Universitat Autònoma de Barcelona

DEPARTAMENT DE FILOLOGIA ANGLESA I DE GERMANÍSTICA

AOC'S Prosodic Variation between Two Contexts: Celebrity Interview versus Political Speech

Treball de Fi de Grau/ BA dissertation

Author: Izaskun Berrade Echecón

Supervisor: Dr Maria Josep Solé Sabater

Grau d'Estudis Anglesos

June 2021

ACKNOWLEDGEMENTS

During this part of my academic life, I have been fortunate to receive support in many ways. Without it, directly or indirectly, this dissertation would not have been possible. Firstly, I would like to thank my supervisor, Dr Maria Josep Solé Sabater, for her guidance, her eye for improvement, her patience and kindness and for believing in me. Secondly, I want to acknowledge how greatly these three institutions have shaped me academically and personally: the Autonomous University of Barcelona with special praise to Dr Andrea Huerta, Dr Elisabet Pladevall, Dr Hortènsia Curell and Dr David Owen for their wonderful teaching; the University of the Basque Country, and Dr Eva Delgado for her enthusiasm while I dived into Pragmatics; and finally, the University of Edinburgh for the experience and growth. Thirdly, I am grateful for my classmates and friends Mónica Vazquez, Helena Sánchez, and Marta Galera and for those who have supported me through and through, Uxua Esteban and Colin O'Shea. Finally, I wish to say thank you to my grandparents Rosario and Jesús for their unconditional support, my family for their words of encouragement and Mushu for always being by my side.

TABLE OF CONTENTS

Index of Figures

Abstract

The paper aims to compare the main features in Congresswoman Alexandria Ocasio-Cortez's speech in the context of a celebrity interview, specifically in the *Late Night with Seth Meyers* show and a political speech, specifically in a House Floor speech. On the one hand, I anticipate political speech to be slower-paced, though more emphatic than common speech. On the other hand, common speech is expected to be quicker-paced as there is turn-taking, change of topics and is more open to showing emotions such as humour.

Following Holliday et al. (2020) and Miyauchi (2017) Tigue et al. (2012), Reich (1980), Strangert (2005) and Vadimovich (201), I aim to analyse AOC's speech focusing mostly on the use of F0 variation. While analysing F0 I came up with 6 variables: F0 mean and range, F0 variation, pauses, elongated final -s, creaky voice, and HRT. Results confirm that the celebrity interview context has a fast-paced, high-pitched speech and continuous pitch shifts while political speech favours a slow-paced, although with a wider F0 range as emphasis is used to impact the audience. Creaky voice and elongated final -s are comparable in both contexts whereas the four remaining variables vary.

Keywords: political speech, celebrity interview, F0 variation, creaky voice, pitch, HRT, Ocasio-Cortez.

1. Introduction

In this project, I will analyse the prosodic features present in Congresswoman Alexandria Ocasio-Cortez's speech. AOC, as she is famously known for, is a thirty-one-year-old Latina woman from The Bronx, a New York district known for gang activity, poverty-stricken neighbourhoods, and language diversity –with over seventy-five spoken languages on its streets. Moreover, she was brought up in a Puerto Rican household as her mother immigrated to the United States from Puerto Rico and her father was Puerto Rican raised in The Bronx, making her fluent in Spanish. She went to Boston University and graduated with degrees in both Economics and International Relations. As well as this, she worked for the late Senator Ted Kennedy and also worked on Bernie Sanders' 2016 presidential campaign. I must mention her background as it contributes to what I expect her prosodic features to be as well as possible code-switching either between dialects or languages.

The paper aims to compare the main features in Alexandria's speech in the context of a celebrity interview (CI from now on), specifically in the *Late Night with Seth Meyers* show on the twenty-second of March 2019, and a political speech (PS), specifically in a House floor speech on the twenty-third of July 2020. I chose to analyse these two speeches because I wanted to examine how speech style would affect the use of intonational devices. I expected to find intonational differences between the two settings. On the one hand, I anticipated political speech to be slower-paced, though more emphatic than common speech. It is normally planned to draw the listener's attention towards specific information. Moreover, I expected to find code-switching as it is a tool broadly used in political speech. On the other hand, common speech is expected to be quicker-

paced as there is turn-taking, change of topics and is more open to showing emotions such as humour.

