How Language Proficiency Affects the Phonetics of Singing: A Case Study on Multi-Lingual Korean Pop Group *Stray Kids*

By Nicolette Warren
Student ID 14391295
Supervisor Prof. dr. P.P.G. Boersma
For Master's of Arts in General Linguistics
University of Amsterdam
Submitted 16 June 2023

How Language Proficiency Affects the Phonetics of Singing: A Case Study on Multi-Lingual Korean Pop Group *Stray Kids*

I. Introduction

For almost as long as humans have been a species on this Earth, music has been an important aspect of how we communicate with one another. Music allows people to communicate feelings and ideas in a way that normal speech cannot. Rhythm, melody, and harmony transcend language, and open our minds to a multitude of different thoughts and feelings. In the modern day and age, there are hundreds, if not thousands, of different instruments used to create music. Even computers have been utilized as musical instruments in the last several decades. However, the most rudimentary musical instrument we have access to is our own voices, through the act of singing. The practice of singing has been recorded in every culture, regardless of time, region, or values, since the beginning of traceable human history (Koopman, 1994). Regardless of language, beliefs, and environment, music permeates our daily lives as a species. Additionally, humans have had the biological ability to produce music with our vocal tracts for at least 500,000 years, and have been theorized to have used the ability for urgent communications alongside the more common instances of spiritual practices and leisure we see today (Prasad, 2021). Knowing that, music seems to be just as ingrained in our DNA as breathing or blinking.

In modern life, music is a huge part of entertainment and leisure for the average person, and provides some of the most lucrative and star-studded careers possible for those willing to pursue the craft professionally. Some of the biggest names in music can make millions of dollars per show, with acts like Taylor Swift averaging a revenue of $9 million USD, all because of the vast network of fans they garner, through the prevalence of music in the average person's life (Sager, 2023). In fact, the average American spends nearly 17% of their time listening to music, with estimations reaching 3 hours and 50 minutes of music play time per day (Ferjan, 2023). Taking into consideration that several hours a day must also be dedicated to sleep, spending nearly 4 hours listening to music indeed assumes a large portion of a person's life. Truly, music is an extremely present aspect of modern human society.

While the origins of music are largely recorded as being related to religion or spiritualism, today there are thousands of different genres that rely on different instruments and moods to send a message. In fact, with the proliferation of popular music culture today, there are thousands of individual genres and subgenres of music (Wikipedia, "Music genre", 2023). Each genre tends to have its own unique lyrical and melodic content, to help distinguish the genres from one another and to target different audiences. From country-western to reggae, jazz to screamo, there is a genre out there to suit any one person's needs or desires at any given moment. And with the advent of online streaming services, the consumer of today has access to an incredible wealth of music from different regions, and in different languages, than their own. In particular, a category of music referred to as "World Musics," which just refers to popular music of different countries and cultures, is starting to become extremely popular amongst listeners of today.

The genre of interest in this paper is one of these "World Musics": Korean Pop Music. Korean Pop Music (K-Pop) has seen a massive boom in popularity with the advent of social media and online streaming services, along with COVID-19 lockdowns increasing the amount of time people spend online. From *Gangnam Style* taking Youtube by storm in 2012, to BTS and BLACKPINK breaking into the charts of the mainstream Western Music Market in a way never seen before over the last 5 years, K-Pop has officially become a commonplace concept. In the

2023 online article titled *A Brief History of K-Pop* by Arnaud, K-Pop is defined as a genre of music that originates in South Korea, and revolves heavily around stage performances and dance beats. The article continues on to describe the overall history of K-Pop. It is stated that the genre as it is known today was born in the 1990's, starting with a group called Seo Taiji and The Boys. Their anti-establishment hip-hop tracks and well-rehearsed choreographies appealed to the youth of Korea at the time, and they ultimately took their home country by storm. From that point, hundreds, if not thousands, of groups and solo acts have been thrust into the public spotlight. In the early 2000's and early 2010's, groups like TVXQ and Girls' Generation helped K-Pop grow in popularity across all of Asia, holding massively successful concert tours and selling masses of digital and physical copies of their music. In the mid 2010's through now, K-Pop has begun to break into the mainstream of the West, with new groups setting insane Youtube, Spotify, and Billboard Chart records almost daily.

In terms of actual musical style, there is a large variety of what elements can be considered aspects of the K-Pop genre, even including the languages used for the lyrics, so many modern music professionals regard K-Pop more as an industry than a genre. The most common component is some sort of danceable beat, and these beats are blended with any genre under the sun, sometimes multiple, to try and create a hit. The lyrics are crafted largely in Korean, but most songs also include catchy English phrases, and sometimes other languages too, like Spanish or French. Each song has an immaculate choreography to match, and the stage performances include flashy outfits, impeccable makeup, and tons of confetti. The performance is not just about the song, but rather it is a full package deal; the focus lies on the overall entertainment experience.

Within the K-Pop industry, it is often that you see individuals who are part of the Korean Diaspora, often hailing from the United States or Australia, or some who are not even ethnically Korean at all (usually being Chinese, Japanese, or Thai), moving to South Korea to train to become popstars. This diversity of ethnicity and nationality amongst these performers leads to a variety of people with different language backgrounds coming together to create music.

As a fan of K-Pop and a linguist myself, this diversity of language background within the industry inspired the idea for this project. The goal of this project is to perform a case study on one of these groups that has a particularly diverse composition, and compare the phonetics of their singing to one another. In particular, I will be looking at the first and second formant values (F1 and F2) of singing data across a multitude of tokens for native and non-native speakers of the target languages, Korean and English, and will determine if there are any statistically significant differences between the groups through the data collected.

One reason why a project like this may be important to tackle is because of biases and discrimination seen in different music markets all over the world. It is not uncommon to see both consumers and professionals in the music industry in different parts of the world have biases against those they consider to be "others". Take the Grammy's for example; they are one of the most prestigious music award shows ever, and historically almost all winners for the main categories have been Caucasian and/or American. Additionally, there is plenty of evidence of discrimination and bias in the "consumer" or listener aspect of Western Music. Take the classical music industry, for example. In a study titled *Gender and Race Bias in the Judgement of Western Art Music Performance* conducted by Davidson and Edgar in 2012, 36 young musicians of various gender and ethnic backgrounds were played clips of piano performances by trained professionals and asked to score them. In these clips, some of the examples were audio only, and some had a video component so that the judges could subconsciously gather data on ethnicity and gender. There was no statistically significant difference in the scores for the audio-only clips or, surprisingly, based on race. However, there was an unexpected bias to score women higher than men. Though not what typically is thought of when discussing bias, this data shows that the perception of art is influenced by the perception of the artist.

This is not exclusive to Western culture though. There are examples of discrimination like this globally. In Korea, there is an ideological culture war between embracing globalization for the sake of economic success, and maintaining intense ethno-nationalist ideals for the sake of self preservation in the context of a history of occupation. For example, a study was done in 2012 by Sookyung Kim that looked at the portrayal of foreigners in Korean media from 1990 to 2009, whether this be in a discussion of foreigners coming into the country, foreigners already in the country, or foreigners staying in their home countries. Additionally, the data included a discussion of both average civilians and important figures such as politicians, celebrities, musicians, and others. In this study, it was found that the portrayals of foreigners in the media were never outright negative, which is in line with the idea of embracing globalization. However, the portrayals often victimized or objectified foreigners in some manner, which falls in line with the ideals of ethno-nationalism. This is important in the context of my project because this ideal of ethno-nationalism often manifests in foreign-born or raised, but ethnically-Korean individuals still being considered foreigners in Korea. This would mean that many musical artists in the K-Pop industry, including several of the test subjects in my case study, would be considered as foreigners in Korea despite being ethnically Korean. Regardless of whether their phonetic data that will be analyzed later is in line with that of a “true native”, they may still be seen as “other” within the music industry in Korean. This idea is extremely problematic, as it uses language in the context of music as a tool to divide people, as opposed to unite them. As such, the potential results of this project can further influence how discrimination in the music industry may be understood, and how to approach statistical analyses that show differences in the least divisive way possible.

Before any work can be done with the data like building a corpus or analyzing any tokens, it is absolutely vital to take a moment to look at the ways in which singing as a form of production differs from speaking. In spoken word, it is the job of the speaker to try to balance perceptibility and articulatory effort so that the acoustic properties of the target sound can be expressed correctly to the listener. However, when someone is producing singing, there are additional concerns for resonance, expression, and style. These steer the focus of singing away from perceptibility and articulatory effort, which in turn can create differences in the acoustic data collected (Bradley, 2018). This means that what may be considered typical formant values, vocal tract position, and mouth shape for any given vowel while speaking may not match what is collected in singing data. In turn, this shift in focus in favor of stylistic elements also causes a difference in the interpretation of vowels, both in perception from listeners, as well as a singer’s own thought process behind production.

In regard to perception, there are several keystone studies that have been conducted on the categorization of vowels from singing data. These studies have been conducted using both “untrained” ears, or what may be considered the average consumer of music, as well as professionals such as singing instructors and speech pathologists, who have more experience in this line of work compared to other test subjects. In particular, the work of Gregg and Scherer (2006) stands out because of the highly detailed graphics that illustrate the relationship between the target vowel and the common interpretations from the audience. This experiment used both male and female English-speaking singers in order to record the various tokens in both low and high intonations (pitches). The panel of listeners, and subsequently interpreters, for these recordings was composed of both speech pathologists and singing instructors. In Figure 1 below, it can be seen that across all types of intonation, a large majority of the vowels were commonly misinterpreted by listeners. In most cases, the percentage of misinterpreted tokens exceeded 20%. In the diagram, the direction an arrow points indicating the direction in which the mismatched vowel compares to target vowel (i.e. "ɪ → i" would indicate that [ɪ] was misinterpreted as [i]) and the value next to each arrow show the proportion of incorrect matches as a decimal. The thickness of an arrow also indicates if the data is from data for that vowel recorded at a lower pitch or a higher pitch comparatively. The decimal proportions shown in the diagram are an average taken across both male and female singers within the same pitch category. The vowels that are circled and do not have arrows are indicated as such because their mismatch percentage did not exceed 20%. Based on this diagram, it is clear that when singers focus on stylistic elements, it not only alters their production, but also alters a listeners perception, even those with years of training in the field.

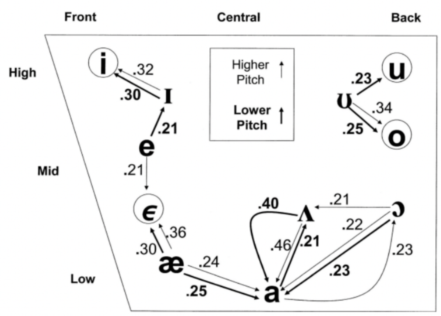

Figure 1, Diagram indicating vowel mismatch by listeners (Gregg and Scherer, 2006)

In regard to production, there are also several keystone studies that cover the differences seen in the vowel space while singing. One particularly comprehensive study regarding production of singing is the work of Thomas Cleveland titled *A Clearer View of Singing Voice Production: 25 Years of Progress*, published in 1994. In this work Cleveland compiles linguistic research from the 1970s to 1990s that discuss singing in English. In his paper, Cleveland spends a great deal of time discussing the finer details when it comes to producing singing. He explained that "the singer's formant helps to boost the singer's voice above the sound of the orchestra and is manifested by a clustering of the third, fourth, and fifth formants in the acoustic spectrum. The literature suggests that the singer's formant is typically derived from a lowering of the larynx and widening of the laryngeal tube. However, it remains to be seen what other articulatory modifications as well as phonatory adjustments promote the singer's formant," (pg. 19). Additionally, Cleveland states that the "observation of female singers singing higher pitches reveals that larger openings of the mouth are typically utilized for singing higher pitches... As the mouth opening is adjusted, the first formant is raised in concert with the fundamental frequency," (pg. 19). In short, Cleveland is explaining the exact ways in which people can alter their articulation using their various speech organs like the larynx and mouth in order to make their singing sound more pleasant, while still maintaining an amount of intelligibility necessary for interpretation from the audience. In addition, Cleveland briefly brings up the idea of a "singer's formant" used to aid projection, which is again an aid to the stylistic elements of signing. The details of the ways in which people utilize their speech organs differently for signing compared to speech, as well as acoustic differences like the “singer’s formant”, clearly show just how much singing varies from speech, and why the two need to be studied separately.

