

Fiat Lingua

Title: "Lekhmove tawak?": A phonological comparison between a constructed language, Dothraki, and mixed languages.

Author: Kato Kuijpers

MS Date: 03-18-2024

FL Date: 01-01-2025

FL Number: FL-000100-00

Citation: Kuijpers, Kato. 2024. "'Lekhmove tawak?": A phonological comparison between a constructed language, Dothraki, and mixed languages." FL-000100-00, *Fiat Lingua*, <<https://fiatlingua.org>>. Web. 01 January 2025.

Copyright: © 2024 Kato Kuijpers. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License.



<http://creativecommons.org/licenses/by-nc-nd/3.0/>

**“Lekhmove tawak?”¹: A phonological comparison between a constructed language,
Dothraki, and mixed languages.**

Kato Kuijpers

S3772225

MA in Multilingualism

Faculty of Arts

University of Groningen

Supervisors:

Dr. M. Mazzoli

Dr. F. H. Eiró de Oliveira

17 June, 2022

MA thesis, Departments of Applied Linguistics and Frisian Language and Culture, Faculty of
Arts, Rijksuniversiteit Groningen

15.686 words

¹ Dothraki for: “is a constructed language real?/ a real constructed language?”

Acknowledgements

First and foremost, I would like to express my gratitude to my supervisor, dr. Maria Mazzoli, whose constructive feedback and knowledge of mixed languages have helped me countless times in the process of writing this thesis. Secondly, I am extremely grateful to prof. dr. F. Gobbo whose expertise in the field of interlinguistics and constructed languages was important for setting up my research and who pointed me in the right direction in terms of the existing literature on the topic.

Thanks should also go to dr. Hanneke Loerts for her advice on the statistical aspects of my study and to the Language Diversity Research Group for letting me present my initial findings and giving me feedback on it. Lastly, I would also like to thank dr. F. H. Eiró de Oliveira for being my second reader.

Table of Contents

0. Abstract.....	5
1. Introduction.....	6
1.1 Outline of the Study	7
2. Theoretical Background.....	7
2.1 Constructed Languages	8
2.1.1 <i>Defining Constructed Languages</i>	8
2.1.2 <i>Phonology of Constructed Languages</i>	14
2.1.3 <i>Dothraki: An Overview</i>	15
2.1.3.1 Dothraki Phonology	16
2.2 Contact Languages	18
2.2.1 <i>Defining Contact Languages</i>	18
2.2.2 <i>Mixed Languages: An Overview</i>	20
2.2.3 <i>Phonology of Mixed Languages</i>	23
2.3 Statement of Purpose.....	26
3. Methods.....	29
3.1 Corpus	29
3.2 Materials and Design.....	29
3.2.1 <i>Vowels</i>	30
3.2.2 <i>Consonants</i>	31
3.2.3 <i>Other Phonological Aspects</i>	32
3.3 Procedures and Analyses.....	33
4. Results.....	34
4.1 Identifying Conflict Sites	34
4.1.1 <i>The Phonology of Dothraki</i>	34
4.1.1.1 English Phonology	34
4.1.1.2 Spanish Phonology.....	35
4.1.1.3 Arabic Phonology	36
4.2 Acoustic Analyses	37
4.2.1 <i>Single Vowels</i>	37
4.2.1.1 Stressed Vowels	37
4.2.1.2 Unstressed Vowels.....	40
4.2.2 <i>Vowel sequences</i>	43
4.2.3 <i>Plosives</i>	45
4.2.4 <i>Fricatives</i>	46

4.2.5 <i>Liquids</i>	47
5. Discussion	49
5.1 The Resolution of Conflict Sites in Dothraki	49
5.2 Conflict Site Resolution: A Comparison between Dothraki and Mixed Languages.....	52
5.3 A Comparison between Dothraki and Mixed Language Phonology.....	53
5.4 Limitations of the Study	55
6. Conclusion	56
7. References	58
8. Appendices.....	66
Appendix A	66

0. Abstract

The study of constructed languages (or conlangs), i.e. languages that are consciously created by an individual or a group of people in a short time span, is a relatively new and upcoming field in the area of linguistics (Schreyer, 2021). Therefore, the current study aims to make a contribution to this field by examining a specific kind of constructed language, an artistic language, which is a constructed language created as a form of artistic expression (Canepari, 2018; Libert, 2018). These artistic languages are often made to mimic natural languages, begging the question of whether a comparison can be made between constructed and natural languages. Several studies have argued that artistic languages most closely resemble contact languages (Van Oostendorp, 2019; Schreyer, 2021), which is why the current study will investigate this more closely.

Specifically, the current study will compare one artistic language, Dothraki, with mixed languages, a subtype of contact languages, as they can be argued to resemble artistic languages most closely on the basis of their conscious creation (Mazzoli & Sippola, 2021; Van Oostendorp, 2019). Additionally, the current study will focus on a phonological comparison as there is a lack of research in the area of phonetics and phonology in both the field of constructed languages as well as that of mixed languages. Through the identification of phonemic conflict sites, i.e. sites where multiple sounds compete for a place in the language's phoneme inventory (Stewart & Meakins, 2021), and the acoustic examination of vowels and consonants, it was expected that Dothraki would resolve these conflict sites in the same way as mixed languages, through processes of assimilation, merging and overlapping categories. Contrary to expectations, a well-defined split in the source of the phonemes of the Dothraki sound system was found, which can be attributed to a radical difference in the development and context between artistic and mixed languages.

Keywords: constructed languages, conlangs, mixed languages, Dothraki, phonology

“Lekhmove tawak?”²: A phonological comparison between a constructed language, Dothraki, and mixed languages.

1. Introduction

As Versteegh (1993) stated: “it is not an exaggeration to say that most linguists feel that Esperanto, Ido, Volapük, etc., not being natural languages, do not belong to the domain of linguistics” (pp. 539-540). A few decades before, Martinet (1946, p. 37) expressed a similar sentiment, speaking of a prejudice that he experienced from other linguists against those who are interested and involved in the creation of languages. Fortunately, the study of artificial languages, referred to as interlinguistics, is an up-and-coming field in linguistics, receiving more and more serious academic attention in recent years (Sanders, 2020; Van Oostendorp, 2019). The current study is of the belief that languages of this kind do have a place in the field of linguistic study, being able to offer important insights into the human language capacity that might not be found through the examination of natural languages. That is why the current study is focussed on artificial languages, commonly referred to as constructed languages (conlangs), which can be defined as languages that are deliberately created by an individual or a small group of people with a specific goal in mind (Schreyer, 2021, p. 328).

Specifically, the current study aims to examine one kind of constructed language, an artistic language (to be explained in detail in section 2.1.1), in light of a natural language type, mixed languages, which can be argued to resemble constructed languages most closely in comparison to other natural language types (more on this in sections 2.1.1 and 2.2.1). In order to do this, a specific artistic language was chosen to examine as a case study. This language is called Dothraki, an example of which can be seen in the title of this study, and

² Dothraki for: “is a constructed language real?/ a real constructed language?”

will be further explained in section 2.1.3. The specific research questions and hypotheses of the study will be explained in detail in section 2.3, and by answering these questions, the study aims to fill a gap in the literature surrounding constructed languages, as well as mixed languages, by examining the aspect of linguistics that has been studied the least in both linguistic fields to this date, that is phonetics and phonology. Therefore, a comparison will be made between the phonology of an artistic language with that of mixed languages, and will not consider the morphosyntactic, lexical or any other properties of the languages in question.

1.1 Outline of the Study

The current study aims to reach its objectives through the following structure: the first chapter gives a general introduction to the study, why this topic was chosen and what it aims to achieve. The second chapter will provide an extensive theoretical background of the two language types involved, ending in a statement of purpose in which the specific research questions and corresponding hypotheses will be substantiated and explained in detail. Chapter 3 will then provide an overview of the chosen methods for the study, describing where the data comes from, how measurements were done, and how the data will be analysed. Following the methodology, chapter 4 will give an overview of the results of the different aspects of the analysis. In chapter 5, the discussion will take place that interprets the results of the current study and ties it back to previous research in the relevant linguistics fields. Finally, chapter 6 will provide a brief conclusion to the entire study and will give some directions for future research. The references used in this study can be found in chapter 7, and the appendices in chapter 8.

2. Theoretical Background

The current study will make a comparison between one type of language: constructed languages, and another type of language, mixed languages. In order to do this, both language types will have to be explained in detail. The focus will first be placed on constructed

languages in section 2.1. Subsection 2.1.1. will provide an explanation of what constructed languages are, the different kinds there are, and how they are developed. The section will end with a closer examination of artistic languages as this subtype of constructed languages has been chosen as the focus of this study, and the ways in which they may resemble natural languages. A link between artistic languages and mixed languages will be made and a gap in the literature will be identified, i.e. the phonological resemblance of artistic and natural languages. Subsection 2.1.2 will give an overview of the studies done so far in terms of the phonology of constructed languages, and thereby artistic languages, and subsection 2.1.3 will provide a description of one specific artistic language that the current study will examine: Dothraki.