1.1. Literature Review

Drawing on established literature has been especially challenging since most papers tackle specific features of specific people. I used two studies about Barack Obama's political speech as a starting point. Holliday et al. (2020)'s and Miyauchi (2017). The first study aims to explore Barack Obama's speech in the context of intonational variation between Mainstream U.S. English (MUSE) and African American Language (AAL) focusing on the Tones and Break Indices used, then subsequently annotated for affect, defined as "the kind of emotion about some referent or proposition". The results indicated that affect had no apparent association with boundary tones, and affective meaning contrast was not predictive of said tones either. Nevertheless, affect is indeed associated with pitch accent patterns. For instance, L+H* and its phonetic characteristics such as steeper rises and higher peaks are predicted by negative affect. Even though I have not analysed my chosen prosodic features in association with pragmatic meaning, I found this study's methodology interesting. Firstly, how they draw on the linguistic background of the subject in question and secondly, how they structure their data between variables and constants. In this second study carried out by Miyauchi (2017), the speech delivered by President Obama at the Hiroshima Peace Memorial Park is examined and analysed mostly focusing on the tones employed. What the results indicate is that Barack Obama frequently makes use of the Fall-Rise and Mid-Level nuclear tones. The Fall-Rise tone makes the speech delivery sound rhythmical and easy to understand by laying emphasis appropriately. The Mid-level tone makes the speech delivery sound not only calm and impressive, but it also makes other adjoining emphatic tone groups stand out by

contrast and produces a memorable impression. This second study on Barack Obama's speech helped me develop a background of what to expect from a political speech, i.e., to expect a wider pitch range but also less F0 variation as the use of Mid-Level nuclear tones is prominent. This paper is the reason why I expected AOC's political speech to be more serious and slow-paced which translates into a broad use of Mid-Level nuclear tones thereby making tone variation not more frequent but indeed more emphatic.

Furthermore, Tigue et al. (2012) looks at voting behaviour based on its relation to voice quality perceptions of politicians. In Study 1 they changed the voice pitch of recordings of US presidents and requested participants to associate personality traits to the voices and to name the voice they preferred to vote for. The results of Study 1 show that lower-pitched voices were linked to favourable personality traits more than were higher-pitched voices and therefore preferred. In Study 2, they discovered that participants are more prone to choose the candidate with the lower-pitched voice when choosing between two unknown men's voices uttering a neutral sentence. Hence, it raised my interest in looking at F0 mean in my chosen settings. I wish to see if these findings hold in AOC's speech.

Reich (1980) considers pauses as a way "to facilitate certain operations involved in the production and in the perception of speech". Pauses have been proved to improve the accuracy in detecting and recalling lists of digits and letters. Thus, Reich aims to examine the effect of pauses in the perception of sentences firstly through a semantic categorisation task and secondly a sentence recall task. The results illustrate that in sentences where pauses were used between clauses words were categorised quicker and propositions were recalled more accurately. This paper points out the importance of a feature I underestimated, and I decided to test its findings on my study.

Prosodic devices have been widely studied and it does not come as a surprise that speakers use them to change their speech to match their context. Strangert (2005) examines the informative and argumentative functions of prosody through speech samples from two highly skilled speakers: a professional news announcer in a news reading and a famous politician in an interview. Results show that even though both speakers use prosodic devices efficiently, in the news reading, they are used mainly to highlight important information whereas the interview favours a wider range of argumentative and emotional expressive acts reflected in the prosody. These results led me to think that I would indeed find different use of intonational devices in different settings and showed me what to expect from the celebrity interview and the political speech.

Vadimovich (2016) aims to analyse David Cameron's public speeches by focusing on its nuclear tones. Based on an autosegmental approach to intonation, the author proposes a step-by-step method for the analysis of nuclear tones, while also considering the register and the range. The findings show a clear preference for the following tones: fall (H*L), rise (L*H) and fall-rise (H*LH) which have in common the emphatic function on negative factors although a falling tone also marks positive factors – this ambivalence could explain its high usage. I will not be analysing specific tones in my paper, since I am interested in several prosodic variables which have been found to characterize political speech, but I found this paper very interesting because it establishes a clear methodology, identifies the main nuclear tones and their function and provides a good description of the context in which these occur.