After establishing the basics of how singing may be both produced and perceived differently than speech, it is important to look at the work already done on the topic of phonetics in singing and L2 speech. There are a variety of different things that must be established before a project like this can take shape; the exact oddities of vowel formants and other spectral data in singing, the differences seen in vowels between native and L2 speakers, and several other niche studies that aid the understanding of these two ideas and other aforementioned material. This background will lay the foundation for the discussion of differences in vowel formants while singing between native and L2 speakers of a language that is to be discussed later.

The most foundational research for this project is that which distinguishes the vowels of singing from the vowels of speech. As briefly discussed earlier, Cleveland’s work shows that singers take liberties in articulation when producing vowels while singing. There are several subsequent keystone works on the topic. First, there is *Spectral Analysis of Sung Vowels*, a three-part series by Bloothooft and Plomp (1983-1985). In the first part, *Variation due to differences between vowels, singers, and modes of singing*, the pair conducted an experiment using 7 male and 7 female professionally trained singers, in which the singers were asked to sing a series of vowel sounds in several different “modes”, such as “light” and “dark”, which are common classifications used in the professional music world. The result of this experiment established that the factors “vowel”, “singer”, and “mode”, as well as their interactions, could have their spectral variations expressed as a function of the fundamental frequency of a musical note [F0]. In the second part of this series, *The effect of fundamental frequency on vowel spectra*, the pair studied how the aforementioned relationships between “vowel”, “singer”, and “mode” varied when the fundamental frequency varies, as a means to take their experimental data into the world of theory. The recordings used in the second study were the same ones from the previous study: however, they were newly represented in the ⅓-oct spectra method, to allow the researchers to focus more on the variation between each spectrum due to their fundamental frequency, and to project their findings onto data points that did not exist in their corpora. The results of this research confirmed that the relationship seen in experiments from the first study holds up in the world of theory. The third and final part of this series, *Characteristics of singers and modes of singing*, aimed to break down the data along the categories of “gender” and “mode” to allow the researchers to thoroughly describe the differences within each category. As a result, the researchers were able to break the category of “mode” into different dimensional properties, particularly the “soft-loud” dimension which shows the potential for an articulatory interpretation regarding glottal interference, and the “pressed-dark” dimension which shows a slight lowering or raising of [F1] depending on stylistic interpretation. Overall, this trio of studies establishes a relationship between different factors of singing and their effects on vowel production that are not typically seen in speech.

Another keystone project on this topic is *The Acoustics of the Singing Voice* by Johan Sundberg (1977). This project serves almost like a guide to producing a high-quality singing voice: Sundberg describes in depth the aspects of articulation, vowel placement, resonance, and other factors that create what is considered the “ideal” singing voice in the professional world. This project is so important because it helps to further the understanding of the resonance of vowels in singing versus speech. In particular, Sundberg discusses the idea of a “singing peak” formant that is not seen in speech. This formant was also discussed earlier with the work of Cleveland and seen as a means to aid auditory projection while someone is singing live. The example in the paper is that of a male singer, and this peak of energy in the sound occurs at around 2500 Hz. This is important to note for my project in particular because it may sometimes (though not always) get conflated with other formant values when using computer programs to extract formant values and depending on the clarity of the audio clip used. In particular, this could cause issues with the second formant, F2, as 2500 Hz is near to typical F2 values for several vowels in several languages. Thus, this peak in energy could cause difficulties with the interpretation of differences in vowel fronting between proficiency groups in each target language.

A final keystone project that studied the differences between singing and speaking is *The Harmonic Structure of Vowels in Singing in Relation to Pitch and Intensity* by Barret Stout (2005). Stout's project focuses on the effects of intensity when pitch is constant and the effects of pitch when intensity is constant for different vowel sounds. Though not identical, Stout's results seem to mirror those of the "singing formant" finding from the aforementioned Sundberg paper. In his experiment, Stout's analysis of the data shows a relationship in which an increase in intensity is found in any formant or "peak" above 1800 Hz for male singers holding a constant note. This again establishes a relationship we do not see in speech – vowels in singing have "atypical" energy peaks in their spectral data when compared to what is typically understood of speech. Such information is critical for the analysis of anything beyond F0 and F1, because these energy peaks may not all lie on the true formants of the data, and could cause interpretation issues when dealing with spectrograms.

Another extremely crucial precedent in the literature for this project is work that has been done on differences in the formant measurements of vowels between L1 and L2 speakers of the same language. Studies like this have been done on a multitude of different languages, with L2 speakers of all backgrounds imaginable. The projects that will be discussed below will focus specifically on speakers of Korean and English.

To start on a disappointing note, there unfortunately seems to be very little research published in English or in accessible places that studies the differences in measured formant values of Korean values in which the test subjects are native Korean speakers and L2 learners whose native language is English. This means that there is no study for comparison against any analyses I will conduct on the data in which Korean is the target language. Thus, there is no established precedent to show how English speakers will handle Korean vowels in speech, let alone in singing. This gap in the research pool unfortunately thins out the available information for the project to follow, but it is not significant enough such that the research cannot continue.

Despite this gap in the research on native versus L2 Korean speakers, there is a very robust repertoire on native versus L2 English speakers, as well as native bilingual Korean-English versus native monolingual Korean speakers. Up for discussion first will be native versus L2 English speakers.

One keystone project on native versus L2 English (native Korean) speakers is *Relative distances among English front vowels produced by Korean and American speakers* by Byunggon Yang (2014). As the title suggests, Yang's experiment was one that tested for the formant values of the front vowel contrasts [i-ɪ] and [ɛ-æ], neither of which exist in Korean, for both native speakers of English, and L2 speakers of English whose first language was Korean. Yang's findings showed that the L2 speakers on average had a significantly less distinguishable contrast between [i-ɪ] than native speakers. Any differences seen in the data for the [ɛ-æ] contrast were not statistically significant. Additionally, L2 speakers who were at a more "advanced" level as defined by a proficiency test given before the data collection had a significantly more similar formant pattern to native speakers than less "advanced" L2 learners. Yang's project is important for 2 reasons. First, the suggestion that advanced L2 learners can overcome deficits in vowel inventory based on their L1 to produce vowels phonetically very similar to native speakers of their L2 is important because the speakers to be studied in my project have varying proficiency levels in their L2. Consequently, those that have more advanced proficiency levels in their L2 are likely to have differing results than those that are less advanced, which may cause the data to have much more varied results than if the L2 speakers were all at a similar proficiency level. Additionally, the suggestion that, even if a contrast in an L2 does not exist in a participant's native language, their data may not be significantly different than native speakers of said L2, is extremely important because it will heavily affect the interpretation of similarities seen in the data I collect. That is, if the analysis does not prove to have statistically significant differences in formant measurements between proficiency groups, there could still be some sort of relationship between fluency and speech at play, but it is indeterminable. It may very well be a result of the phenomenon in which L2 learners can near-master sounds or contrasts not found in their native languages, but the data would neither confirm nor deny this fact.

Another keystone set of works revolve around English vowels produced by bilingual Korean-English speakers. This set of works belong to Eunjin Oh, and were published from 2009 to 2013. In the first work, *The High Vowel Systems of Korean-English Bilingual Speakers*, Oh documents the English vowel inventory, with a focus on high vowels, as produced by natively bilingual speakers. The findings of Oh's study actually differ slightly from the aforementioned work by Yang, and Oh finds that there are notable differences in the high vowel contrasts of bilingual Korean-English speakers when compared to what would be expected of monolingual English speakers. However, the context of Oh's study is important to remember, because she only made comparisons between natively bilingual and monolingual speakers, whereas Yang's study compared different levels of L2 proficiency. Additionally, Oh went on to write several other papers like *Dynamic spectral patterns of American English front monophthong vowels produced by Korean-English bilingual speakers and Korean late learners of English*, and *Learning Dynamic Transitions of the American English Vowels [eɪ] and [oʊ]: A Comparison between Korean-English Bilingual Speakers and Korean Learners of English*, in which she compared the data from natively bilingual speakers to data from late L2 learners. The results of these papers from Oh do fall more in line with those of the aforementioned Yang project, and suggest there is a significant difference between the vowel formant inventory of natively bilingual speakers versus the inventory of late L2 learners. In particular, Oh's findings regarding [eɪ] and [oʊ] are insightful because these vowels are considered to be "dynamic monophthongs" in American English (commonly notated as just [e] and [o]) and this shows that this aspect is seemingly a lot harder for later L2 learners to master when compared to natively bilingual speakers. Overall, the work of Oh brings insight to the way in which natively bilingual Korean-English speakers differ from both monolingual English speakers AND late L2 learners of English.

Lastly, there are a few additional works regarding speech versus singing that may need to be discussed before delving into the project at hand. They cover a variety of topics like the theory of "basis of articulation", speech synthesis, globalization of music, and others.

The first of these works is *Basis of Articulation and the Phonetics-Phonology Interface* by Božović and Kašić (2021). In this work, Božović and Kašić work to define and discuss everything relevant to the theory of "basis of articulation". As they explain, "basis of articulation" (sometimes referred to as "articulatory setting") has taken on two possible definitions in the field of linguistics. The first of these definitions is “a language specific ‘default’ or ‘baseline’ (neutral) position of the speech organs, prior to the moment of speaking, as they prepare to engage in the production of speech,” (pg. 2). The second is “the sum of all speech habits in a language, which constitute the native speakers’ typical pronunciation, i.e. all the language-specific typical (habitual) positions of the speech organs when put into action,” (pg. 2). The pair assert that these two definitions refer to very different explanations of the same phenomenon that has been anecdotally documented for centuries; people who speak different mother-tongues pronounce the same words differently. These definitions are then used as a framework for their specific research on Slavic languages, which leads to an important conclusion – if viewed as a “system of speech habits” the theory of “basis of articulation” can encompass both the physiological aspects from the first definition along with the psychological aspects of the second definition. This theory of “basis of articulation” is extremely important to my project because it is the framework for my own statistical analysis; I will look for differences in this “basis of articulation”, averaging my results across speaker and vowel to obtain results that approach the “bigger picture” as opposed to smaller details.

Another of these works is *Voice conversion: From spoken vowels to singing vowels* by New et. al (2010). In this project, the goal was to attempt to create a synthesis program that could convert spoken vowels into sung vowels, using the parameters for singing as discussed above. This includes modifications to F0, duration, and a variety of other spectral properties, similarly discussed in several of the aforementioned projects. Two systems are proposed for making these changes: one involves scaled variances to the spoken data, and the other is a linear transformation. After a number of trials, this study concluded that the proposed system has the potential to be quite successful, given that the singing data they generated looked and sounded very similar to raw singing data collected from the participants who also recorded the spoken data. This is relevant to my project because vocal effects like reverb and autotune in the recording studio can cause changes to the spectral data collected from recordings in this context. However, the research suggests that synthetically generated or altered singing may be “close enough” to raw singing data that we can regard these differences as negligible. This will be extremely important for the interpretation of data in my project going forward.