The discussion will then shift in section 2.2 to the second type of language considered in this study: that of contact languages. First, section 2.2.1 will give a definition of contact languages, explaining what they are and how they are developed, as well as the different kinds of contact languages that are currently identified. Section 2.2.2 will examine one type of contact language further, that of mixed languages. This section will explain what they are, the different kinds of mixed languages, and theories regarding the development of mixed languages. As with constructed languages, a section on phonology is necessary for mixed languages, which will be provided in section 2.2.3. Lastly, section 2.3 will conclude the theoretical background by drawing a comparison between artistic language phonology and mixed language phonology and will give a statement of purpose.

2.1 Constructed Languages

2.1.1 Defining Constructed Languages

The human capacity for language is argued by some to be one of the most, if not the most important aspect that sets us apart from other species (Fitch, 2010, p. 5-6; Friederici, 2017, p. xi; Sanders, 2020, p. 7). While our natural languages have developed over a span of

thousands of years, human intelligence, and our need to control and manipulate natural processes, has given us the ability to interact with the process of language development, allowing people to start creating languages themselves (Sanders, 2020, p. 7). These languages are called constructed languages, “conlangs” for short, which is a term coined by Danish linguist Otto Jespersen during the early 20th century (Vandenberg, 2019, p. 7). Conlangs can be defined as languages of which all aspects of their typology are consciously created by a group of people or an individual with a specific goal in mind. Additionally, they are created in a relatively short amount of time in comparison to the development of natural languages (Destruel, 2016, p. 1; Schreyer, 2021, p. 328).

Despite the name and definition for these languages being relatively new, conlangs have been around for hundreds of years, with the first historical record of a constructed language stemming from the 12th century (Destruel, 2016, p. 3; Knežević, 2018, p. 10). This language, which was called “lingua ignota” (literally, ‘the unknown language’), was created and used during the 12th century by Benedictine abbess Hildegard of Bingen in some of her texts (Destruel, 2016, p. 3; Knežević, 2018, p. 10). This first record of a conlang was followed by evidence of numerous experiments with language from the Renaissance period onwards, but the interest in the creation of conlangs appeared to flourish during the second half of the 19th century (Knežević, 2018, p. 12).

Most conlangs from before the 19th century were little experiments, which functioned more as intellectual exercises instead of having a specific goal for which the language would be used. From the 19th century onwards, constructed languages became known to the wider public. Esperanto, for instance, which was created during the late 19th century, even resulted in something that can be considered a social movement (Forster, 1982, p. 1-11; Janton & Tonkin, 1993, p. 113-128; Knežević, 2018, p. 13). The language gained a wide following, both on an institutional level as well as an informal level, resulting in a community of

speakers, texts written in the language, and international congresses to speak about the language and its creation (Knežević, 2018, p. 13). Esperanto, and many other conlangs during this time, were created as auxiliary languages to function as a lingua franca to support international communication (Canepari, 2018; Sanders, 2020).

Aside from auxiliary languages, other types of conlangs were also developed, categorized by their purpose of creation (Adelman, 2014; Vandenberg, 2019). One of those categories of conlangs is referred to as engineered languages, which includes logical languages and philosophical languages such as Thomas More's utopian language and John Wilkins' philosophical language of which the latter was supposed to be a language that did not have the ambiguity of natural languages (Libert, 2018; Okrent, 2009, p. 22; Smith, 2017). Another category of conlangs, and one that has gained more attention in recent years, is the category of artistic languages (artlangs). These are languages created as a form of artistic expression, usually for novels, television, and film (Canepari, 2018; Libert, 2018). Gobbo (2017) makes a further distinction within the group of artistic languages by adding the category of Hollywood languages, artistic languages that are created for mainstream movies and television that often receive a large following as long as the respective movie or television series remains popular. The current study will focus on one of these Hollywood languages but will consider it from the category of artlangs in general as the public fanbase will not be part of this study.

Aside from differentiating between different kinds of conlangs through their goal of creation, another distinction can also be made based on the way the languages are formed. On the one hand, there are the *a posteriori* languages, which are created from existing languages by rearranging or picking aspects of one or more languages and putting them together to create a new language (Van Oostendorp, 2019). An example of an *a posteriori* language that is created from only one language is *Latino sine flexion*, which as the name suggests is Latin

without inflection in order to simplify the language (Adelman, 2014, p. 546). Examples of a posteriori languages that are created from multiple languages include Esperanto which is based on several Indo-European languages (Libert, 2018), and another international auxiliary language called Volapük which is based on English, French and German (Reagan, 2019, p. 211).

On the other hand, there are the a priori languages which are not based on existing languages and created from scratch (Van Oostendorp, 2019). A unique example is the language Solresol, which was developed by Jean François Sudre by using only musical vocal notes, and could both be spoken as well as played on instruments (Adelman, 2014, p. 546). While Solresol is an extreme case of an a priori language as it is not based on any aspects related to language, most a priori languages are meant to be similar to natural languages, which raises the question of to what extent an a priori conlang is actually influenced by natural languages. Libert (2018) and Van Oostendorp (2019) argue that even a priori conlangs are always influenced by existing languages to some extent, as they are created by people using language. Additionally, they argue that instead of a dichotomy between the two types, it can be better viewed as a scale, as there are many a posteriori languages with a priori elements and vice versa (Libert, 2018; Van Oostendorp, 2019).

As mentioned, many a priori languages are created to resemble the way natural languages are used. Especially artlangs created for works of fiction are created with this goal. According to Barnes and Van Heerden (2006), as authors of artlangs essentially create fictional native speakers, artlangs can be considered “‘virtual’ natural languages” (p. 103). Evidence of this can be found in Destruel’s (2016) study, where he examined several artlangs according to Greenberg’s (1963) Universals of Language. The Universals of Language were established by Greenberg (1963) after he set out to discover linguistic truths that are found in all natural languages. By evaluating a total of 30 languages, Greenberg (1963) created a list

of 45 universals, mostly based on lexical, syntactic, and morphological qualities of languages. By applying these universals to artlangs, Destruel (2016) found that languages such as Dothraki and Quenya, which are created as natural languages in a fictional world similar to our own, follow nearly all Universals of Language.

An example can be seen in Greenberg's first universal of grammar, which holds: "In declarative sentences with nominal subject and object, the dominant order is almost always one in which the subject precedes the object" (as quoted in Destruel, 2016, p. 32; Greenberg, 1963). This is confirmed in both Dothraki and Quenya as, for both languages, SVO (subject-verb-object) is the dominant word order. Of the fifteen universals that were relevant to Dothraki, the language meets all fifteen, which is of importance to the current study as it will focus on Dothraki.

In contrast, artlangs like Na'vi and Klingon, which were created as alien languages and were therefore not intended to be similar to natural languages, do break several language universals, as could be expected as they can be seen as attempts to test the limits of language creation (Destruel, 2016; Vandenberg, 2019). For example, Greenberg's second universal of grammar holds: "In languages with prepositions, the genitive almost always follows the governing noun, while in languages with postpositions it almost always precedes" (As quoted in Destruel, 2016, p. 34; Greenberg, 1963). This universal is disproved in Na'vi as the language has free variation, which causes both adpositions to occur.

The supposed similarity of some artlangs with natural languages, as seen in the confirmed language universals in artlangs such as Dothraki and Quenya, begs the questions of which natural language family or category these artlangs could belong to. One attempt has been made by Peak (2014) who made the argument that the artlang Dothraki could be sorted into the Nostratic superfamily based on the resemblance of phonological and lexical similarities between Dothraki and the languages that are part of the Nostratic superfamily.

Nonetheless, this is a rather broad generalization as the Nostratic superfamily is a large group that includes many Eurasian language families and therefore has a broad range of phonological and lexical properties that Dothraki could be compared to.

Instead of comparing an artlang to a natural language family, it might be better to compare artlangs to a natural language category (e.g., contact languages or lingua francas). From this perspective, Van Oostendorp (2019) argues that conlangs, and thereby also artlangs, can be considered extreme cases of language contact, as they are essentially created from a combination of different languages or influenced by different languages that come into contact with each other. Especially the a posteriori artlangs, which are deliberate combinations of several existing languages, can be considered contact languages, but even the a priori types that are meant to resemble natural human languages are influenced by existing languages to some extent simply because they are meant to resemble natural languages. Van Oostendorp (2019) further argues that the deliberate planning and the easily identifiable point of origin causes them to be of interest to the study of language contact as natural contact languages are less clear in these regards. Similarly, Schreyer (2021) points out that conlangs, and thereby artlangs, may resemble contact languages such as pidgins in the way that their vocabulary is often small, it has relatively simple grammar, and is used in specific contexts, but the difference lies in the reason for creation (p. 239). Where pidgins are naturally formed out of a communicative necessity (Jourdan, 2021, p. 364), conlangs are created intentionally. Therefore, another type of language resulting from language contact, mixed languages, may be more comparable to constructed languages as they differ from other contact languages due to the intentionality of their creation (Mazzoli & Sippola, 2021). The current study aims to examine to what extent an artlang resembles mixed languages in terms of phonology.