Following Holliday et al. (2020), Miyauchi (2017), Tigue et al. (2012), Reich (1980), Strangert (2005) and Vadimovich (2016), I analyse Alexandria Ocasio-Cortez's

speech in the context of a political speech and a celebrity interview focusing mostly on F0 behaviour (i.e., F0 mean, variation, range, tones, etc). This way I will know if those two contexts trigger different features in my subject's speech.

2. Methodology

I decided to carry out an acoustic analysis of prosody in the context of a celebrity interview context and a political speech. To do so, I chose congresswoman Alexandria Ocasio-Cortez as my subject as she seemed a person who would consciously or unconsciously change her speech depending on context.

The data I have collected comes from two videos from the *YouTube* platform. From each video, I analysed a minute and a half. The first one is a celebrity interview in the *Late Night with Seth Meyers* show between Alexandria Ocasio-Cortez and Seth Meyers and the second one is a recording of a House floor speech given by the congresswoman herself. To analyse the prosodic features of both videos, I used the program *Praat* thus I transformed the MP4 files into WAV files. Then, I separated the intonational phrases into tiers and transcribed the speech into TextGrids. While doing so, I discovered how pauses behaved in each context and decided to add them as a variable in my study. Following this, I chose to analyse the frequency and duration of pauses and F0 behaviour in both contexts which led me to notice the presence of High Rising Terminal (HRT). Moreover, while looking at the spectrogram in the *Praat* program, both creaky voice and elongated final -s became features worth looking at because of their frequency of occurrence. While creaky voice and elongated final -s appear to be constant features in her speech, F0 mean, variation and range, frequency and duration of pauses, and HRT appear to vary in each context. For F0 variation I chose specific intonational phrases in

both contexts and focused on how frequent F0 changes are. The pitch range used in each context was calculated as the difference between the lowest and highest pitch. The F0 mean variable indicates which context uses a higher/lower tone overall. The creaky voice feature can be appreciated in the spectrogram and the struggle of the program to identify pitch during creak. I also labelled the instances of HRT and elongated final -s I encountered and finally, I took into account pauses equal to or over 100ms, listing their duration and frequency in each of the contexts.

Therefore, my dataset consists of two long audio WAV files with their respective TextGrid, a basic transcription of the script and a list of the frequency and/or duration of the features within each context, making it easier for me to draw a comparison.

Praat is a free computer software package for speech analysis designed by Paul Boersma and David Weenink of the University of Amsterdam. It allowed me to measure F0 (fundamental frequency or number of glottal pulses per second) at specific points of my WAV files. Measuring fundamental frequency allows us to analyse features related to tone and intonation, in this case, HRT (High Rising Terminal)–a feature of some variants of English where a declarative sentence ends with a rising intonation– F0 mean (overall tone), variation (intonation), and range (or scope). Additionally, the spectrograms provided by *Praat* help to show creaky voice too.

Even though I understand that analysing a minute and a half from each instance of speech might not seem to be sufficient, due to the large number of variables I wanted to analyse I had to restrict myself to rather short excerpts. I encountered some struggles determining if elongated final -s was ultimately a variable worth looking at as final lengthening of segments before a pause is a well-known effect. Thus, I decided only to consider instances of the variable which lasted more than 100ms. Finally, the presence of

creaky voice was something I could clearly observe during the initial analysis although showing written proof of its presence was a challenge as creaky voice is not a feature that is easy to measure.

3. Results and Discussion

The analysis of my data sheds light on all my chosen variables. Firstly, I will talk about pauses and speaking rate, secondly about F0 mean, variation and range, thirdly about elongated final -s and creaky voice and finally about HRT.