Another work worth mentioning is *Whose World, What Beat: The Transnational Music Industry, Identity, and Cultural Imperialism* by Reebee Garofalo (1993). Garafalo’s essay discusses at length the topic of globalized popular music and its relationship to traditional music styles. In particular, Garafalo mentions English, particularly the American variation, as the Lingua Franca of pop music, and the birth of new genres from fusion of traditional and modern styles. While some people tend to frame the globalization of the music industry as the “death” of traditional styles and diversity in music, Garofalo frames it as the birth of a new diversity: one based around interconnectedness instead of autonomy. These ideas are important to the project at hand because it helps explain a variety of things within the data used. First, it helps explain why there is a mix of Korean and English in a K-Pop group at all. Despite being “Korean” Pop, the K-Pop industry is extremely globalized, with heavy marketing done towards audiences outside of Korea. Therefore, it makes sense to not only release full English versions of certain songs, but to also have English lyrics sprinkled into standard Korean releases. Second, it explains why the vowels in the English data may be analyzed more homogeneously than anticipated, despite the English speakers having a very diverse variety of language backgrounds. American English is seen as a “standard” for globalized pop music, and it is not uncommon to see singers take on a more Americanized accent when singing in English (regardless of their typical English speaking voice). This is not a hard and fast rule, but it happens a great deal of the time.

A final work worth mentioning is *Embracing Anti-Racist Practices in the Music Perception and Cognition Community* by Baker et. al (2020). This is an editorial work that discusses the history of musical theory and that contests the fact that the rules laid forth in that framework are largely based in antiquated racist and imperialist ideals, but for some reason are still used fervently in the modern day and age. The paper goes into detail to explain how not only music theory has its roots in bigotry, but that the more modern academic field of music perception and cognition has issues regarding lack of diversity in authorship and collaboration, as well as heavy gatekeeping based around subconscious biases. Additionally, this idea is applied to the music industry as a whole, in which people of marginalized backgrounds being the head creatives behind popular music is considered outside of the norm. This is important to the project at hand because it helps gives context to the reader as to why the test subject for the case study is interesting or "revolutionary" – a handful of K-Pop groups are a largely self-produced, which means all of their music from instrumentals to lyrics are written by the members themselves. Thus, their musical work falls outside of this prescribed norm, especially in regards to the mixing of languages in the lyrics.

Now that the background of what has already been studied has been covered, it is important to establish who the case study will revolve around. The K-Pop group chosen as the test subject for this project is Stray Kids. Stray Kids are an eight (formerly nine) member boy group signed under JYP Entertainment. Releasing a pre-debut extended play (EP) on 8 January 2018, and making their official debut with another EP on 25 March 2018, Stray Kids has since released three studio albums, three compilation albums, one reissue, eleven EPs, two single albums, and twenty-eight singles (kpopprofiles.com, "Stray Kids Discography", 2023). These releases have been in Korean, English, and Japanese, with some songs being released in multiple languages. Their roster is composed of 2 Australians and 6 Koreans, all of whom have varied experiences with the Korean and English languages (kpopprofiles.com, "Stray Kids Members", 2023). This project will focus only on Korean and English releases, since no one in their line up is a native speaker of Japanese, and thus comparisons of the data would only be between L2 speakers.

In age-order, the line up consists of Bang Chan (born 1997), who was born in Seoul, Korea but has citizenship in Australia and was raised in Sydney, Australia. He is a native speaker of English, and has studied Korean since a very young age while speaking it sparingly at home when he was growing up. Then there is Lee Minho, stage name Lee Know (born 1998), who was born and raised in Gimpo, Korea. He is a native speaker of Korean, and began learning English in the group's management agency's compulsory program at the age of 18. Next is Seo Changbin (born 1999), who was born and raised in Yongin, Korea. He is a native speaker of Korean, and began learning English in the management agency's compulsory program at the age of 16. Following that is Hwang Hyunjin (born 2000), who was born in Seoul, Korea. He is a native speaker of Korean, but lived in Las Vegas, USA from the ages of 5 to 6 where he attended school and learned English, and subsequently returned to living in Seoul. His English education was continued in the management agency's compulsory program. Next is Han Jisung, stage name HAN (born 2000), who was born in Incheon, Korea. He is a native speaker of Korean, but lived in Malaysia for nearly a decade where he attended private Catholic school and learned English from native speakers of a British dialect. His English education was also continued in the management agency's compulsory program. After this is Felix Lee (born 2000), who is the only foreign born member, having been born and raised in Sydney, Australia. He is a native speaker of English, and began learning Korean in the management agency's compulsory program for foreign artists at the age of 17. Following this is Kim Seungmin (born 2000), who was born in Seoul, Korea. He is a native speaker of Korean, but lived in Los Angeles, USA from the ages of 10 to 11 where he attended school and learned English, and subsequently returned to living in Seoul. His English education was also continued in the management agency's compulsory program. Lastly, there is Yang Jeongin, stage name I.N (born 2001), who was born and raised in Busan, Korea. He is a native speaker of Korean, and began learning English in the management agency's compulsory program at the age of 15. In Figure 2 below, a photograph of the members can be seen.

Figure 2, Stray Kids members. Names from left to right: Yang Jeongin, Lee Minho, Han Jisung, Felix Lee, Bang Chan, Hwang Hyunjin, Kim Seungmin, Seo Changbin (JYP Entertainment, 2023)

At this point, it is important to establish a brief history of Korean and English, including the 3 dialect variants of English spoken by the members, and what the typical vowel formants of male speakers of these languages look like before collecting usable data to analyze.

To start, Korean is a part of the Koreanic Language Family, which also includes other languages from the Korean Peninsula, Manchuria, and Jeju Island. It is also the most widely spoken language in said family, with the Koreanic Language Family being the 13th largest in the world by number of speakers, having over 77 million in and around the region of origin. While this language family is well documented in modern history, little is known about its origins and evolutions over time due to a lack of writing system in any of the major languages of the family until the 15th century, as well as uncertain relatedness to other language families in the region (Lee, 2015). In Figure 3, there is an image detailing the spread and usage of the Koreanic Language Family via a dialect map of Korean (Morgan, 2019).

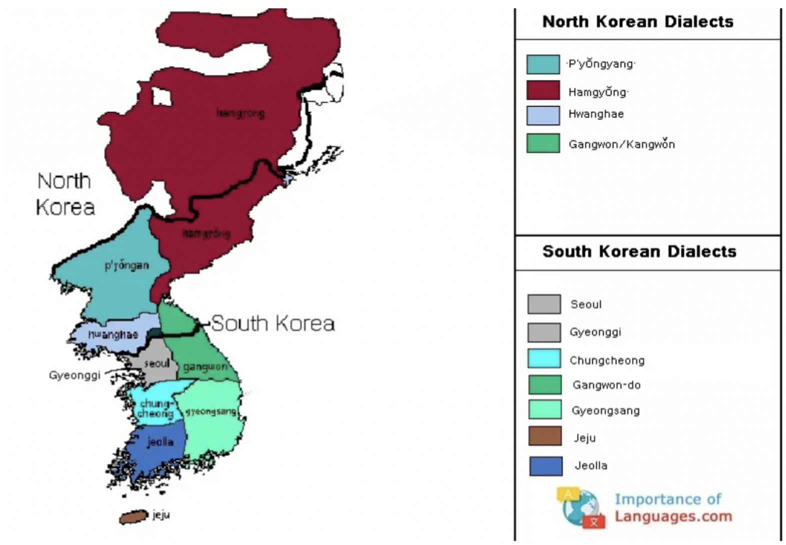

Figure 3, Map of Korean Dialect Usage (Morgan, 2019)

To properly analyze vowel data in Korean, it is important to establish the vowel inventory. In Korean, there are seven standard monophthong vowels. In the International Phonetic Alphabet (IPA), these would be notated as [i], [ɯ], [u], [o], [e], [ʌ], and [a]. Additionally, Korean features 12 diphthong vowels. In IPA, these would be notated as [je], [jɛ], [ja], [jo], [ju], [jʌ], [ɥi], [we], [wɛ], [wa], [wʌ], and [ɰi]. It is important to note that [ø] was once a part of the language's inventory but has since gone effectively extinct, and [ɛ] only occurs in diphthongs and not as a monophthong. Below in Figure 4, you can see the full list of monophthong and diphthong vowels condensed into one table (wikipedia.com, "Korean Phonology", 2023). Additionally, in Figure 5, you can see average formant values of adult male speakers of Korean, including vowels not native to their language (Yang, 1996).

Monophthongs	[i]	[ɯ]	[u]	[o]	[e]	[ʌ]	[a]	
Diphthongs (sorted by final vowel sound)	[ɥi] [ɰi]		[ju]	[jo]	[je] [we]	[jʌ] [wʌ]	[ja] [wa]	[jɛ] [wɛ]

Figure 4, vowels in Korean (Wikipedia, 2023)

Vowel	f_0		F_1		F_2		F_3	
a	162	(25)	738	(87)	1372	(124)	2573	(127)
ɛ	165	(22)	591	(75)	1849	(106)	2597	(110)
e	167	(26)	490	(105)	1968	(150)	2644	(94)
i	172	(24)	341	(29)	2219	(176)	3047	(146)
o	170	(25)	453	(47)	945	(134)	2674	(156)
ø	166	(24)	459	(69)	1817	(163)	2468	(134)
u	174	(27)	369	(43)	981	(141)	2565	(173)
y	174	(26)	338	(30)	2114	(140)	2729	(213)
ʌ	165	(25)	608	(76)	1121	(110)	2683	(145)
ɨ	174	(26)	405	(37)	1488	(176)	2497	(80)

Figure 5, average vowel formants of Korean speakers for a selection of vowels (Yang, 1996)

Next, English is part of the Germanic branch of the Indo-European Language Family, and is one of the most widely spoken languages in the world, with over 2 billion speakers currently. It has been spread as a Lingua Franca throughout the planet through colonization and imperialism on the part of England in the past, and the United States in more modern times. English is often the first foreign language of choice in countries all over the world because of this history (Crystal et. al, 2023). There are many dialectal variations of English based on place of usage, and three are important to this project; British English, American English, and Australian English. Below in Figure 6, a map showing the percentage of population per country that self-reports as English Speakers is given (WikiMedia Commons, 2023).