2.1.2 Phonology of Constructed Languages

As mentioned, artlangs that are created to mimic or appear as human languages (e.g., Dothraki and Quenya) can resemble natural languages by following Greenberg's Universals of Language (Destruel, 2016). Nonetheless, the universals considered in Destruel's (2016) study mostly concern themselves with grammar, syntax, and morphology. Research into the phonology of conlangs and artlangs in relation to natural languages appears to be lacking. That is why the current study will focus on the phonological aspect of artlangs.

So far only a few studies have examined phonological aspects of conlangs in general and even less studies have specifically focussed on artlangs. In terms of conlangs in general, Salas (2012) found that when they are spoken by actual speech communities or individuals, the phonology of the conlang can be influenced by a person's native language if some phonetic nuances of the conlang are not clearly defined. One possible cause for this is the way conlangs are created. Natural languages originate as spoken languages and eventually develop a written version, while conlangs are usually created and first seen by others in a written form that is only later developed into a spoken version, which allows for native language influences of the people attempting to speak it, especially before the phonetic details of the spoken version are fully defined (Salas, 2012). Similarly, Ng and Schwendiman (2017) examined the distribution of phonemes across the sound inventories of several conlangs and found that over half of the phonemes in a conlang's sound inventory are often shared with the native language of the conlang creator. Even Klingon, which is supposed to be an alien language that is consciously created to sound different from natural languages, still shares approximately 40% of its phonemes with English, its parent language.

In terms of artlangs specifically, the idea of sound symbolism appears to play a role in the creation of an artlang's sound system. According to Beinhoff (2015), who describes sound symbolism as the mental connections that are made when hearing certain sounds, the

phonological realisation of an artlang can have an influence on the characterization of the language and the culture of the fictional people the language is spoken by. Nonetheless, the associations with certain sounds can also be caused by sociocultural meaning, which is why sound symbolism is considered a contested theory in the field of linguistics (Beinhoff, 2015).

2.1.3 Dothraki: An Overview

The current study will focus on the artlang Dothraki for a phonological analysis and a comparison with mixed languages. Dothraki is a language created for and spoken in the television series *Game of Thrones*. While some words and phrases of the language were already present in the book series the television series is adapted from, it only became something that can be considered a full language for the television show because it needed full dialogue for the first time. The person eventually tasked with creating Dothraki was American linguist David J. Peterson, who is the founder of the Language Creation Society which organizes congresses on all matters that concern conlangs (Gobbo, 2017; Meluzzi, 2019). He worked on developing the language for two years before it was used in the show. At the end of the post-production of the first season of the show in 2011, the language had over 3000 words, over 600 audio files as a guide in pronunciation, and nearly 200 pages of descriptions about the Dothraki language and culture (Gobbo, 2017, p. 46).

According to Gobbo (2017), one of the main reasons the language became such a quick success is the way the language is connected to the diegetic culture of the group of people who speak it. In the context of the fictional world, it is spoken by a nomadic group of warriors called the Dothraki, whose culture centres around horses (Meluzzi, 2019). Consequently, their language also revolves around horses which can be seen in many of its key expressions. For example, what they call themselves, Dothraki, literally translates to ‘those who ride.’ Gobbo (2017, p. 46) attributes this to Peterson’s consideration of the Sapir-

Whorf Hypothesis, which holds that the language people speak can influence the way they view the world. Additionally, Peterson (2015b) explains that his aim was to create a language that mimics natural languages in order to make it appear as naturalistic as possible and thereby give it linguistic authenticity. He tried to do this by allowing irregularities in the language that all could be explained by a fictional evolutionary process of the language. An evolutionary process that is supposed to mimic the evolutionary process of natural languages (Peterson, 2015b, p. 15).

Another reason why Dothraki may have become such an international success, having developed its own fanbase who share forums and webpages to talk about and in the language, is that in linguistic terms, the language is relatively easy to learn (Gobbo, 2017, p. 46). The morphology of the language is relatively regular in comparison to many other languages as it does not have a lot of suppletion or allomorphy (Gobbo, 2017, p.46).

2.1.3.1 Dothraki Phonology. In terms of its phonology, Dothraki is somewhat more complicated from the perspective of an English-speaking audience as it contains several sounds that do not occur in any variety of English. According to Peterson (2015a), he added these sounds to make the language sound harsher and foreign from a Western perspective, which was the main assignment he was given in his task to develop the language. Even though Dothraki itself is not a written language but an exclusively spoken one, Peterson (2015a) indicates that he based the pronunciation of the language mostly on the way an English-speaking audience would pronounce the small number of words and phrases that were originally used in the books and added foreign consonants whenever the original spelling was distinctly non-English. The full phonetic inventory he eventually landed on can be seen in table 1, showing the consonant system, and table 2, showing the vowel system.

Table 1*Phonemic inventory of Dothraki: consonants (Peterson, 2015a, p.92).*

	Labial	Dental	Alveolar	Palatal	Velar	Uvular	Glottal
Stop		t, d			k, g	q	
Affricate				tʃ, dʒ			
Fricative	f, v	θ	s, z	ʃ, ʒ	x		h
Nasal	m	n	n	ɲ	ŋ	ɴ	
Glide				j	w		
Lateral		l					
Tap/Trill			ɾ, r				

Table 2*Phonemic inventory of Dothraki: vowels (Peterson, 2015a, p.92).*

	Front	Back
High	i	
Mid	e	o
Low		a

As can be seen from the tables, the phonetic inventory of Dothraki is similar to English, with only a few exceptions due to the elimination of the plosives [p] and [b], and the addition of phonetic elements from Arabic and Spanish to make it sound harsher (Peterson, 2015a; Vinodh, 2019). The Arabic influences on the language can be seen in the addition of consonants such as the velar consonant [x] and the uvular [q]. The Spanish influences can be seen in the use of a trilled [r], and the dental consonants, with [t], [d], and [l] being pronounced [t̪], [d̪], and [l̪]. For the realisation of the orthographic letter r, Dothraki uses either [r], or [r̪] depending on the situation. If a word starts with an r that is followed by a vowel, has an r at the end of the word, or has a double r, it is pronounced like the Spanish [r̪]. In all other instances, it is pronounced as a tapped [r] (Peterson, 2014, p. 15).

In terms of the vowels, Dothraki uses the cardinal pronunciations as in Spanish, and does not have [u]. Additionally, Spanish influences can be seen in the way Dothraki vowels are always pronounced separately even if they appear in a cluster of vowels (Peterson, 2015a; Vinodh, 2019). According to Vinodh (2019), while Dothraki consists of four vowels in terms

of orthography, its phonetic inventory actually has seven vowels with the addition of the following three: [ɑ], [ɛ], and [ɔ] which only appear after [q] due to vowel laxing.

In terms of stress, Peterson specifically designed the language to sound different from English (Peterson, 2015a). Where it is rare to have stress on the final syllable in English, all words in Dothraki that end with a consonant have final stress as it looks to the right of the word to determine the stress. If a word does not end in a consonant but the penultimate syllable is heavy, the penultimate syllable will be stressed. In all other instances, the first syllable is the one that is stressed.

2.2 Contact Languages

2.2.1 Defining Contact Languages

As mentioned, constructed languages can be considered of interest to the field of contact linguistics, which is why the current study will consider conlangs from a contact language perspective. Naturally, this field will first have to be defined before a comparison can be made. According to Bakker and Matras (2013) many languages could be considered contact languages in a way as most languages are influenced by their surrounding languages to some extent, but they define contact languages in the field of contact linguistics as “new languages that have emerged in situations in which the repertoires of languages available to the people in contact did not provide a sufficiently effective tool for communication” (p. 1). This definition indicates that not all situations of language contact will always lead to the development of a new contact language (Bakker & Matras, 2013). A somewhat different definition is given by Thomason (1997) who focusses more on the traceability of the contact language and defines a contact language as a language consisting of linguistic material that cannot be attributed to only a single language in terms of its source and thereby are not easily sorted into a single language family.

While the field of contact linguistics is relatively new, a shift can already be seen. According to Siemund (2008), interest used to lie mostly in analysing contact languages individually, whereas now it has shifted to a comparison of different contact languages across different contact situations. Even though every contact situation is different, a trend can be seen which can be considered a characteristic that sets apart contact languages from other languages. Thomason and Kaufman (1988) refer to this aspect as “broken transmission”, the phenomenon that parent language is not fully transmitted to the newer generations who instead create a new language variety through the influence of the language contact situation. Still, different kinds of contact situations can lead to the development of various kinds of contact languages. Usually, a distinction is made between two contact situations, the first being a situation where multiple languages are present, but bilingualism is limited. In the second contact situation, speakers do have full access to two or more languages (Bakker & Matras, 2013). Additionally, social factors such as the size of the linguistic group, the social status of the group, and language prestige are considered important for the outcome of contact situations (Siemund, 2008).