3.1. *Pauses and speaking rate*

When using *Praat* to take measurements and analyse my data, I made the most of it and used the TextGrid file to label the pauses too because it struck me as a variable of interest. Pauses as a variable were not one of the features I intended to analyse when I designed my study, but I am glad I added it as I was analysing the data. In the celebrity interview context, we find 78 instances of pause while in the political speech context we only find 62. Given that I am analysing an equivalent span of time, –after removing the interviewer speech from the dataset– and the same person, one would expect to find the same pause pattern. Suffice to say, I did not. In the CI context, there are more instances of pauses, but the mean duration is only 200ms whereas in the PS context, with only 62 instances, the mean duration of a pause is 500ms. I interpret from these findings that pausing is a tool that speakers may use not only to plan their next phrase and facilitate processing but also for effect. For example, pausing might have many pragmatic meanings such as creating expectations or giving time to the receiver to process the information or even add seriousness to a situation. As a result, my findings match those of Reich (1980)'s experiment, where it was shown that "in sentences containing pauses

between clauses [...] propositions were recalled more accurately" this could be the reason political speech tends to have longer pauses, to let people process the speech better. Nevertheless, I would like to focus on what pausing might mean to a person's prosodic field or pattern. Although the speaker may be manipulating this tool, there is no doubt that a faster-paced speech will have shorter pauses. I, therefore, measured speaking rate in words per minute in both contexts. Alexandria speaks at a rate of 142 words per minute in the CI context while in the PS context, she speaks at 124 words per minute. Since in her political speech, AOC is speaking at a faster rate than in CI, yet her pauses in the former are more than twice the duration than in the latter context (500ms vs 200ms), we can conclude that she is exploiting pauses for rhetorical effect in political speech.

3.2. *F0 mean, F0 variation and pitch range*

Pitch, measured as F0 mean, is also a tool speakers use to present themselves. Low-pitched voices are associated with being more trustworthy and confident than high-pitched voices. AOC's mean F0 value in the political speech is 26% lower than in the celebrity interview analysed -218Hz versus 293HZ-. The overall lower pitch used in her Political Speech relative to her Celebrity Interview can be observed in Figure 1 below. The pitch shift indicates that the congresswoman wants to present herself as confident and trustworthy, delivering the speech in a lower pitch than what she would use in a less-controlled environment. In Tigue et al. (2012), a study was carried out to test the influence of Voice pitch on voting-related perceptions. Even though the study was performed on only male politicians, the results show that "lower-pitched voices were associated with favourable personality traits more often than were higher-pitched Voice", ergo, preferred. The outcome of this study holds in AOC's speech; therefore, we can conclude that pace

and pitch are devices speakers change to convey different impressions or attitudes in different contexts.

Moreover, not only is the pace faster and the tone higher but the F0 varies a significant amount more in the CI context:

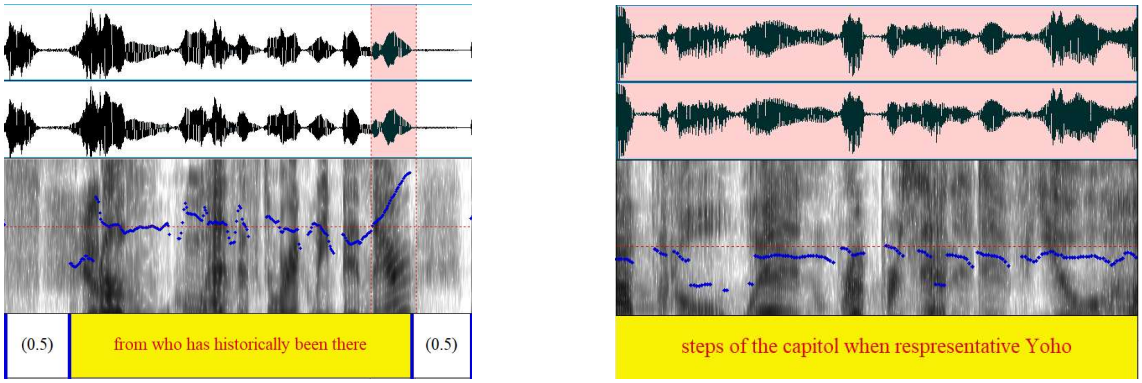

Figure 1. The left panel shows F0 variation with HRT in the CI context, the right panel shows F0 variation in the PS context.