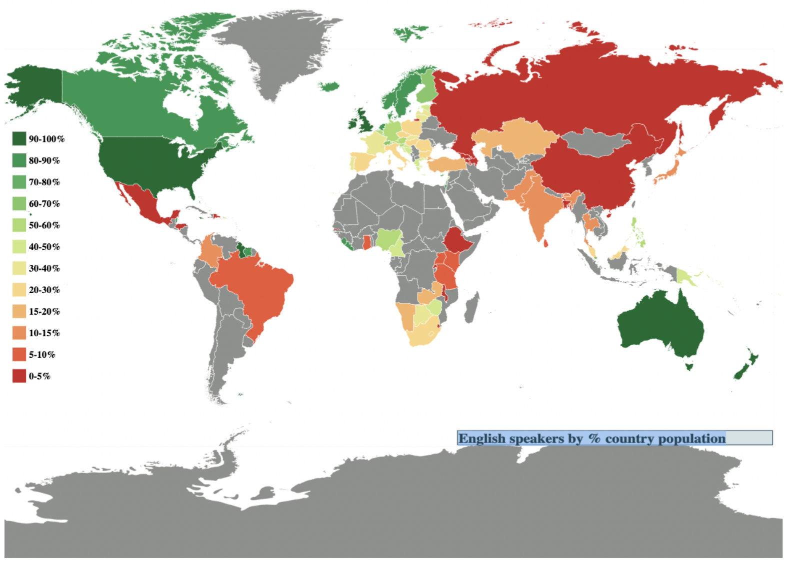

Figure 6, English speakers by % country population (WikiMedia Commons, 2023)

Though there is no truly "standard" dialect in any of the aforementioned regions, there are common traits in the speech of inhabitants in these areas that create distinctions between dialects of different places, and one of these traits is typical vowel space and vowel inventory. To try and analyze the phonetic content of Stray Kids' music, all three of these major English vowel inventories may need to be referenced due to their diverse language experiences.

British English contains an inventory of thirteen standard monophthong vowels. In IPA, these would be notated as [ɪ], [iː], [ʊ], [uː], [ɔː], [e], [ɛː], [ə], [ɜː], [ɒ], [æ], [ʌ], and [ɑː]. There are also seven standard diphthongs . In IPA, these would be notated as [eɪ], [aɪ], [ɔɪ], [aʊ], [əʊ], [ɪə], and [ʊə]. Interestingly, there are also five standard triphthongs. In IPA, these would be notated as [eɪə], [aɪə], [ɔɪə], [aʊə], and [əʊə].

Meanwhile, American English contains an inventory of thirteen standard monophthongs as well. In IPA, these would be notated as [ɪ], [i], [ʊ], [u], [ɛ], [eɪ], [ə], [ɜ], [oʊ], [æ], [ʌ], [ɑ], and [ɔ]. Two of these, [eɪ] and [oʊ], appear to be notated as diphthongs, but are transcribed and analyzed as monophthongs in other literature, often appearing as [e] and [o] respectively, as they will in the remainder of this paper. There are additionally three true diphthongs in this dialect. In IPA, these would be notated as [aɪ], [ɔɪ], and [aʊ].

Lastly, in Australian English, the inventory of vowels contains fourteen standard monophthongs. In IPA, these would be notated as [ɪ], [iː], [ʉː], [ʊ], [oː], [e], [eː], [ə], [ɜː], [ɔ], [æ], [æː], [a], and [aː]. Additionally, there are seven standard diphthongs. In IPA, these would be notated as [æɪ], [ɑɪ], [oɪ], [æɔ], [əʉ], [ɪə], and [ʊə].

Below in Figures 7, 8, and 9, you can see the full list of vowels by dialect (Wikipedia, “English Phonology”, 2023). Additionally, in Figures 10, 11, and 12, you can see the average formant values of male speakers of each dialectal variation, with individual citations.

Received Pronunciation[44][45]

	Front		Central		Back	
	short	long	short	long	short	long
Close	ɪ	iː	ʊ	uː[a]		ɔː[a]
Mid	e	ɛː	ə	ɜː	ɒ[a]	
Open	æ		ʌ			ɑː
Diphthongs	eɪ aɪ ɔɪ aʊ əʊ ɪə ʊə					
Triphthongs	(eɪə aɪə ɔɪə aʊə əʊə)					

Figure 7, vowels in British English (Wikipedia, 2023)

General American

	Front		Central		Back	
	lax	tense	lax	tense	lax	tense
Close	ɪ	i			ʊ	u
Mid	ɛ	eɪ[b]	ə	(ɜ)[c]		oʊ[b]
Open	æ		(ʌ)[c]	ɑ		(ɔ)[d]
Diphthongs	aɪ ɔɪ aʊ					

Figure 8, vowels in American English (Wikipedia, 2023)

General Australian

	Front		Central		Back	
	short	long	short	long	short	long
Close	ɪ	iː		ʉː[a]	ʊ	oː[a]
Mid	e	eː	ə	ɜː	ɔ[a]	
Open	æ	(æː)[e]	a	aː		
Diphthongs	æɪ ɑɪ oɪ æɔ əʉ ɪə (ʊə)[f]					

Figure 9, vowels in Australian English (Wikipedia, 2023)

Vowel	F1(Hz)	F2(Hz)	F3(Hz)
iː	280	2620	3380
ɪ	360	2220	2960
e	600	2060	2840
æ	800	1760	2500
ʌ	760	1320	2500
ɑː	740	1180	2640
ɒ	560	920	2560
ɔː	480	760	2620
ʊ	380	940	2300
uː	320	920	2200
ɜː	560	1480	2520

Figure 10, average vowel formants of male British English speakers for a selection of vowels (UCL Division of Psychology and Language Sciences, 2020)

Vowel	f_0		F_1		F_2		F_3	
æ	126	(16)	687	(83)	1743	(113)	2497	(137)
a	125	(15)	638	(46)	1051	(74)	2318	(185)
ɔ	128	(20)	663	(62)	1026	(57)	2527	(171)
e	128	(18)	469	(36)	2082	(130)	2636	(168)
ɛ	132	(24)	531	(52)	1900	(84)	2561	(148)
i	136	(21)	286	(32)	2317	(104)	3033	(191)
ɚ	130	(20)	490	(32)	1363	(99)	1787	(165)
ɪ	130	(15)	409	(32)	2012	(110)	2671	(148)
ɑ	127	(15)	694	(89)	1121	(85)	2548	(136)
o	129	(18)	498	(41)	1127	(93)	2375	(131)
ʊ	135	(21)	446	(46)	1331	(102)	2380	(125)
ʌ	127	(15q	592	(45)	1331	(71)	2494	(167)
u	135	(17)	333	(33)	1393	(213)	2282	(114)

Figure 11, average vowel formants of male American English speakers for a selection of vowels (Yang, 1996)

	Men		
Vowel in	F1	F2	F3
beat	270	2300	3000
bit	400	2000	2550
bet	530	1850	2500
bat	660	1700	2400
part	730	1100	2450
pot	570	850	2400
boot	440	1000	2250
book	300	850	2250
but	640	1200	2400
pert	490	1350	1700

Figure 12, average vowel formants of male Australian English speakers for a selection of vowels (synSinger, 2016)

II. Methods and Results

To reiterate, I will be conducting a case study on the potential differences in recorded F1 and F2 values while singing vowels in Korean and English between the members of the K-Pop group Stray Kids. As mentioned in the introduction, the 8 members have diverse backgrounds in both languages due to their upbringings. Thus, they are the perfect candidates for a case study on how language background may affect the vowel space while singing in both of the target languages.

To conduct an analysis on the formant values while singing for the 8 different members, a corpus has been constructed using a cappella versions of songs from all of the group's Korean and English releases. At least one song from every major release was used, from their EP *Mixtape*, released 8 January 2018, to their compilation album *SKZ Replay*, released 21 December 2022 (kpopprofiles.com, "Stray Kids Discography", 2023). Note: at the time the data was collected and analyzed, their latest studio album, *5-STAR*, had yet to be released. At the time of submission, said album has been released to the public for approximately two weeks (released June 2, 2023). The song or songs selected from each of these major works were chosen based on two criteria. The first criteria was the frequency of shared lines between members. To make the cut, a song had to include either English or Korean parts where at least two members with different language backgrounds sang the same lyrics and melody, but at different times, not simultaneously. For example, if a song had Felix, who is fluent in English and an L2 speaker of Korean, sing a specific Korean line in the first chorus, and then Hyunjin, who is fluent in Korean and an L2 speaker of English, sing the same line in the second chorus, it would make the cut. On the other hand, if a song had Chan, who is fluent in English and an L2 speaker of Korean, and Seungmin, who is fluent in Korean and an L2 speaker of English, sing the same English line during the first chorus simultaneously, it would not make the cut. The second criteria utilized was that any songs that were originally released before their 9th member left in 2019 had to have officially released versions with only 8 members to be considered. This was to ensure that no data from the ex-member would be collected, as his contribution to the data pool would ultimately be very small due to the timing of his departure, and therefore may constitute an "outlier" in the data. In total, using these criteria, there were 23 songs selected. The full list of songs and lyrical content used in this project will be available in Appendix 1, for reference purposes. Across these 23 songs, by member, there were 871 tokens for Bang Chan, 296 tokens for Lee Know, 108 tokens for Changbin, 207 tokens for Hyunjin, 186 tokens for HAN, 132 tokens for Felix, 648 tokens for Seungmin, and 370 tokens for I.N. By language, there were 1822 tokens for Korean and 996 tokens for English. Thus, a total of 2818 tokens were selected overall.

Unfortunately, the raw studio recordings of just their vocals are not available online due to copyright and intellectual property laws, so the a capella files had to be manually created. The base WAV files were generated using officially released audios sourced from the Stray Kids official Youtube (youtube.com, all original links in “Works Cited”, 2023), which were then run through the MotionBox file converter (motionbox.io, “Youtube to WAV”, 2023). From here, these base WAV files were run through a Vocal Remover and Isolation AI (Melnik, 2023) to isolate the vocals from the instrumentals. These secondary WAV files were then saved to be used as part of the corpus. The one caveat to this method was that, because there are limitations to the power of AI like this, the a capella files created all have some level of interference from instrumentals and vocal effects unlike raw audio from the studio would, which may affect the measured formant values.

To be able to use the manually generated a capella files as a corpus, the data needed to be heavily annotated. First, all tokens unrelated to the project needed to be discarded or ignored. This was done using a Praat text grid (Boersma and Weenink, 2023). Each of the relevant sections of music were isolated and then annotated by member and lyrical content. Additionally, each larger section of music was further annotated by syllable to isolate the vowel sounds. Then, the Praat function “Get formants” was used to find the first and second formants of each vowel within each syllable. Because of the nature of vocal technique, diphthongs were analyzed as the longer of the two vowel sounds, to allow the data to be more easily used when creating a graph of the contours of the vowel space. For example, the English word “I” is normally transcribed as a diphthong [aɪ], but to hold a musical note for a given duration of time, the [a] sound would be extended. This [a] sound is what would be analyzed in the transcription. After all of this was completed, all 2818 tokens and their formant values were broken down into two tables, one for each language, labeled by, speaker, proficiency, and vowel. The tables were also sorted alphabetically by vowel recorded for each token.

Within each language there were two proficiency categories that the members were split into based on the demographic information detailed in the introduction above. The groupings are visible in Figure 13. Additionally, the list of recorded vowels is visible in Figure 14. In total, 7 vowels were recorded for Korean, and 14 vowels were recorded for English.