On the basis of the different contact situations and social factors mentioned, the field of contact linguistics usually distinguishes between the following three contact language types: pidgins, creoles, and mixed languages (Bakker & Matras, 2013; Thomason, 1997). The first two, pidgins and creoles, are generally accepted as two distinct types of contact languages, but some researchers are still hesitant to distinguish mixed languages as well (Bakker & Matras, 2013, p. 1). For example, a leading study on contact languages by Sebba (1997) does not consider mixed languages at all but only discusses pidgins and creoles. Many other studies also dismiss mixed languages as a contact language category, but in the past two decades, mixed languages have gotten more attention and are increasingly considered as a distinct type of language category within the group of contact languages (see Bakker &

Mous, 1994; Matras & Bakker, 2003; Mazzoli & Sippola, 2021; Velupillai, 2015). Naturally, there are also cases where a distinction is not as easily defined between pidgins, creoles, and mixed languages. These languages can be seen as borderline cases, having characteristics from two or all three of the different contact language types (Thomason, 1997). Nonetheless, mixed languages are the main focus of the current paper as they are expected to resemble artistic languages more closely, and will therefore be discussed in more detail in the following section.

2.2.2 Mixed Languages: An Overview

What sets mixed languages apart from other contact languages is that they do not arise out of a communicative need but rather as a conscious markers of identity (Mazzoli & Sippola, 2021, p. 1; Meakins, 2013, p.181; Velupillai, 2015, p. 69). They are usually the result of community bilingualism after some form of social upheaval that causes community members to want to set themselves apart from the rest, thereby forming a new social category with a new language (Mazzoli & Sippola, 2021). Regardless of the new social category, mixed languages are often spoken in the same communities as their parent languages with most speakers of mixed languages also having fluency in one or both of the parent languages the mixed language is created from (Meakins, 2013, pp. 200-201; Velupillai, 2015, p. 70). These mixed languages are referred to as symbiotic mixed languages (Smith, 2000). While mixed languages usually have a clear split ancestry with easily identifiable features of two sources languages, they should be considered as autonomous systems in their own right for several reasons (Velupillai, 2015, p.70). For example, mixed languages may follow a different pattern of change than their parent languages despite being present in the same linguistic community (Velupillai, 2015, p. 70).

Naturally, as most language contact situations are unique, several different types of mixed languages can be identified. The most common way to distinguish mixed languages is

by looking at which elements from their source languages they mix to create a new language. In terms of intertwined languages, which is one structural type of mixed language that combines two separate components of each source language into one whole, a distinction can be made between languages that have a divide in their grammar and lexicon and languages that have a divide between the nouns and their respective grammar rules on one side, and the verbs and their respective grammar rules on the other side (Meakins, 2013; Velupillai, 2015). Initially, the term was only used in reference to the first type, which are commonly called G-L mixed languages (Grammar-Lexicon) (Bakker, 2003). In this type, the lexicon usually comes from one of the source languages and the grammar from the other. Examples of this include the mixed languages Angloromani, which has a clear lexicon/grammar divide, and Media Lengua, a language that also has a clear divide and takes its lexicon from Spanish, and its grammar from Quechua (Velupillai, 2015). Currently, another type is also considered an intertwined mixed language, and is referred to as N-V mixed languages (Noun-Verb). Examples include Michif and Gurindji Kriol. Aside from intertwined mixed languages, there are also converted languages (Mazzoli & Sippola, 2021, p. 9; Velupillai, 2015, p. 75) which have the grammar and lexicon elements from one language but the syntactic from the other language. These can be referred to as F-S mixed languages (Form-Structure) (Velupillai, 2015, p. 75), or F-R mixed languages (Frame-Root) (Bakker, 2017, p. 221). An example of this type is Sri Lankan Malay.

While there is a general consensus on the division of different types of mixed languages, there is an ongoing debate about how mixed languages actually develop. One of the earliest approaches is the one by Bakker (1997), who proposed a theory of language intertwining. In contrast to earlier views on mixed languages, his approach is not based on codeswitching and borrowing process that commonly occur in language contact situations, but rather looks at it from a perspective of language genesis. He does not view it as the

shifting of one language into another, but rather as the combination and intertwining of elements from two languages that create a whole new language (Bakker, 1997, p. 210). According to Bakker (1997), this intertwining is influenced by a grammar/lexicon distinction as well as a bound/free distinction in terms of morphemes. The grammatical morphemes, which are considered bound, tend to come from one language, mostly the one that is most familiar to the speakers (Bakker, 1997, p. 211). The lexical morphemes, which are considered free, are usually taken from the other language.

An approach that is based on contact processes of codeswitching and borrowing is the one developed by Auer (1999). He created a model that shows language change through different stages from codeswitching at stage 1, to language mixing at stage 2, to a mixed language at stage 3. According to Auer, what eventually differentiates a mixed language to language mixing is the loss of alternational codeswitching, and that mixed languages have a lot less variation in their syntax in comparison to the ongoing process of language mixing. Additionally, Auer (1999) argues that speakers of mixed languages are not always speakers of the two source languages while that is the case in the process of language mixing (p. 321).

After Auer's (1999) model, a second approach to the development of mixed languages that is also based on codeswitching is proposed by Myers-Scotton (2003). She argues for a shift from insertional codeswitching to the development of a mixed language, which she illustrates through her Matrix Language Frame (MLF) model. The model distinguishes two language types, the Matrix Language (ML) that is used as a grammatical frame, and the language that provides the lexical material, which she refers to as the Embedded Language (EL) and is considered the weaker language. Myers-Scotton (2003) hypothesizes that when a shift takes place between a dominant and weaker language, i.e., the ML and the EL, a turnover can interrupt the language shift and causes it to stabilize into a mixed language that combines elements from the two sources languages.

Despite being the major approaches in the view on the development of mixed languages, there used to be a lack of empirical evidence to either confirm or deny these claims, especially for the approaches that are based on the transition from codeswitching to a fully formed and autonomous mixed language (Meakins, 2013, p. 212). Auer (1999) even stated that his approach, to which codeswitching processes are essential, is based on “plausible guesses rather than empirically based” (p. 324). For a time, it was uncertain whether evidence could even be found due to a lack of empirical data from the start of the development of mixed languages (Backus, 2003, p. 241).

Nonetheless, in the last two decades, concrete evidence has been found that codeswitching practices are at the core of the emergence of several mixed languages in Australia. One of such languages is Gurindji Kriol. Through their diachronic examination of Gurindji speech McConvell and Meakins (2005) demonstrated documented evidence does exist for the transition from codeswitching to a mixed language. They looked at adult speakers of Gurindji in the seventies, Gurindji children speaking a mixed variety in the eighties, and Gurindji Kriol speech of middle aged and young speakers in the nineties and found clear evidence that codeswitching is at the core of the emergence of Gurindji Kriol as a mixed language. Since then, similar evidence was found for the mixed language Light Warlpiri (see O’Shannessy, 2011; 2020).

2.2.3 Phonology of Mixed Languages

So far, most research into mixed languages has concerned itself with lexical and morphological aspects as those are the aspects that most clearly set mixed languages apart from other contact languages. The fusion of these aspects in from different languages into one languages begs the question of how the phonological systems of the two languages are combined, in a stratified way or combined through processes of assimilation (Meakins, 2013, p. 209). While the consequences of language mixing have not been investigated extensively

yet in terms of phonology, the past decade has seen an increase in studies that concern themselves with the phonological aspect of the mixed languages (Meakins, 2013, p. 209-210; Onosson & Stewart, 2021, p.2).

Despite concerning different mixed languages and different phonological aspects, these studies all have more or less the same starting point, which is the identification of phonemic conflict sites after which acoustic comparisons can be made (Stewart & Meakins, 2021, p. 58). They define these conflict sites as “areas of phonological convergence where two or more sounds compete for a position in the phoneme inventory of a language” (Stewart & Meakins, 2021, p. 58). An example could be a language that includes only three vowels such as /a/, /u/, and /i/ and is the base for most of the mixed language’s phonology, but which is combined with a language that has more vowels, the same three plus /e/ and /o/. /e/ and /o/ can be seen as conflict sites as the speakers of the language will have to find a solution for these sounds, for example through a process of assimilation that changes the sounds to become more similar to familiar vowel sounds (/a/, /u/, or /i/ in this case), or by completely replacing them with familiar sounds (Stewart & Meakins, 2021, p.58). The notion of “conflict site” does appear to be a term rooted in a Western or structural conception of language which many studies using this term seem to disregard. This begs the question of why it is assumed that multiple linguistic elements are in “conflict” at these sites instead of merely being a site of variation. Nonetheless, as the current study is based on existing research in the field of mixed languages, this debate is out of the scope of this study.

