (2021a, January 29). ⛔ *No Kids Allowed* ⛔ [Video]. YouTube. https://www.youtube.com/watch?v=Zi4ELCJpxsU&ab_channel=TomSka%26Friends

TomSka & Friends. (2021, May 14). *Will there be an asdfmovie14??? (Update Q&A)* [Video]. YouTube. https://www.youtube.com/watch?v=pceLll-cJ0A&ab_channel=TomSka%26Friends

Unidad de Currículum y Evaluación. (2012). Bases Curriculares para la Educación Básica. Ministerio de Educación de Chile.

Vallejos, S. (2012). Significados que otorgan a los resultados obtenidos en prueba estandarizada SIMCE, Docentes de Matemática y Lenguaje y Comunicación de dos liceos de Enseñanza Media Técnico Profesional de la región Metropolitana.(Master Degree Thesis) Universidad de Chile, Santiago, Chile.

Vanity Fair. (2019, November 6). *Emilia Clarke Takes a Lie Detector Test | Vanity Fair* [Video]. YouTube. https://www.youtube.com/watch?v=yGE71YTSGt8&ab_channel=VanityFair

Watson, E. (2017, February 17). *New interview of Emma Watson with People/EW Network!* [Video]. YouTube. https://www.youtube.com/watch?v=mWSVB7FMmoE&ab_channel=EmmaW.Thailand

WIRED. (2019, March 21). *Markiplier Answers the Web's Most Searched Questions | WIRED* [Video]. YouTube. https://www.youtube.com/watch?v=3c03hBfakms&t=19s&ab_channel=WIRED

Wyrwoll, C. (2014). User-Generated Content. Social Media, 11–45. https://doi.org/10.1007/978-3-658-06984-1_2

Yule, G. (2014). The Study of Language. Cambridge: Cambridge University Press.

Yurtbaşı, M. (2015). Why should Speech Rate (Tempo) be Integrated into Pronunciation Teaching Curriculum?. Journal of Education and Future , 0 (8) , 85-102 .