SINGER	ENGLISH PROFICIENCY	KOREAN PROFICIENCY
Bang Chan	native	learner
Lee Know	learner	native
Changbin	learner	native
Hyunjin	learner	native
HAN	learner	native
Felix	native	learner
Seungmin	learner	native
I.N	learner	native

Figure 13, table showing the groupings of individuals by proficiency level

LANGUAGE	VOWEL RECORDED
English	a
English	ae
English	ɑ
English	e
English	ə
English	ɛ
English	ɜ
English	i
English	ɪ
English	o
English	ɔ
English	u
English	ʊ
English	ʌ
Korean	a
Korean	e
Korean	i
Korean	o
Korean	u
Korean	ʌ
Korean	ɯ

Figure 14, vowels recorded in analysis for both English and Korean

To determine any potential differences between proficiency groupings, statistical analysis needed to be conducted on the resulting formant values. Using RStudio (RStudio Team, 2020), the data from the aforementioned tab-separated files was run through a mixed-model formula to determine whether or not the difference in values is significant. This was done in two parts, once for English and once for Korean. Within each language, two tests were performed. One test mapped F1 as a function of proficiency, and one test mapped F2 as a function of proficiency. Additionally, each test considered the category of vowel as a random variable, with random slopes for proficiency, as well as considering the speaker a random variable, with random intercepts. This was done so that the results can help determine a generalized result for articulation overall, as opposed to results for individual vowels, or conflating all vowels or all speakers into one single variable. By doing so, this establishes any potential differences in "basis of articulation", as referenced in the introduction, which would give an overall picture of physiological and psychological differences in the articulatory systems of the speakers from different language backgrounds. The specific functions used can be seen in Figure 15, and the analyses of the data can be seen below in Figures 16 - 19.

````
```{r}
model <- lmerTest::lmer (F1..HZ. ~ PROFICIENCY + (PROFICIENCY | VOWEL) + (1 | SPEAKER), data=table)
summary (model)
confint (model)
```

```{r}
model2 <- lmerTest::lmer (F2..HZ. ~ PROFICIENCY + (PROFICIENCY | VOWEL) + (1 | SPEAKER), data=table)
summary (model2)
confint (model2)
```
````

Figure 15, functions used to perform the analyses

```
## Random effects:
##  Groups   Name            Variance Std.Dev. Corr
##  SPEAKER  (Intercept)      1551.4   39.39
##  VOWEL    (Intercept)      6417.5   80.11
##           PROFICIENCY-N+L   155.6   12.47   -1.00
##  Residual                 26076.8  161.48
## Number of obs: 1822, groups:  SPEAKER, 8; VOWEL, 7
##
## Fixed effects:
##                 Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)      686.490     34.791   9.073  19.732  9.2e-09 ***
## PROFICIENCY-N+L  -67.008     34.442   5.533  -1.945    0.104
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr)
## PROFICIENCY 0.122
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')

##                       2.5 %     97.5 %
## .sig01            18.51232  69.789366
## .sig02            47.27457 141.210316
## .sig03            -1.00000   1.000000
## .sig04             0.00000  33.645606
## .sigma           156.36517 166.901993
## (Intercept)      616.23658 756.469705
## PROFICIENCY-N+L -137.99246   3.482973
```

Figure 16, Analysis of Korean data, part 1

```
## Random effects:
##  Groups   Name            Variance Std.Dev. Corr
##  SPEAKER  (Intercept)      1289.5   35.909
##  VOWEL    (Intercept)     67590.8  259.982
##           PROFICIENCY-N+L    15.3    3.911  1.00
##  Residual                 46086.1  214.677
## Number of obs: 1822, groups:  SPEAKER, 8; VOWEL, 7
##
## Fixed effects:
##                 Estimate Std. Error       df t value Pr(>|t|)
## (Intercept)     1643.922     99.647    6.293  16.498 2.04e-06 ***
## PROFICIENCY-N+L  -86.576     32.806    5.159  -2.639   0.0446 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr)
## PROFICIENCY 0.125
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')

##                      2.5 %     97.5 %
## .sig01                  NA         NA
## .sig02                  NA         NA
## .sig03                  NA         NA
## .sig04                  NA         NA
## .sigma                  NA         NA
## (Intercept)      1436.9122 1852.71905
## PROFICIENCY-N+L -157.9292  -19.91613
```

Figure 17, Analysis of Korean data, part 2

```
## Random effects:
##  Groups   Name             Variance  Std.Dev. Corr
##  VOWEL    (Intercept)      3.119e+03  55.8469
##           PROFICIENCY-N+L 2.138e-01   0.4624 1.00
##  SPEAKER  (Intercept)      3.155e+03  56.1676
##  Residual                  2.326e+04 152.4978
## Number of obs: 996, groups:  VOWEL, 14; SPEAKER, 8
##
## Fixed effects:
##                 Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)      696.832     28.397  11.098  24.539 5.07e-11 ***
## PROFICIENCY-N+L   93.035     47.648   5.943   1.953   0.0992 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr)
## PROFICIENCY -0.414
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```
##                       2.5 %    97.5 %
## .sig01            37.378850  86.77189
## .sig02            -1.000000   1.00000
## .sig03             0.000000  18.32292
## .sig04            29.844306  92.25755
## .sigma           145.975460 159.51271
## (Intercept)      641.837301 751.37091
## PROFICIENCY-N+L  -1.010007 187.77556
```

Figure 18, Analysis of English data, part 1

```
## Random effects:
##  Groups   Name             Variance Std.Dev. Corr
##  VOWEL    (Intercept)      39189    197.96
##           PROFICIENCY-N+L  2142      46.28   0.94
##  SPEAKER  (Intercept)       7987     89.37
##  Residual                  40099    200.25
## Number of obs: 996, groups:  VOWEL, 14; SPEAKER, 8
##
## Fixed effects:
##                 Estimate Std. Error       df t value Pr(>|t|)
## (Intercept)     1604.505     65.119   17.884  24.640 2.98e-15 ***
## PROFICIENCY-N+L   94.088     76.041    5.762   1.237    0.264
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr)
## PROFICIENCY -0.155
```

```
##                        2.5 %     97.5 %
## .sig01           135.9020581  295.99205
## .sig02             0.1370031    1.00000
## .sig03            10.1588140   96.65805
## .sig04            46.9900347  153.58876
## .sigma           191.6539790  209.54824
## (Intercept)     1477.0241674 1731.58049
## PROFICIENCY-N+L  -61.2144191  249.11794
```

Figure 19, Analysis of English data, part 2

The mixed-model function “lmer” from R was used on both the Korean and English data collected to fit the data into linear models, so that F1 and F2 could be viewed as a function of proficiency, with random variable vowel, to indicate the potential relationship between fluency and formant values while singing. The two levels of this predictor (native speakers and learners) were coded into orthogonal contrasts, with native speakers (N) being negative and learners (L) being positive. In English, the results for the test on F1 showed that measurements from learners were not significantly different than those of native speakers (t = 1.953, p = 0.0992). It cannot be concluded with any amount of confidence that fluency in English has any relationship to the formant values collected (estimated difference = 93.035 Hz, standard error = 47.648 Hz, 95% confidence interval = -1.010007… 187.77556 Hz). Also in English, the results for the test on F2 showed that measurements from learners were not significantly different than those of native speakers (t = 1.237, p = 0.264). It cannot be concluded with any amount of confidence that fluency in English has any relationship to the formant values collected (estimated difference = 94.088 Hz, standard error = 76.041 Hz, 95% confidence interval = -61.2144191… 249.11794 Hz). In Korean, the results for the test on F1 showed that measurements from learners were not significantly different than those of native speakers (t = -1.945, p = 0.104). It cannot be concluded with any amount of confidence that fluency in Korean has any relationship with the formant values collected (estimated difference = -67.008 Hz, standard error = 34.442 Hz, 95% confidence interval = -137.99246… 3.482973 Hz). Also in Korean, the results for the test on F2 showed that measurements from learners were significantly lower than those of native speakers (t = -2.639, p = 0.0446). It can be concluded with a small margin of confidence that fluency in Korean affected the formant values (estimated difference = -86.57586 Hz, standard error = 32.80630 Hz, 95% confidence interval = -157.9292… -19.91613 Hz). This means that in only one of four instances, the results showed a statistically significant difference.

III. Conclusion and Discussion

Based on the results of the statistical analyses above, there were no consistent statistically significant differences in the vowel formant data collected between native Korean and native English speakers while singing in either of the two languages. There is evidence to suggest that the native Korean speakers recorded higher F2 values a significant amount of the time in English than native English speakers, but this significance does not hold up for F2 values in Korean or F1 values in either language. This does mean, however, that in at least one context, the native Korean speakers potentially seem to sing more fronted tongues than the native English speakers (high F2). This would mean the Korean speakers have a different “basis of articulation” than the English speakers, but this result is unfortunately not consistent across the data. Additionally, this does not necessarily coincide with what is seen in the average formant value tables by language in the introduction, as there are a handful of examples where F2 is lower on average for Korean speakers for corresponding English vowels in all dialects. For instance, Korean [u] is stated to, on average, have an F2 value of 876 Hz. Whereas English [u] is stated to, on average, have an F2 value of 920 Hz in British English, 1393 Hz in American English, and 1000 Hz in Australian English. Thus, the results in which native Korean speakers produce higher F2 values on average in their English singing data appears to be “atypical”, or at least not wholly representative of a difference solely based on native vowel inventory. Additionally, because of the context in which this statistically significant data was collected, along with how close the derived *p-value* (0.0446) for the data was to the typical threshold for significance (0.05), it may be appropriate to adjust the threshold used for this experiment and consider this data insignificant as well. When compared to the results of the other three tests along with the data from previous experiments as detailed above, it would not be a stretch to assume that this "significance" may just be a fluke based on interference in the data, an outlier that was unaccounted for, or a variety of other issues.

There are a variety of factors that could have an unaccounted-for effect on the results of the statistical tests performed. First, the quality of the audio recordings used could have had an effect on the results. As mentioned in the methods section, the a capella clips had to be manually created using an AI program that separates instrumentals from vocals. However, just like any other technological program, AI is not without its own faults and limitations. Due to the musical style and production of songs by Stray Kids, the instrumentals were likely very hard to completely separate from the vocals, even if it would appear that way to the standard human ear. Additionally, vocal effects like reverb and autotune cannot be removed at all, since they are part of the vocal line and note the instrumental line. These issues could create an amount of interference in the clips used in the corpus. Depending on how much interference is left in a given clip, this may have affected the formant values extracted by Praat's interface. There is no real way to tell if this would affect each token in exactly the same manner, but it does suggest that there may at least be some level of inaccuracy to the formant values collected because of interference.

Another factor that could have affected the formant values extracted from the data is the concept of a "singing formant". As noted in the introduction, there is research to suggest that, when singing vowels, there is a region of high energy within the spectrum of formants that we do not see in speech. In male singers, like those of Stray Kids, that region of high energy lies somewhere around 1800 to 2500 Hz according to studies conducted on the matter. This is, unfortunately, very close to the target frequencies for F2 values in most, if not all, vowels in both languages. This means that there is a chance that Praat's formant extraction command could potentially extract data based on this "singer's formant" instead of the real F2 value. However, the singing formant is not uniformly above or below the value of F2 according to either aforementioned study. Thus, this cannot be corrected for by also selecting for a potential third formant, and then either disregarding F2 as the "singer's formant" if it uniformly occurs below the real F2, or just disregarding that third formant all together if the "singer's formant" routinely occurs above the real F2. Additionally, because this would mean two regions of high energy potentially occurring very close to one another, close enough to overlap, it could mean that Praat's system merges both into one large formant region, and takes the average value between the two. Which, again, is not ideal. In any of these cases, the "singer's formant" could be obstructing the analysis of the data from this corpus.

A third and final factor that could have affected some of the formant values collected is the condensing of diphthongs into monophthongs. As previously explained, proper vocal technique often requires extended duration of vowel sounds to maintain a musical note for the specific duration. For monophthongs this is simple – the single vowel sound is extended. However, for diphthongs this is not the case. In the case of diphthongs, one of the two parts must be extended, and the transition between the vowel sounds will either come at the very beginning or the very end of the extension. Generally speaking, it will be the first of the two vowel sounds that extends for "true" diphthongs like [ɑɪ] and the second sound for diphthongs that include an approximant like [jʌ]. For this project, only the extended part of the diphthongs was measured. Diphthongs are of course dynamic though, which means that the transition from one vowel to the other will affect the extracted formant vowels if the annotated interval contains even a fraction of that change. Thus, this would mean that there is a chance that the annotated portions of diphthongs might include some of the transition noise, which would affect their measurements.