The earliest (and still ongoing) debate in terms of mixed language phonology concerns the phonological analysis of a specific mixed language, Michif, of which Papen (2003) and Bakker and Papen (1997) claim that it has a stratified system, meaning that the components of the different languages retain their own phonological properties. The other side of the debate argues against a split phonology of Michif. For example, Prichard and

Shwayder (2014) examined French and Cree vowels within the Michif vowel system, as well as the application of a distinctly French rule of liaison to the Cree elements of Michif and found for both aspects that a clear separation between the French and the Cree systems cannot be made.

The first to compare several mixed languages in terms of their phonology was Van Gijn's (2009) study in which he compared Michif, Media Lengua, Callahuaya, and Mednyj Aleut, which differ in mixed language type (i.e., N-V, G-L, or F-S mixed languages). He concluded that the way the phonological properties of the mixed languages are realised can be predicted based on the typology of the mixed language. Similar to Papen (2003) and Bakker and Papen (1997), Van Gijn (2009) argues that Michif has a stratified phonology, and attributes this to the Noun-Verb split of the language. As the elements from the two different languages are kept relatively separate, the French noun phrases keep their French phonology whereas the Cree verb phrases retain Cree phonology. In contrast, a language like Media Lengua, which is classified as a G-L language, is argued to have only one phonological system as the elements from the two different languages mix at the prosodic level within words. According to Van Gijn (2009) the mixed words conform to the phonology of the ancestral language, which means that the mix of Quichua and Spanish elements conform to the phonology of Quichua in the case of Media Lengua.

While Van Gijn's (2009) model serves as a starting point for a general theory regarding the phonology of mixed languages, the phonological situation of most mixed languages is more complex than having either a completely stratified system or a system that is based on the ancestral language. As Stewart (2018) argues, where Media Lengua is expected to retain the Quichua three-vowel system and assimilate the two extra Spanish vowels, this is not exactly what happens. Instead, speakers of Media Lengua appear to keep the two extra Spanish vowels to some extent (Stewart, 2014). Additionally, in the case of

Gurindji Kriol, a N-V mixed language, Van Gijn's model predicts a separation of Kriol voiced stops and Gurindji voiceless stops, but Jones and Meakins (2013) have shown through VOT (voice onset time) measurements that Kriol stops are actually assimilated to Gurindji stops.

Aside from the two studies mentioned above, an increasing number of studies that examine specific phonological aspects of different mixed languages reveal the complex nature of mixed language phonology, both for the supposed stratified systems of N-V languages as well as the supposed integrated systems of G-L languages (for VOT studies see Jones & Meakins, 2013; Stewart, 2018; for vowel studies see Jones et al., 2011; Onosson & Stewart, 2021; Rosen et al., 2020). Stewart and Meakins (2021) give an analysis of several mixed languages and conclude that a clear-cut division is indeed hard to find, instead they found systems that were only partially integrated such as with the Michif vowel system, overlapping categories or vowels that nearly merged in the vowel system of Media Lengua, and contrasts that only occur in certain positions and not in others such as the contrasts between the voicing of stops in Gurindji Kriol.

2.3 Statement of Purpose

The main focus of the current paper is to compare and contrast the phonological behaviour of an artlang that is supposed to mimic a natural language, with a natural language category, mixed languages. As has been discussed previously in section 2.1.1, even a priori artlangs are influenced by elements of existing languages simply because they are meant to resemble natural human languages (Van Oostendorp, 2019). This is why artlangs could be considered cases of language contact as they essentially mix certain elements from a variety of natural languages, either consciously in their creation, or unconsciously through the influence of the languages the creators themselves speak or have knowledge of (Van Oostendorp, 2019). The current study puts forward the hypothesis that artlangs may resemble

a specific variety of contact languages, mixed languages, due to the mixing of different languages into one new language and the intentionality of creation that can be seen in both artlangs and mixed languages (Destruel, 2016; Schreyer, 2021; Mazzoli & Sippola, 2021). Thus, the current study aims to examine to what extent an artlang resembles mixed languages, specifically by looking at the phonology of the artlang chosen as a case study, i.e. Dothraki.

By looking at the phonology, the current study aims to fill a gap in the literature of artistic languages, and thereby also constructed languages to some extent, as phonology is the aspect that is currently considered the least in the field of interlinguistics. In order to perform an in-depth analysis of several phonological aspects, the current study uses the artistic language Dothraki, because it is one of the artlangs that most closely resembles natural languages in terms of morphosyntax and lexicon, as proven by Destruel's (2016) application of Greenberg's (1963) Universals of Language. Therefore, the current study aims to achieve this goal by answering the following research question: To what extent does the phonology of Dothraki, an artistic language, resemble the ways in which the phonology of mixed languages is realised with respect to the source languages?

In terms of mixed language phonology, some studies can be found that believe the ancestral or source language provides most of the phonology for the mixed language, regardless of the mixed results that are found through several studies examining the phonology of mixed languages (see Stewart & Meakins, 2021; Van Gijn, 2009). Other studies have shown a completely stratified phonology based on which elements of the lexicon belong to which language (Papen, 2003; Van Gijn, 2009). Even though the phonology of mixed languages has been shown to be more complex than the dichotomy between a completely stratified phonology or a fully combined phonology based on the ancestral language (Stewart, 2018; Stewart & Meakins, 2021), the current study poses that in terms of phonology, English

can be seen as the ancestral or source language for Dothraki which is influenced by Spanish and Arabic phonology in different aspects of the phonological system. The main reasons for this decision are based on the fact that English is the first language of Dothraki's creator, and according to Peterson (2015a), he based Dothraki mostly on the English sound system to make it easier to pronounce for the English-speaking actors and audience.

Therefore, phonemic conflict sites can be identified and examined similar to the way studies on the phonology of mixed languages are set up (Stewart & Meakins, 2021). This allows for a more reliable comparison between the phonology of artlang and the phonology of mixed languages. Therefore, the following sub questions will help answer the main question:

1. How are the resolutions of conflict sites (vowels, consonants) in the Dothraki sound system resolved in relation to its source languages (English, Spanish, Arabic)?
2. Is the way in which conflict sites are resolved in the Dothraki sound system similar to how they are resolved in mixed languages?

As mentioned, in the literature of mixed languages a dichotomy is often made between a lexically stratified phonological system or a phonological system integrated into the source languages (Stewart & Meakins, 2021). The former theory may not be applicable to an artistic languages such as Dothraki as the lexical items of the language are completely invented and do not derive their phonological shape from any source language. Thus, a lexically-based stratification would not be expected for Dothraki. Additionally, based on findings by Prichard and Shwayder (2014) that even stratified mixed language phonologies show integration and assimilation of the different elements, and the view that ancestral languages provide most of the mixed language's phonology (Stewart & Meakins, 2021; Van Gijn, 2009), it can be expected that the largest part of Dothraki phonology closely resembles

that of English, and that certain Spanish and Arabic elements are assimilated in some way.

This hypothesis is made from the consideration that English can be identified as the ancestral language of Dothraki.

3. Methods

3.1 Corpus

The current study is a corpus study, which uses a Dothraki language course as its main source for Dothraki audio. The course is called *Living Language Dothraki* and is created by the linguist who developed the language, David J. Peterson (2014). The corpus consists of over 60 minutes of audio that gives the pronunciation of about 250 Dothraki words and phrases. Additionally, a spoken conversation of approximately 4 minutes is included in the corpus as well. The course features three speakers, one of which is Peterson himself, a male speaker who was in his late thirties at the time of recording. He voiced the audio of all the word lists and half of the spoken conversation. The other half of the conversation is voiced by another unidentified male speaker, which did not give enough data to be used in the current study. The third speaker is an unidentified female whose data will not be used either in this study as only a small number of words and phrases are voiced by her. Due to timing constraints, only a selection of words and phrases of the language course was used. The full word and phrase list used in the current study can be found in appendix A.

3.2 Materials and Design

Stewart and Meakins (2021) claim that any study that wishes to investigate mixed language phonology must first identify “phonemic conflict sites”, which they define as “areas of phonological convergence where two or more sounds compete for a position in the phoneme inventory of a language” (p. 58) (for a more detailed description see section 2.2.3). These conflict sites can be resolved in a number of ways. Either one of the competing sounds is completely replaced. For example, in a situation where a conflict site occurs between /i/

and /e/, replacement would mean that all instances of /e/ will become /i/ or vice versa.

Another way of resolving a conflict site is through assimilation. In the same situation, that could mean that either /i/ or /e/ will change to sound more like the other, but will not completely become the other vowel.