Figure 1 clearly illustrates the difference in F0 variation between the two contexts. While in CI we see many rises and falls, in PS there are mostly falls or Mid-Level nuclear tones with an almost constant fundamental frequency mean value.

I also measured the pitch range in both types of speech by calculating the difference between the highest and lowest pitch used. The results show that the pitch range is wider in the PS context (517Hz) than in CI (433Hz), suggesting a more emphatic, persuasive, and motivating speech. The F0 features analysed indicate that Political speech tends to be more constant, calmer and more serious although much more emphatic than normal speech as is determined in Strangert (2005). These qualities would translate to fewer words per minute, longer pauses and fewer pitch shifts, yet more drastic ones. “These findings shed light on how speech behaviour can be optimally shaped to the demands of the situation” Strangert (2005). Therefore, we have discovered three features

that CI has that PS lacks: fast pace, high F0 and high F0 variation. By contrast, we have disclosed what Political Speech favours.

3.3. *Elongated final -s and creaky voice*

The variables analysed so far have been shown to differ in both types of speech. However, my data illustrate that not all variables differ across contexts. I also found constants. During my first review of the WAV files, the feature of elongated final -s seemed to stand out but when taking a closer look, I discovered that the sound [s] when realised as /s/ in final position always gets lengthened to an average value of 220ms, whereas in initial and medial position /s/ never goes beyond 100ms which makes it a constant in our speaker's speech-. When it is realised as /z/ it has a duration similar to that in other syllable and word positions, although there are some instances of it being elongated to utterance finally.

As for the use of creaky voice, although the presence of creaky voice was more prominent in a PS context, the difference was not as significant for it to mean that creaky voice varies across contexts.

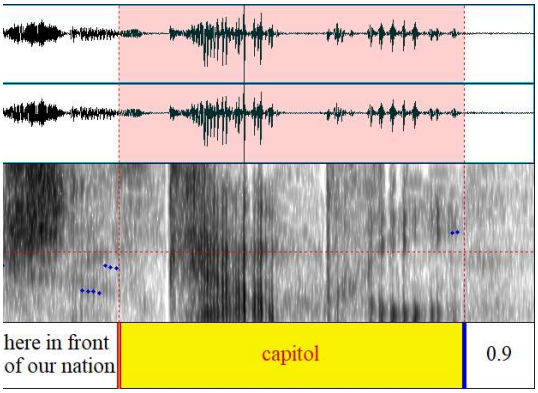

Figure 2. Creaky voice.

Figure 2 illustrates AOC's creaky voice and shows the challenge of analysing it as sometimes *Praat* does not even identify the glottal pulses in the waveform. The

spectrogram is not as clear, making it difficult to analyse vowel quality –through F1, F2, F3 formants– should anyone want to do so. Fortunately, in this study, the aim was just to identify creaky voice as constant –same feature use in both contexts– or as variable –different feature use in both contexts–.

3.4. *Use of High Rising Terminal*

Finally, HRT was not a feature I was expecting to find in the congresswoman speech. Recall that HTR has pragmatic meanings associated with the stereotypical "Valley talk" – that is, uptalk, or upspeak– and gives a sense of unfinished speech, insecurity, etc. AOC has built a reputation around hard, honest, and serious political speeches and many people have only seen this side of her. Although her political speeches show no instances of HRT, there are 7 instances of HRT in her CI speech with a mean rising in pitch of 49%.

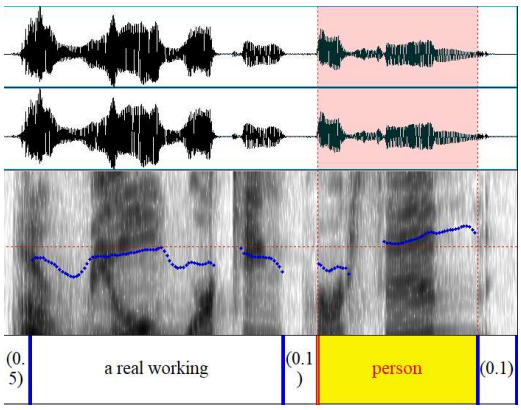

Figure 3. HRT in *person* from 184Hz to 329Hz.