To determine if this lack of difference in any F1 values in either language and F2 values in Korean, as well as the significant in F2 values in which F2 are higher for native Korean speakers than native English speakers in English, is truly applicable to the larger population or if the differences recorded only hold true in this one small context, it would be beneficial to try to create a more advanced mixed effects model so that formant values can be looked at as both a function of proficiency and vowel. However, due to the number of vowels involved in each language, establishing meaningful orthogonal contrasts for this variable would prove to be extremely complicated. Creating a 14-way contrast or layered contrast for the English vowels might genuinely be impossible. This is especially true given that since the vowel inventory for this project includes monophthongs from multiple dialects of English, and therefore some of the attribute labels like height and fronting may “overlap” between some vowels based on their typical dialect. Therefore, at this time the project will remain limited to the simple mixed-model in which formant values are mapped only as a function of proficiency.

Additionally, another way to approach this limited model issue and see if the trend of native Korean speakers having higher F2 measurements in English yet finding no significant differences for other measurements persists in the population is to split the data into smaller parts by vowel. For instance, a mixed-model test could be run solely on the [ʌ] tokens from Korean, where the formants are again mapped as functions of proficiency, but this time with participant as the random variable instead of vowel. This type of set-up would allow the analyses to show individualized data for each vowel type within each language. However, comparisons across all vowels in a language would be difficult to perform in this case, since they are now being run through separate models with different details (like degrees of freedom). Also, just because the analyses of each language as a whole with vowels and speakers mapped as random variables yields statistically significant results in some contexts does not mean that tests for each individual vowel would do the same. There could be certain vowels within either of the languages that simply do not yield anything significant, or even be statistically significant in the opposite direction to what would be expected based on the whole-language results. Thus, as positive as this approach may seem at first glance, it seems to further complicate the results instead of fixing the issue it aimed to correct.

In the future, I think that a plethora of work could be done on similar topics to this project regarding vowel measurements while singing in the context of different languages and different proficiencies. In particular, there are two projects I think would be fruitful to add to the existing literature. First is a project quite similar to the one conducted for this paper. The difference would be that, instead of taking pre-existing data full of vocal editing and instrumental interference to try and work around and forcing it to become a corpus, new data would be recorded solely for the purpose of the experiment. The experimenters could then have control over exactly what lyrics and melody were used, how many iterations of each vowel would occur, and there would be no noise interference from digitally removed instrumentals and effects. In particular, children’s songs or folk songs would probably make a good choice for the corpus, given that such songs are often short in length and can be easily recorded in one stretch, the melodies are easy for untrained singers to replicate, and the lyrics are usually not difficult to remember, so that the participants could sing from memory without the need to reference any written material (though that could be provided if needed). In this way, it could be determined if the results of my experiment could be replicated with other data, or if the statistical significance observed here is just an anomaly.

Another project that would be great to see in the future would be one in which proficiency's effect over time on the formants of sung vowels was tested. By this I mean that singers who are learning an L2 would be the test subjects, and data would be collected as their proficiency in their L2 changes. I attempted a project on this myself, with a similar setup to the current project in terms of using already existent data to make my corpus, and the results I extracted were statistically significant. However, the data for that project had such significant instrumental interference in comparison to the data used for the project in this paper, so those results may not be replicable. Instead, it would be beneficial to take several subjects with no fluency in an L2 and record their singing data in said L2 language. Then, over time, have the test subjects take courses to improve their fluency in their L2, and re-record their singing data at several chosen time milestones. Then, the values could be compared. Additionally, this aforementioned longitudinal study could be studied in "reverse", in which the test subjects would be proficient in an L2 to begin with, but over time would lose their proficiency due to lack of practice or environmental factors. This would be a much more complicated approach, but I think it would make for a fascinating compliment to the more traditional longitudinal study.

It would also be fruitful to conduct either or both of these aforementioned projects on languages within and outside of the same language families. For instance, the first proposed follow-up experiment that most closely replicates my own in this paper could be conducted once comparing English and German data, two languages in the Germanic branch of the Indo-European language family, and then once again with German from the Indo-European family, and Xhosa from the Niger-Congo language family, which are, of course, two languages in dramatically different language families, Then the results of both studies could be compared in a third sub-study. This dichotomy could also be used for the proposed longitudinal experiment. This way, researchers could document any differences based not only on native speakers versus late learners, but also see whether or not the relatedness of the languages has an effect on the results recorded.

Despite the flaws that may exist in some of the data, this case study proves to be useful in the larger context of the literature. Depending upon interpretation, the results in which the differences in formant values measured for native speakers of Korean and English are statistically significant can support the findings of other previous works in the literature. In particular, the theory of a "singer's formant" and "basis of articulation" could very well be at play in these results. Also, this project opens the door for a multitude of follow up studies that could also prove fruitful additions to the literature. Lastly, as mentioned in the introduction, the results of this project and follow up projects could be important within the sphere of social issues, particularly discrimination and bias in the world of music, both professionally and as consumers. Music is meant to be something that unites human kind, not something that causes division. If the biases people hold about people having "better" or "more correct" singing in a language based on their language background, and use this as a means to ostracize L2 learners as artists, this could be detrimental. Instead, any statistically significant results of this project and any further project should be used as a way to show the unique approach to production that L2 learners bring to modern music, and to celebrate the diversity of sound we get to experience in today's world of entertainment.

Works Cited

Arnaud, Jourdan. "A Brief History of K-Pop." *The Los Angeles Film School*, 20 Feb. 2023, www.lafilm.edu/blog/a-brief-history-of-kpop/.

Baker, David John, et al. "Embracing Anti-Racist Practices in the Music Perception and Cognition Community." *Music Perception*, vol. 38, no. 2, 2020, pp. 103–105., https://doi.org/10.1525/mp.2020.38.2.103.

Bloothooft, Gerrit, and Reinier Plomp. "Spectral Analysis of Sung Vowels. I. Variation Due to Differences between Vowels, Singers, and Modes of Singing." *The Journal of the Acoustical Society of America*, vol. 75, no. 4, 1984, pp. 1259–1264., https://doi.org/10.1121/1.390732.

Bloothooft, Gerrit, and Reinier Plomp. "Spectral Analysis of Sung Vowels. II. the Effect of Fundamental Frequency on Vowel Spectra." *The Journal of the Acoustical Society of America*, vol. 77, no. 4, 1985, pp. 1580–1588., https://doi.org/10.1121/1.392001.

Bloothooft, Gerrit, and Reinier Plomp. "Spectral Analysis of Sung Vowels. III. Characteristics of Singers and Modes of Singing." *The Journal of the Acoustical Society of America*, vol. 79, no. 3, 1986, pp. 852–864, https://doi.org/10.1121/1.393423.

Boersma, Paul & Weenink, David (2023). Praat: doing phonetics by computer [Computer program]. Version 6.3.09, 25 March 2023 from http://www.praat.org/

Božović, Đorđe, and Zorka Kašić. "Basis of Articulation and the Phonetics-Phonology Interface." *Philologia*, vol. 19, no. 19, 2021, pp. 1–10, https://doi.org/10.18485/philologia.2021.19.19.1.

Bradley, Evan D. "A Comparison of the Acoustic Vowel Spaces of Speech and Song." *Linguistic Research*, vol. 35, no. 2, 2018, pp. 381–394., https://doi.org/10.17250/khisli.35.2.201806.006

Cleveland, Thomas F. "A Clearer View of Singing Voice Production: 25 Years of Progress." *Journal of Voice*, vol. 8, no. 1, 1994, pp. 18–23., https://doi.org/10.1016/s0892-1997(05)80315-7.

Crystal, David, et al. "English Language." *Encyclopædia Britannica*, 25 Apr. 2023, www.britannica.com/topic/English-language.

Davidson, Jane W., and Richard Edgar. "Gender and Race Bias in the Judgement of Western Art Music Performance." *Music Education Research*, vol. 5, no. 2, 2003, pp. 169–181., https://doi.org/10.1080/1461380032000085540.

"English Phonology." *Wikipedia*, Wikimedia Foundation, 25 Mar. 2023, https://en.wikipedia.org/wiki/English_phonology.

Ferjan, Matija. “30+ People Listening to Music Statistics & Trends (2023)M.” *HeadphonesAddict*, 18 Apr. 2023, headphonesaddict.com/listening-to-music-statistics

Garofalo, Reebee. “Whose World, What Beat: The Transnational Music Industry, Identity, and Cultural Imperialism.” *The World of Music*, vol. 35, no. 2, 1993, pp. 16–32. *JSTOR*, http://www.jstor.org/stable/43615564.

Gregg, Jean Westerman, and Ronald C. Scherer. “Vowel Intelligibility in Classical Singing.” *Journal of Voice*, vol. 20, no. 2, 2006, pp. 198–210., https://doi.org/10.1016/j.jvoice.2005.01.007.

Huckvale, Mark. “Week 5.” *PLIN0065 Intermediate Phonetics*, 31 Oct. 2020, www.phon.ucl.ac.uk/courses/plin0065/week5.php.

JYP Entertainment, “Stray Kids 5-STAR Teaser Image”, *Twitter*, 16 May 2023, https://twitter.com/Stray_Kids/status/1658487751883501573?s=20.

Kim, Sookyung. “Racism in the Global Era: Analysis of Korean Media Discourse around Migrants, 1990–2009.” *Discourse & Society*, vol. 23, no. 6, 2012, pp. 657–678., https://doi.org/10.1177/0957926512455381.

Koopman, John. “Antiquity to 1950.” *A Brief History of Singing*, John Koopman, Appleton, WI, 1994.

“Korean Phonology.” *Wikipedia*, Wikimedia Foundation, 25 Mar. 2023, https://en.wikipedia.org/wiki/Korean_phonology.

Lee, Sean. “A Sketch of Language History in the Korean Peninsula.” *PLOS ONE*, vol. 10, no. 5, 2015, https://doi.org/10.1371/journal.pone.0128448.

Melnik, Dmitry. “Vocal Remover and Isolation.” *[AI]*, 2023, vocalremover.org.

Menegaz, Felipe, and Peter Fitzgerald. “File:World Map Percentage English Speakers by Country.Svg.” *Wikimedia Commons*, 15 Dec. 2009, commons.wikimedia.org/wiki/File:World_map_percentage_english_speakers_by_country.svg.

Morgan, Sally. “Korean Language Dialects List / Map.” *ImportanceofLanguages.com*, 15 Feb. 2019, https://www.importanceoflanguages.com/common-korean-language-dialects/.

“Music genre.” *Wikipedia*, Wikimedia Foundation, 25 Mar. 2023, https://en.wikipedia.org/wiki/Music_genre.

New, Tin L, et al. “Voice Conversion: From Spoken Vowels to Singing Vowels.” *2010 IEEE International Conference on Multimedia and Expo*, 2010, https://doi.org/10.1109/icme.2010.5582961.

Oh, Eunjin. “Dynamic Spectral Patterns of American English Front Monophthong Vowels Produced by Korean-English Bilingual Speakers and Korean Late Learners of English.” *Linguistic Research*, vol. 30, no. 2, 2013, pp. 293–312., https://doi.org/10.17250/khisli.30.2.201308.007.