Thus, finding conflict sites is the first step the current study will include in the research design. These conflict sites will be identified between the English and Arabic phonemic elements and the English and Spanish phonemic elements within Dothraki as English can be considered the ancestral language which the language sound inventory is mostly based on. Once the phonemic conflict sites have been identified, acoustic analyses as described in the following sections (3.2.1, 3.2.2) can be applied to the different types of phonemes.

3.2.1 Vowels

For the acoustic analysis of the vowels, all single vowels will be examined in terms of their first (F1) and second (F2) formant frequency measurements, with the exception of vowels after [q], in order to investigate the placement of the vowels in the vowel space similar to the case studies of mixed languages in Stewart and Meakins (2021), Rosen et al. (2020), and Onosson and Stewart (2021). The vowels after [q] are not examined in the current study as they are subject to a vowel laxing process and thus not pronounced in their cardinal positions. Additionally, not enough vowels after [q] could be measured for a reliable analysis.

The F1 is indicative of the height of the vowel, the lower the F1, the higher the vowel is pronounced in the vowel space. The F2 is indicative of the frontness/backness of the vowel, and the lower the F2, the more backed the vowel is pronounced in the vowel space. Thus, in case of assimilation, the F1 and F2 of the vowel would change in the direction of another vowel, and in case of replacement, the F1 and F2 of the two vowels will overlap.

Both stressed and unstressed vowels will be measured and analysed separately. A difference can be expected as English has a number of short lax vowels, especially in unstressed positions whereas, according to the phonology as described by Peterson (2015a), Dothraki only has long and tense vowels, even in unstressed positions.

Additionally, as the language includes double vowel which are pronounced separately, these vowels will be acoustically analysed as diphthongs to examine the movement from one vowel to the next one, similar to Onosson and Stewart (2021). In order to do this, the F1 and F2 will be measured for both vowels.

The vowel measurements for the single vowels will be compared to baselines of English and Spanish vowel frequencies as taken from Bradlow (1995), in order to see how closely the Dothraki vowels resemble either of these vowel systems, as these are the two phonological systems between which the expected phonemic conflict sites will occur. The vowel measurements for the double vowels cannot be compared to a Spanish baseline as no such baseline could be found in any previous studies. Instead, the vowels movement will be visualised in vowel charts in order to see whether a clear movement from one vowel to the next is visible or whether they are pronounced like single vowels or diphthongs. The Dothraki vowels will not be compared to an Arabic baseline as Peterson (2015a), indicated that he only used Spanish elements for the Dothraki vowels.

3.2.2 Consonants

In terms of consonants, voice onset time (VOT) measurements will be done for the plosives, similar to Stewart (2018) and the case studies described in Stewart and Meakins (2021). The VOT is the time between the release of the plosive, and the start of vocal fold vibration. VOT can be both positive, if the vibration starts after the release of the plosive, or negative if the vibration starts before the release of the plosive. The latter is often referred to as pre-voicing. The duration of the VOT of plosives is different for every language and can

therefore be used as a point of comparison. In case of assimilation, the VOT of one plosive will become either shorter or longer in order to resemble another plosive. In case of replacement, the VOT of both sounds will be approximately the same. In the current study, only word-initial VOT will be measured to make sure there is no interference from surrounding sounds.

As with the vowels, these measurements will be compared to an English baseline (Chodroff et al., 2015), and in this case an Arabic baseline as taken from Al-Ansari and Kulikov (2021), and a Spanish baseline as taken from Face and Menke (2020), as these are the two phonological systems between which phonemic conflict sites are expected to occur.

For the fricatives, the number of occurrences of the “right” Dothraki fricative [x] will be examined in comparison to the number of English fricatives or stops that might be produced instead. This approach is similar to Stewart (2020). The same will be done for the liquids such as the Spanish trilled [r], and the Spanish dental [l], in comparison to its English counterparts, [ɹ], [r], and [l].

3.2.3 Other Phonological Aspects

While other phonological aspects such as nasals, prosody, and stress are relevant for Dothraki, they will not be part of the current study. The reason for this is the lack of knowledge on which languages influenced these aspects. In his description of the Dothraki sound system, Peterson (2015a) is clear that Dothraki vowels, and liquids are influenced by Spanish and some Dothraki consonants, and fricatives, are influenced by Arabic and Spanish, with the remaining sounds coming from English. No such information was mentioned in terms of stress, prosody, or the nasals of the language. Therefore, these aspects are out of the scope of the current study.

3.3 Procedures and Analyses

The audio files will be cut and converted to WAV files through the audio editing program Audacity (Audacity Team, 2021). All acoustic measurements will be done manually through the program Praat (Boersma & Weenink, 2020). In terms of the F1 and F2, these measurements will be taken in the middle of the vowel, at 50% so that it will be the same for all vowels. In terms of the double vowels, both will be measured at 50% of their duration. Another way to measure formant frequencies would be at the highest peak of intensity of each vowel, but as these can differ greatly in terms of how far into the vowel it occurs, sometimes the highest peak is right at the start, or sometimes more towards the end, this could cause too much variation in the data. The mean of both the F1 and F2 for each type of vowel will then be plotted to give a visualization of the placement of these vowels, and the standard deviation will be given through the addition of an ellipsis. For the double vowels, the movement from the first to the second vowel will be visualized through a similar vowel plot using the mean F1 and F2 of each vowel.

In terms of the VOT measurements, the measurement will start at the release of the plosive and end when vocal fold vibration begins. This can be seen in the sound waves of the audio file as vocal fold vibration gives a regular periodic waveform. The mean VOT's for each different plosive can then be compared through a statistical analysis. All charts and figures will be created using the program R (R Core Team, 2020).

In terms of statistics, they will all be done through the program R (R Core Team, 2020) and the following statistical test will be used in this study: a one-sample t-test, which compares the gathered data from Dothraki with a sample or population means, in this case the Spanish, English and Arabic baseline values.

4. Results

4.1 Identifying Conflict Sites

4.1.1 The Phonology of Dothraki

As mentioned in section 2.1.3.1, the Dothraki sound system is mostly based on English, with some Spanish and Arabic influences (Peterson, 2015a). For a detailed description of the Dothraki sound system see section 2.1.3.1, but for an easier comparison with the sound systems of the languages it is influenced by, the consonant and vowel inventory will be repeated here:

Table 3

Phonemic inventory of Dothraki: consonants (Peterson, 2015a, p.92).

	Labial	Dental	Alveolar	Palatal	Velar	Uvular	Glottal
Stop		t, d			k, g	q	
Affricate				tʃ, dʒ			
Fricative	f, v	θ	s, z	ʃ, ʒ	x		h
Nasal	m	n	n	ɲ	ŋ	ɴ	
Glide				j	w		
Lateral		l					
Tap/Trill			ɾ, r				

Table 4

Phonemic inventory of Dothraki: vowels (Peterson, 2015a, p.92).

	Front	Back
High	i	
Mid	e	o
Low		a

Additionally, Dothraki has 3 allophone vowels due to a process of vowel laxing after [q]: [ɑ], [ɛ], and [ɔ].

4.1.1.1 English Phonology. As mentioned in section 2.1.3.1, most of the consonant inventory of Dothraki is taken from English, American English more specifically. In total, American English is made up of 25 consonants (table 5) and 13 single vowels (table 6). In addition to the single vowels, American English also has a number of diphthongs: [eɪ, oʊ, aɪ, oɪ, aʊ, iə, eə, əə] (Kretzschmar, 2008, p. 44).

Table 5

Phonemic inventory of American English: consonants (Kretzschmar, 2008, p. 48; Olive et al., 1993, p. 28).

	Bilabial	Labiodental	Dental	Alveolar	Palato-alveolar	Palatal	Velar	Glottal
Stop	p, b			t, d			k, g	
Affricate					tʃ, dʒ			
Fricative		f, v	θ, ð	s, z	ʃ, ʒ			h
Nasal	m			n			ŋ	
Glide	w					j		
Lateral				l				
Tap/Retroflex				ɾ/ɽ				

Table 6

Phonemic inventory of American English: single vowels (Kurath, 1977, p. 18-19).

	Front	Central	Back
High	i, ɪ		u, ʊ
Mid	e, ε	ə, ɜ	o, ɔ, ʌ
Low	æ		ɑ

Of all these phonemes, 16 English consonants overlap with Dothraki: [k, g, f, v, θ, s, z, ʃ, ʒ, h, m, n, ŋ, j, w, ɾ], and 3 English vowels overlap with Dothraki: [i, e, o].

4.1.1.2 Spanish Phonology. The Spanish phoneme system consists of 19 consonants and 5 vowels, as can be seen in table 7 and 8 respectively.

Table 7

Phonemic inventory of Spanish: consonants (Hualde et al., 2012, p.93).