Figure 3 shows one instance of HRT in AOC's CI context speech where F0 rises 145Hz in only one word. Most people have perceived HRT over the years as indicating insecurity and uncertainty, for example, Braga et al. (2004) state that "rising pitch means uncertainty while falling pitch conveys certainty and assertiveness". However, the study by McConnell-Ginet & Eckert in 2003 stated that HRT might have other meanings, such

as that the person is not giving the final word on the matter, that they are open to the topic continuing, or even that they are not yet ready to cede their turn. We could most likely conclude that statements made in the CI context by AOC do not aim for the same type of impact as they do in the PS context.

4. Conclusion

I started this project by looking at six variables: pauses, F0 mean and range, F0 variation, elongated final -s, creaky voice, and HRT. In the end, only four fluctuate between the context of the celebrity interview and political speech. The data suggests that the features elongated final -s and creaky voice are constants in congresswoman Alexandria Ocasio-Cortez' speech regardless of their context. I discovered that [s] was only elongated if it was in utterance final position and creaky voice appeared throughout all the data analysed even if it is the case that it becomes more prominent with lower F0. Mean pitch, pitch variation and range and HRT diverge in frequency and/or duration in both contexts and they all contribute to the hypothesis that the CI context gives rise to a fast-paced, higher-pitched speech with rich F0 variation whereas the PS context calls for a slow-paced, low F0, and stable tones without prominent F0 spikes and falls.

Each of these variables are prosodic devices used by the speaker to convey different meanings. If we consider the context of a celebrity interview as the one where AOC has less control over her own prosody, we could say that this speech is the one chosen unconsciously. As both speeches are inherently different in almost every aspect, we could assume that the prosodic variations in the political speech context are done consciously. This would confirm that AOC lengthens her pauses to deliver a speech more easily understood and likely to cause impact. She also lowers her pitch to sound more

trustworthy and confident. The use of Mid-Level nuclear tones and Fall tones makes her speech sound more definite and serious. Also, the use of these tones makes her pitch shifts even more emphatic in contrast.

In summary, this study has characterized some prosodic features of AOC's political speech and more informal speech, and thus contributes to our understanding of the prosodic features that are actively manipulated by speakers to convey specific meanings in different contexts.

References

Braga, D., & Marques, M. A. (2004). The pragmatics of prosodic features in the political debate. In *Speech Prosody 2004, International Conference*.

Holliday, N., Bishop, J., & Kuo, G. (2020). Prosody and political style: The case of Barack Obama and the L+ H* Pitch accent. In Proceedings of the 10th International Conference on Speech Prosody (pp. 670-674).

Late Night with Seth Meyers. (2019, March 22). Rep. Alexandria Ocasio-Cortez Responds to Fox News' Weird Obsession with Her [Video]. YouTube. https://www.youtube.com/watch?v=IjLoHgvBhCI&t=1s

McConnell-Ginet, S., & Eckert, P. (2003). *Language and gender*. Cambridge University Press.

Miyauchi, S. (2017). A Study on the Intonation Used in a Speech by President Obama.

NBC News. (2020, July 23). AOC Slams Rep. Yoho In House Floor Speech, Says 'Having A Daughter Does Not Make A Man Decent' [Video]. YouTube. https://www.youtube.com/watch?v=tSZcdMCHn2o&t=185s

Reich, S. S. (1980). Significance of pauses for speech perception. *Journal of Psycholinguistic Research*, *9*(4), 379-389.

Strangert, E. (2005). Prosody in public speech: analyses of a news announcement and a political interview. In *Ninth European Conference on Speech Communication and Technology*.

Tigue, C. C., Borak, D. J., O'Connor, J. J., Schandl, C., & Feinberg, D. R. (2012). Voice pitch influences voting behavior. *Evolution and Human Behavior*, *33*(3), 210-216.

Vadimovich, P.M. (2016). The step-by-step methodology of analyzing nuclear tones and their functions (as exemplified by David Cameron's discourse of public speeches). Педагогическое образование в России, (4), 51-55.