Oh, Eunjin. “Learning Dynamic Transitions of the American English Vowels [Eɪ] and [Oʊ]: A Comparison between Korean-English Bilingual Speakers and Korean Learners of English.” *The Journal of Studies in Language*, vol. 29, no. 2, 2013, pp. 323–343., https://doi.org/10.18627/jslg.29.2.201308.323.

Oh, Eunjin. “The High Vowel Systems of Korean-English Bilingual Speakers.” *Korean Journal of English Language and Linguistics*, vol. 9, no. 4, 2009, pp. 633–661., https://doi.org/10.15738/kjell.9.4.200912.633.

Prasad, Aarathai. “When Did Humans First Make Music?” *BBC Science Focus Magazine*, 11 Feb. 2021, www.sciencefocus.com/the-human-body/when-did-humans-first-make-music/.

RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/

Sager, Jessica. “Taylor Swift Net Worth 2023 - Parade: Entertainment, Recipes, Health ...” *Parade*, 22 May 2023, parade.com/1133777/jessicasager/taylor-swift-net-worth/

Stout, Barrett. “The Harmonic Structure of Vowels in Singing in Relation to Pitch and Intensity.” *The Journal of the Acoustical Society of America*, vol. 10, no. 2, 1938, pp. 137–146., https://doi.org/10.1121/1.1915969.

Stray Kids (JYP Entertainment), *Youtube* — 2023,

https://www.youtube.com/watch?v=GjuJ8LlGEoE&ab_channel=StrayKids-Topic

https://www.youtube.com/watch?v=NlX5i9oZN1I&ab_channel=StrayKids-Topic

https://www.youtube.com/watch?v=TgnAn_OVQtA&list=PLBXl392FRRgztl3cMgkQ56iiInAqlPwO2&index=4&ab_channel=StrayKids-Topic

https://www.youtube.com/watch?v=r_G906iQID8&ab_channel=StrayKids-Topic

https://www.youtube.com/watch?v=jiQFAC3FWg0&list=PLDhajrZgo0TLwm2eq2XaqxGwM1TTrJTC0&index=9&ab_channel=StrayKids-Topic

https://www.youtube.com/watch?v=i-ju-Auf-C8&list=PLKkZQUgBQ1JM5epHzZretq_U4Dh5sohoO&ab_channel=StrayKids-Topic

https://www.youtube.com/watch?v=8PruSzsNDlw&list=PLR8wLBnjpTpHhFk-0Qon6Vfl4q9ihR7S0&index=16&ab_channel=StrayKids-Topic

https://www.youtube.com/watch?v=yXs1bro7wvw&ab_channel=StrayKids-Topic

https://www.youtube.com/watch?v=ysHAB8O_dQs&list=PLR8wLBnjpTpHhFk-0Qon6Vfl4q9ihR7S0&index=14&ab_channel=StrayKids-Topic

https://www.youtube.com/watch?v=-xwP0DAzqdY&list=PLR8wLBnjpTpHhFk-0Qon6Vfl4q9ihR7S0&index=18&ab_channel=StrayKids-Topic

https://www.youtube.com/watch?v=_r78BsRU9Zw&list=PLR8wLBnjpTpHhFk-0Qon6Vfl4q9ihR7S0&index=9&ab_channel=StrayKids-Topic

https://www.youtube.com/watch?v=Fvv2RtQc6S4&ab_channel=StrayKids-Topic

https://www.youtube.com/watch?v=bbDenQ8jEAs&ab_channel=StrayKids-Topic

https://www.youtube.com/watch?v=WVPd9dZqo8Y&ab_channel=StrayKids-Topic

https://www.youtube.com/watch?v=_BeIFoLfiXg&ab_channel=StrayKids-Topic

https://www.youtube.com/watch?v=HcyaBfXU1wk&ab_channel=StrayKids-Topic

https://www.youtube.com/watch?v=hpq9nZpU34I&list=OLAK5uy_lZDk4rd2jtBxv8Yt4k0u19IDcu8oncpXU&ab_channel=StrayKids-Topic

https://www.youtube.com/watch?v=-OofaCbwiow&ab_channel=StrayKids-Topic

https://www.youtube.com/watch?v=xd-uC02Qmmc&ab_channel=StrayKids-Topic

https://www.youtube.com/watch?v=E9PNeH9kJ9w&list=PLR8wLBnjpTpHhFk-0Qon6Vfl4q9ihR7S0&index=6&ab_channel=StrayKids-Topic

https://www.youtube.com/watch?v=xoSXk690IS4&ab_channel=StrayKids-Topic

https://www.youtube.com/watch?v=wGsYHSBbVV4&list=PLhURll6XRcxevwHsEDPpTY_XYr7nNNduP&index=5&ab_channel=StrayKids-Topic

https://www.youtube.com/watch?v=ngBk3v4FWTM&list=PL4w0tSEAwBih2XBA1hJ_oKDHKlydQBwlj&ab_channel=StrayKids-Topic

Sundberg, Johan. “The Acoustics of the Singing Voice.” *Scientific American*, vol. 236, no. 3, 1977, pp. 82–91., https://doi.org/10.1038/scientificamerican0377-82.

synSinger. “Men, Woman and Children.” *synSinger*, 3 Dec. 2016, synsinger.wordpress.com/2016/11/24/men-woman-and-children/.

Yang, Byunggon. “A Comparative Study of American English and Korean Vowels Produced by Male and Female Speakers.” *Journal of Phonetics*, vol. 24, no. 2, 1996, pp. 245–261, https://doi.org/10.1006/jpho.1996.0013.

Yang, Byunggon. “Relative Distances among English Front Vowels Produced by Korean and American Speakers.” *The Journal of the Acoustical Society of America*, vol. 136, no. 4, 2014, pp. 2146–2146., https://doi.org/10.1121/1.4899750.

“YouTube to WAV.” *Motionbox*, 2023, https://motionbox.io/tools/youtube-to-wav.

Unknown Authors, *Kpop Profiles*, — 2023, https://kprofiles.com/stray-kids-discography/., https://kprofiles.com/stray-kids-members-profile-facts/.

Unknown Authors, *Color Coded Lyrics*, — 2023,

https://colorcodedlyrics.com/2020/06/17/stray-kids-airplane-bihaenggi/

https://colorcodedlyrics.com/2020/06/17/stray-kids-another-day-ilsang/

https://colorcodedlyrics.com/2020/09/14/stray-kids-any-ani/

https://colorcodedlyrics.com/2020/06/17/stray-kids-blueprint-cheongsajin/

https://colorcodedlyrics.com/2022/10/10/stray-kids-circus-korean-ver/

https://colorcodedlyrics.com/2022/12/21/stray-kids-fam-korean-ver/

https://colorcodedlyrics.com/2020/03/17/stray-kids-i-am-you-ot8-ver/

https://colorcodedlyrics.com/2020/01/23/stray-kids-levanter-english-ver/

https://colorcodedlyrics.com/2020/03/17/stray-kids-m-i-a-ot8-ver/

https://colorcodedlyrics.com/2020/03/17/stray-kids-miroh-ot8-ver/

https://colorcodedlyrics.com/2020/03/17/stray-kids-mirror-ot8-ver/

https://colorcodedlyrics.com/2019/12/26/stray-kids-mixtape-gone-days/

https://colorcodedlyrics.com/2020/03/25/stray-kids-mixtape-on-track-babolado-ala/

https://colorcodedlyrics.com/2022/03/19/stray-kids-muddy-water/

https://colorcodedlyrics.com/2020/06/17/stray-kids-pacemaker/

https://colorcodedlyrics.com/2021/12/23/stray-kids-placebo/

https://colorcodedlyrics.com/2021/12/23/stray-kids-scars-korean-ver/

https://colorcodedlyrics.com/2021/08/23/stray-kids-silent-cry/

https://colorcodedlyrics.com/2019/12/09/stray-kids-stop/

https://colorcodedlyrics.com/2020/03/17/stray-kids-spread-my-wings-eolin-nalgae-ot8-ver/

https://colorcodedlyrics.com/2019/11/07/stray-kids-story-that-won-t-end-ggeutnaji-anheul-iyagi/