	Bilabial	Labiodental	Interdental	Dental	Alveolar	Alveopalatal	Palatal	Velar
Stop	p, b			t, d			ɟ	k, g
Affricate						tʃ		
Fricative		f	θ		s			x
Nasal	m				ɲ		ɲ	
Lateral				l			ʎ	
Tap					ɾ			
Trill					r			

Table 8

Phonemic inventory of Spanish: vowels (Hualde et al., 2012, p. 90).

	Front	central	Back
High	i		u
Mid	e		o
Low		a	
	Unrounded		rounded

As mentioned in section 2.1.3.1., Dothraki takes its vowels from Spanish, with the exception of /u/. Additionally, vowels sequences of two vowels are pronounced separately as in Spanish. In terms of stop consonants, Dothraki [t] and [d] are pronounced dental as in Spanish. In terms of other consonants, Dothraki uses the Spanish trilled [r], and a dental [l]. This causes the following conflict sites between the English and Spanish influences in Dothraki in terms of consonants: [t] vs. [t̪], [d] vs. [d̪], [l] vs. [l̪], [ɹ] vs. [r], and [r]. In terms of vowels, possible conflict sites can be identified between all English vowels: [i, ɪ, u, ʊ, e, ε, ə, ɜ, o, ɔ, ʌ, æ, ɑ] and all Spanish vowels [i, e, a, o, u].

4.1.1.3 Arabic Phonology. The Arabic phoneme system consists of 30 consonants as can be seen in table 9.

Table 9

Phonemic inventory of Arabic: consonants (Amayreh, 2003, p. 518; Watson, 2007, p.19-20).

	Bilabial	Labiodental	Dental	Alveodental	Palatal	Velar	Uvular	Pharyngeal	Glottal
Stop	b			t, d t̪, d̪		k	q		ʔ
Affricate					ɟʒ				
Fricative		f	θ, ð ð̣	s, z ṣ	ʃ	x	χ, ʁ	ħ, ʕ	h
Nasal	m			n					
Lateral				l					
Tap/Trill				r/r̄					
Glide	w				j				

The Arabic vowels will not be considered as they are not used in the Dothraki. Even in terms of consonants, only a small number of Arabic sounds are used in the Dothraki sound system, as explained in section 2.1.3.1. Dothraki takes the voiceless uvular stop [q] from Arabic, and the voiceless velar fricative [x]. Both of these sounds are unfamiliar in English and can therefore be identified as conflict sites with the English [k] and [g] as these sounds are most similar in terms of pronunciation and placement in the vocal tract.

4.2 Acoustic Analyses

4.2.1 Single Vowels

4.2.1.1 Stressed Vowels. The mean first and second formant frequencies for four Dothraki stressed vowels can be seen in table 10 below. Additionally, the total number of formant measurements is given per vowel as well as the standard deviation between brackets. The mean formant frequencies of the Spanish and English baseline are given for the same vowels as Dothraki as well as additional vowels that appear in the Spanish and English vowel system. The number for the Spanish and English baseline is 1 as it is a sample means. A visualisation can be seen in figure 1 below where each colour corresponds to a different speaker.

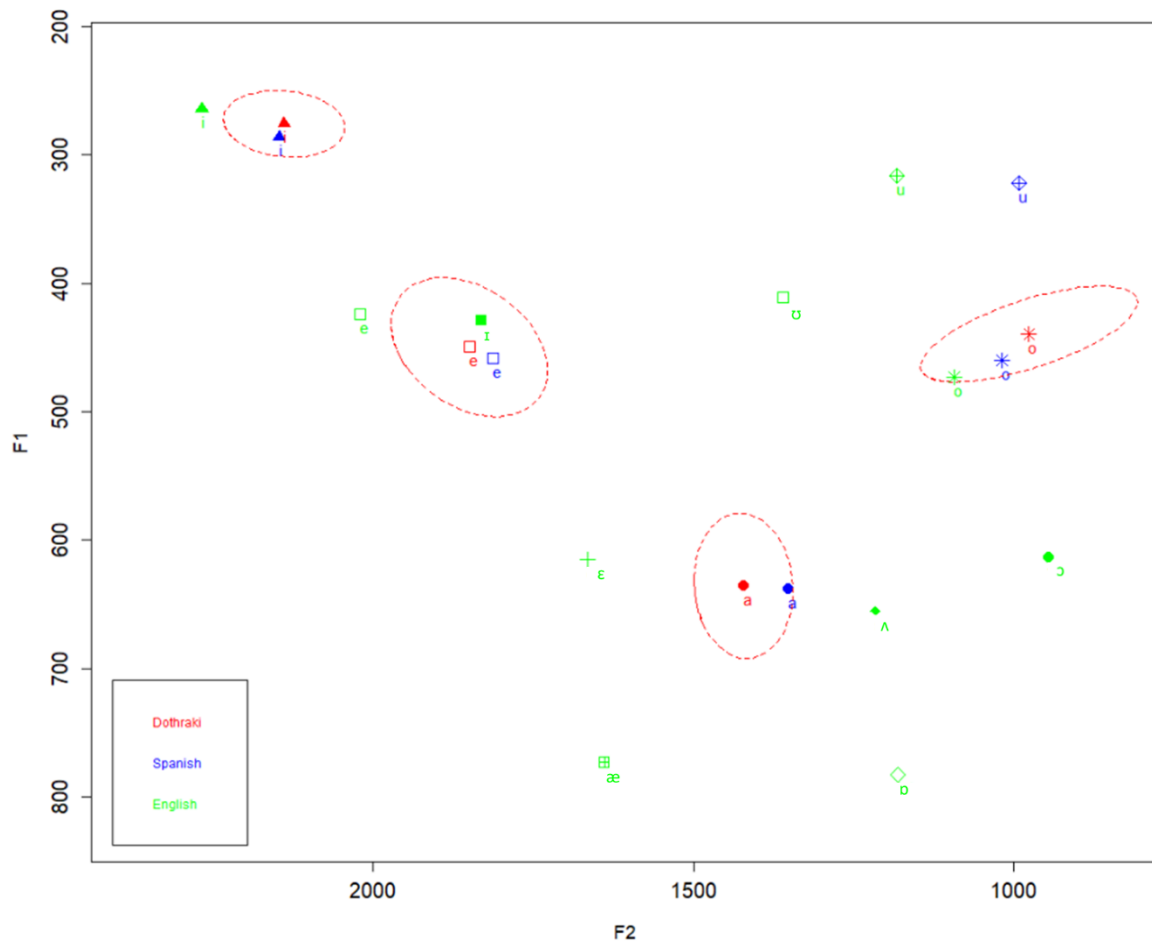
Table 10

Mean Formant Frequencies Stressed Dothraki Vowels with Standard Deviation between Brackets, and Spanish and English Baseline from Bradlow (1995).

Vowel	Dothraki			Spanish			English		
	N	F1	F2	N	F1	F2	N	F1	F2
a	160	636 (57)	1422 (77)	1	638	1353			
e	76	450 (54)	1850 (122)	1	458	1814	1	424	2020
i	60	275 (26)	2139 (95)	1	286	2147	1	264	2268
o	85	439 (37)	977 (169)	1	460	1019	1	473	1094
u				1	322	992	1	316	1183
ɪ							1	429	1831
ɛ							1	615	1665
æ							1	773	1640
ʌ							1	655	1216
ɒ							1	783	1182
ɔ							1	614	945
ʊ							1	411	1361

Figure 1

Mean Formant Frequencies Stressed Dothraki Vowels (P in red) with Standard Deviation and Spanish (in blue) and English (in green) Baseline from Bradlow (1995).



As can be seen from figure 1, three of the Dothraki vowels, [i], [e], and [o], are relatively close to their English counterparts. Nonetheless, when comparing the placement of the Dothraki vowels to their Spanish counterparts, all four Dothraki vowels appear to be produced in approximately the same location in the vowel space. With the exception of [i], the Dothraki vowels are not closely located to any of the additional short English vowels. In order to confirm whether the Spanish and Dothraki vowels can be considered the same, or if there is a significant difference a one-sample t-test was used.

For the vowel [a], a one-sample t-test revealed that there was no significant difference between the F1 of Dothraki ($M = 636$, $SD = 57$) and the Spanish baseline (638), $t(159) = -$

0.53, $p = 0.59$, 95% CI [626.75, 644.47]. The effect was of a small size, $r^2 = 0.002$. A one-sample t-test did reveal a significant difference between the F2 of Dothraki ($M = 1422$, $SD = 77$) and the Spanish baseline (1353), $t(159) = 11.37$, $p < 0.001$, 95% CI [1410.29, 1434.37]. The effect was of a large size, $r^2 = 0.45$. When comparing the Dothraki [a] to the surrounding English vowels [ɛ, æ, ɒ, ʌ], the difference was significant for the F1 ($p < 0.001$, $p < 0.001$, $p < 0.001$, and $p < 0.001$, respectively), as well as for the F2 ($p < 0.001$, $p < 0.001$, $p < 0.001$, and $p < 0.001$, respectively).