https://colorcodedlyrics.com/2021/08/23/stray-kids-the-view/

https://colorcodedlyrics.com/2022/03/18/stray-kids-venom-geomijul/

Appendix 1

1.1 AIRPLANE (2020)

KOREAN/ENGLISH

멀리 떠나 올라타자 비행기

너와 나 함께 타는 비행기

>>>

We are, we are 어디든 가요

We are, we are 너와는 어디든지 가지

>>>

처음부터 넌

내 심장을 흔들고 나를 미치게 해

>>>

멀리 떠나 올라타자 비행기

너와 나 함께 타는 비행기

>>>

We are, we are 어디든 가요

We are, we are 너와는 어디든지 가지

>>>

We are, we are

>>>

처음부터 넌

내 심장을 흔들고 나를 미치게 해

>>>

We are, we are

>>>

We are, we are 어디든 가요

We are, we are 너와는 어디든지 가지

1.2 ANOTHER DAY (2020)

KOREAN

그저 두 발 뻗고 눕는다고

눈이 감기지 않아

눈을 질끈 감아봐도 다시

꿈이 없는 잠을 자

>>>

그저 두 발 뻗고 눕는다고

눈이 감기지 않아

눈을 질끈 감아봐도 다시

꿈이 없는 잠을 자

1.3 ANY (2020)

KOREAN/ENGLISH

이것도 아닌데 저것도 아냐

What do you want baby

What do you want

>>>

별의 별건 많아 근데 다 별로

별의 별건 많아 근데 다 발로

>>>

별의 별건 많아 근데 다 별로

별의 별건 많아 근데 다 발로

>>>

이것도 아닌데 저것도 아냐

What do you want baby

What do you want

>>>

별의 별건 많아 근데 다 별로

별의 별건 많아 근데 다 발로

>>>

별의 별건 많아 근데 다 별로

별의 별건 많아 근데 다 발로

>>>

별의 별건 많아 근데 다 별로

별의 별건 많아 근데 다 발로

>>>

별의 별건 많아 근데 다 별로

별의 별건 많아 근데 다 발로

1.4 BLUEPRINT (2020)

KOREAN/ENGLISH

Watch me do what I want

>>>

날 좀 내버려 둬

>>>

저 멀리 손짓하는 햇살 속에

푸른 꿈이 나를 비춰

>>>

난 계속 멈추지 않고서 달려갈 걸

절대 지치지 않고서 달려갈 걸

저기 닿을 듯한 빛은 꺼지지 않아

>>>

Watch me do what I want

>>>

날 좀 내버려 둬

>>>

저 멀리 손짓하는 햇살 속에

푸른 꿈이 나를 비춰

>>>

난 계속 멈추지 않고서 달려갈 걸

절대 지치지 않고서 달려갈 걸

저기 닿을 듯한 빛은 꺼지지 않아

>>>

저 멀리 손짓하는 햇살 속에

푸른 꿈이 나를 비춰

>>>

난 계속 멈추지 않고서 달려갈 걸

절대 지치지 않고서 달려갈 걸

저기 닿을 듯한 빛은 꺼지지 않아

1.5 CIRCUS KOREAN EDITION (ORIGINAL JAPANESE RELEASE 2022, KOREAN RE-RELEASE 2022)

KOREAN/ENGLISH

Bring out the fire

>>>

링 링 링 저글링 we gon' burn this down

>>>

삐에로 rockstar

>>>

링 딩 딩 디기딩 we gon' run this town

We gon' run this town

>>>

The show starts with a bang bang bang

밤낮없이 놀아 play play play

>>>

고삐를 풀어 ready

끝도 없이

Round and round and round and round

>>>

The show must go on

작은 불씨들을 일으켜

>>>

Focus, stand up

>>>

이제 시작인데 어디가

>>>

Welcome to the zone

>>>

Feelin’ good right now, feelin’ good right now

>>>

멋대로 움직이는 몸

>>>

Feelin’ good right now, feelin’ good right now

>>>

눈과 귀는 이미 사로잡았고

정신 줄마저 집 나갔죠

>>>

So welcome to my home

>>>

Feelin’ good right now, feelin’ good right now

>>>

Bring out the fire

>>>

링 링 링 저글링 we gon’ burn this down

>>>

삐에로 rockstar

>>>

링 딩 딩 디기딩 we gon’ run this town

We gon’ run this town

>>>

The show starts with a bang bang bang

밤낮없이 놀아 play play play

>>>

고삐를 풀어 ready

끝도 없이

Round and round and round and round

>>>

The show must go on

작은 불씨들을 일으켜

>>>

Focus, stand up

>>>

이제 시작인데 어디가

>>>

Welcome to the zone

>>>

Feelin' good right now, feelin' good right now

>>>

멋대로 움직이는 몸

>>>

Feelin' good right now, feelin' good right now

>>>

눈과 귀는 이미 사로잡았고

정신 줄마저 집 나갔죠

>>>

So welcome to my home

>>>

Feelin' good right now, feelin' good right now

1.6 FAM KOREAN EDITION (ORIGINAL JAPANESE RELEASE 2020, KOREAN RE-RELEASE 2022)

KOREAN/ENGLISH

하나둘씩 모여서

하나같이 움직여

Everybody put your hands up

Stray Kids sing it loud loud loud

>>>

우린 서로서로가 너무 잘 알아

이렇게 다 같이 무대 위를 날아

>>>

가족보다 많이 보는 사이

봐도 봐도 안 질리는 타입

>>>

하나둘씩 모여서

하나같이 움직여

Everybody put your hands up

Stray Kids sing it loud loud loud

>>>

우린 서로서로가 너무 잘 알아

이렇게 다 같이 무대 위를 날아

>>>

가족보다 많이 보는 사이

봐도 봐도 안 질리는 타입

>>>

하나둘씩 모여서

하나같이 움직여

Everybody put your hands up

Stray Kids sing it loud loud loud

>>>

우린 서로서로가 너무 잘 알아

이렇게 다 같이 무대 위를 날아

>>>

가족보다 많이 보는 사이

봐도 봐도 안 질리는 타입

1.7 I AM YOU (ORIGINAL RELEASE 2018, RE-RECORD 2020)

KOREAN/ENGLISH

I am YOU, I see me in you

너와 있을 때 난 알 수 있어

>>>

I found YOU, I found me in you

그 안에서 나의 모습이 보여

>>>

같은 공간 속에서, 같은 시간 속에서

>>>

Let me run, let me run, let me run along with YOU

>>>

I am YOU, I see me in you

너와 있을 때 난 알 수 있어

>>>

I found YOU, I found me in you

그 안에서 나의 모습이 보여

>>>

같은 공간 속에서, 같은 시간 속에서

>>>

Let me run, let me run, let me run along with YOU

1.8 LEVANTER ENGLISH EDITION (ORIGINAL KOREAN RELEASE 2019, ENGLISH RE-RELEASE 2020)

ENGLISH

I wanna be myself

Yeah I gotta be myself

>>>

I′m listening to my heart, let it guide me

I feel the light, I feel the light

>>>

I wanna be myself

Yeah I gotta be myself

>>>

I′m listening to my heart, let it guide me

I feel the light, I feel the light

>>>

I wanna be myself (I don′t care)

Yeah I gotta be myself (Just don′t care)

1.9 M.I.A (ORIGINAL RELEASE 2018, RE-RECORD 2020)

KOREAN

대체 어딜 간 거야 너의 그 예쁜 미소

워낙 잘 웃던 너라서 더 그리워져

내겐 다 보여 안 그런 척 어색한 미소

밝았던 너라 더 걱정이 돼

>>>

대체 어딜 간 거야 너의 그 예쁜 미소

워낙 잘 웃던 너라서 더 그리워져

내겐 다 보여 안 그런 척 어색한 미소

밝았던 너라 더 걱정이 돼

1.10 MIROH (ORIGINAL RELEASE 2019, RE-RECORD 2020)

KOREAN/ENGLISH

힘들지 않아 거친 정글 속에

뛰어든 건 나니까 I'm okay

>>>

We goin' higher 다음 도시 속에

빌딩들 내려보며 fly all day

>>>

힘들지 않아 거친 정글 속에

뛰어든 건 나니까 I'm okay

>>>

We goin' higher 다음 도시 속에

빌딩들 내려보며 fly all day

>>>

힘들지 않아 거친 정글 속에

뛰어든 건 나니까 I'm okay

>>>

We goin' higher 다음 도시 속에

빌딩들 내려보며 fly all day

1.11 MIRROR (ORIGINAL RELEASE 2018, RE-RECORD 2020)

ENGLISH

For once I wish I could see myself

>>>

For once I wish I could see myself

>>>

For once I wish I could see myself

1.12 MIXTAPE: GONE DAYS (2019)

KOREAN/ENGLISH

Gone days

왜 자꾸 다 안된다는 건데

지나간 과거 신경 꺼 ey

>>>

This is the new generation, go away

>>>

Gone days

왜 자꾸 다 안된다는 건데

지나간 과거 신경 꺼 ey

>>>

This is the new generation, go away

>>>

Gone days

왜 자꾸 다 안된다는 건데

지나간 과거 신경 꺼 ey

>>>

This is the new generation, go away

1.13 MIXTAPE: ON TRACK (2020)

ENGLISH

One more step, I will never stop

I'll always be on track

>>>

One more step, I will never stop

I'll always be on track

>>>

One more step, I will never stop

I will always be on track

1.14 MUDDY WATER (2022)

Feeling so fly, feeling so fly

Feeling so fly, feeling so fly

This my life style, this my life style

I'm breathing, working, moving

Sit down and listen for once

>>>

Feeling so fly, feeling so fly

Feeling so fly, feeling so fly

This my life style, this my life style

I'm breathing, working, moving

Sit down and listen for once

1.15 PACEMAKER (2020)

KOREAN/ENGLISH

발맞춰 run run

뒤돌아보지 마 이끌어 줄게 널

>>>

Just run run

>>>

너에게 맞춰 그다음 다음

너의 plan대로 가자고 다음다음

>>>

발맞춰 run run

뒤돌아보지 마 이끌어 줄게 널

>>>

Just run run

>>>

너에게 맞춰 그다음 다음

너의 plan대로 가자고 다음다음

>>>

발맞춰 run run

뒤돌아보지 마 이끌어 줄게 널

>>>

Just run run

>>>

너에게 맞춰 그다음 다음

너의 plan대로 가자고 다음다음

1.16 PLACEBO (ORIGINAL RELEASE 2018, RE-RECORD 2020)

ENGLISH

Cause later when I become addicted to life

>>>

Cause later when I become addicted to life

>>>

Cause later when I become addicted to life

1.17 SCARS KOREAN EDITION (ORIGINAL JAPANESE RELEASE 2021, KOREAN RE-RELEASE 2021)

KOREAN/ENGLISH

I'll never cry

Because I know that it'll never change

아무도 모르는 곳 그저 버티고 서 있어 난

주저앉아서 몇 번 손 내밀어 봤지만

결국 내 손은 바닥을 짚고 일어나 다시

>>>

밤새며 난 I turn on the lights 멈추지 않아

>>>

I will never give up

>>>

더 파고들어

구겨진 모습

>>>

I'll never cry

Because I know that it'll never change

아무도 모르는 곳 그저 버티고 서 있어 난

주저앉아서 몇 번 손 내밀어 봤지만

결국 내 손은 바닥을 짚고 일어나 다시

>>>

밤새며 난 I turn on the lights 멈추지 않아

>>>

I will never give up

>>>

더 파고들어

구겨진 모습

1.18 SILENT CRY (2021)

KOREAN/ENGLSIH

말없이 외치는 너의 목소리를 내가 들어 줄게

>>>

서툴게 닫아 둔 너라는 그 공간

문틈 사이로

>>>

내게만 들리는 Silent Cry

그동안 숨겨둔 Silent Cry

>>>

너만 아파하지 마

내게 줘 네 Silent Cry

>>>

말없이 외치는 너의 목소리를 내가 들어 줄게

>>>

서툴게 닫아 둔 너라는 그 공간

문틈 사이로

>>>

내게만 들리는 Silent Cry

그동안 숨겨둔 Silent Cry

>>>

너만 아파하지 마

내게 줘 네 Silent Cry

1.19 SPREAD MY WINGS (ORIGINAL RELEASE 2018, RE-RECORD 2020)

KOREAN/ENGLISH

Naturally

>>>

Spread my wings

>>>

어른인 척 할 때가 좋은듯해 안 변할래

철부지 같다는 말 듣는데도

>>>

뭐든 서툴지만 처음이란 단어가 좋아

어리숙하지만 이대로만

>>>

Naturally

>>>

Spread my wings

>>>

어른인 척 할 때가 좋은듯해 안 변할래

철부지 같다는 말 듣는데도

>>>

뭐든 서툴지만 처음이란 단어가 좋아

어리숙하지만 이대로만

1.20 STOP (2019)

KOREAN/ENGLISH

I’ll go wherever I wanna go

>>>

나침반을 보고 걸어가

어딜 가든 상관없어 나

>>>

벽이 막아서도 다른 길로 가

I'll go wherever I wanna go

>>>

발을 들어 그냥 내딛고 나가

I'll go wherever I wanna go

I'll go wherever I wanna go

>>>

나침반을 보고 걸어가

어딜 가든 상관없어 나

벽이 막아서도 다른 길로 가

I'll go wherever I wanna go

>>>

발을 들어 그냥 내딛고 나가

I'll go wherever I wanna go

1.21 STORY THAT WON'T END (2019)

KOREAN/ENGLISH

Never say goodbye 너와 난 하나니까

같은 꿈 속을 함께 거닐 테니까

>>>

Never say goodbye 너와 난 하나니까

같은 꿈 속을 함께 거닐 테니까

1.22 THE VIEW (2021)

KOREAN/ENGLISH

그동안 몰랐어 앞이 어두웠어

Cause I never ever let it go

>>>

I like The View right now

I like The View right now

>>>

I like The View right now

I like The View right now

그동안 몰랐어 앞이 어두웠어

Cause I never ever let it go

>>>

I like The View right now

I like The View right now

>>>

I like The View right now

I like The View right now

>>>

I like The View right now

I like The View right now

>>>

I like The View right now

>>>

I like The View right now

>>>

I like The View right now

>>>

I like The View right now

1.23 VENOM (2022)

KOREAN/ENGLISH

거미줄

거미줄

>>>

거미줄

거미줄

>>>

No escape 사라진 감각 can’t feel my fingers

걸렸어 You got me wrapped up

Around your fingers

>>>

퍼져 가는 poison 중독되는 potion

>>>

거미줄

거미줄

>>>

거미줄

거미줄

>>>

No escape 사라진 감각 can’t feel my fingers

걸렸어 You got me wrapped up

Around your fingers

>>>

퍼져 가는 poison 중독되는 potion

>>>

거미줄

거미줄

>>>

거미줄

거미줄