For the vowel [e], a one-sample t-test revealed no significant difference between the F1 of Dothraki ($M = 450$, $SD = 54$), and the Spanish baseline (458), $t(75) = -1.35$, $p = 0.18$, 95% CI [437.29, 462.00]. The effect was of a small size, $r^2 = 0.02$. A significant difference was revealed for the F2 of Dothraki ($M = 1850$, $SD = 122$), and the Spanish baseline (1814), $t(75) = 2.58$, $p = 0.012$, 95% CI [1822.27, 1878.08]. The effect was also of a small size, $r^2 = 0.08$. When comparing the Dothraki [e] to the English baseline [e], a significant difference was found for both the F1 ($p < 0.001$) and the F2 ($p < 0.001$). Similarly, when comparing the Dothraki [e] to the English [ɪ], a significant difference was found for the F1 ($p = 0.001$). No significant difference was found between the F2 of the Dothraki [e] and the English baseline [ɪ] ($p = 0.18$).

For the vowel [i], a one-sample t-test found a significant difference between the F1 of Dothraki ($M = 275$, $SD = 26$) and the Spanish baseline (286), $t(59) = -3.14$, $p = 0.003$, 95% CI [268.75, 282.18]. The effect was of a medium size, $r^2 = 0.14$. No significant difference was found between the F2 of the Dothraki [i] ($M = 2139$, $SD = 95$) and the Spanish baseline (2147), $t(59) = -0.61$, $p = 0.54$, 95% CI [2115.08, 2163.94]. The effect was of a small size, $r^2 = 0.01$. When comparing the Dothraki [i] to the surrounding English vowels [i, e, ɪ], a significant difference was found for all F1 ($p = 0.001$, $p < 0.001$, $p < 0.001$, respectively), as well as for all F2 ($p < 0.001$, $p < 0.001$, $p < 0.001$, respectively).

Finally, for the vowel [o], a one-sample t-test revealed a significant difference between the Dothraki F1 ($M = 439$, $SD = 37$) and the Spanish baseline (460), $t(84) = -5.07$, $p < 0.001$, 95% CI [431.35, 447.51]. The effect was of a medium size, $r^2 = 0.23$. Additionally, a one-sample t-test also revealed a significant difference between the Dothraki F2 ($M = 977$, $SD = 169$) and the Spanish baseline (1019), $t(84) = -2.31$, $p = 0.02$, 95% CI [940.05, 1013.09]. This effect was of a small size, $r^2 = 0.06$. When comparing the Dothraki [o] to the English baseline [o], a significant difference was found for both the F1 ($p < 0.001$) and the F2 ($p < 0.001$).

4.2.1.2 Unstressed Vowels. As with the stressed vowels, table 11 below shows the mean first and second formant frequencies with standard deviation for all four Dothraki vowels, as well as the number of measurements per vowel. Additionally, the mean formants for both the Spanish and English baselines are given for the same vowels as in Dothraki, as well as additional vowels that appear in the Spanish and English vowel systems. The number for the Spanish and English baseline is 1 as it is a sample means. A visualisation of these mean formants can be seen in figure 2 below, where each speaker corresponds to a different colour, the same colour as in figure 1 above.

Table 11

Mean Formant Frequencies Unstressed Dothraki Vowels with Standard Deviation between Brackets and Spanish and English Baseline from Bradlow (1995).

Vowel	Dothraki			Spanish			English		
	N	F1	F2	N	F1	F2	N	F1	F2
a	114	610 (67)	1472 (109)	1	638	1353			
e	46	458 (39)	1811 (126)	1	458	1814	1	424	2020
i	97	322 (31)	2046 (95)	1	286	2147	1	264	2268
o	58	472 (36)	1177 (155)	1	460	1019	1	473	1094
u				1	322	992	1	316	1183
ɪ							1	429	1831
ɛ							1	615	1665
æ							1	773	1640
ʌ							1	655	1216
ɒ							1	783	1182
ɔ							1	614	945
ʊ							1	411	1361

Figure 2

Mean Formant Frequencies Unstressed Dothraki Vowels (in red) with Standard Deviation and Spanish (in blue) and English (in green). Baseline from Bradlow (1995).

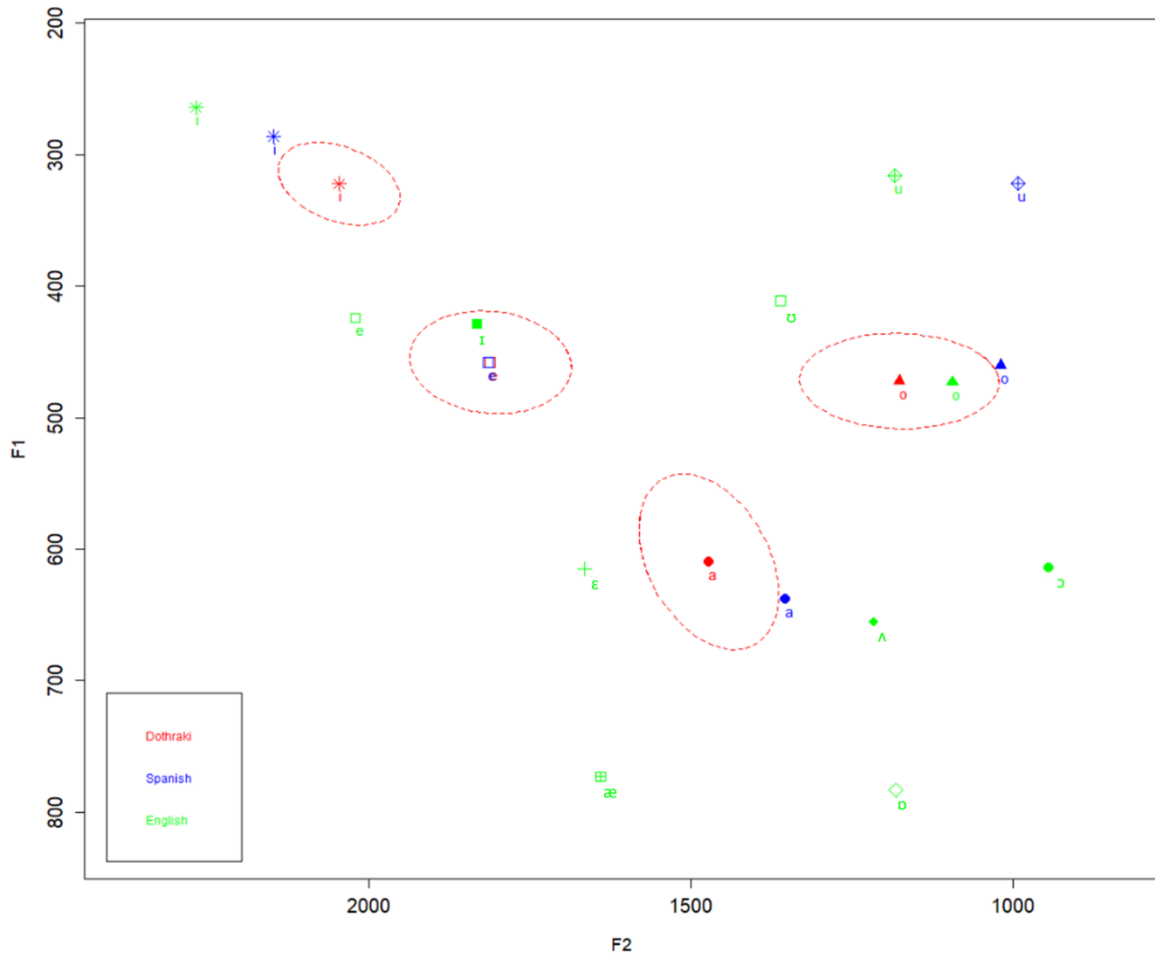


Figure 2 shows that even the unstressed Dothraki vowels are still relatively close to their Spanish counterparts, but not as close as the stressed vowels. Especially the unstressed Dothraki [i], [a], and [o], are further away from their Spanish counterparts as compared to the stressed Dothraki vowels in figure 1. As with the stressed vowels, none of the Dothraki vowels, with the exception of [i] and [o], are close to any of the vowels of the English sound system. In order to confirm whether the Spanish and Dothraki vowels can be considered the same, or if there is a significant difference a one-sample t-test was used.

For the [a] vowel, a one-sample t-test revealed a significant difference between the F1 of the Dothraki F1 ($M = 610$, $SD = 67$) and the Spanish baseline (638), $t(113) = -4.55$, $p <$

0.001, 95% CI [597.11, 621.91]. The effect was of a medium size, $r^2 = 0.15$. Similarly, a significant difference was found between the Dothraki F2 ($M = 1472$, $SD = 109$) and the Spanish baseline (1353), $t(113) = 11.70$, $p < 0.001$, 95% CI [1451.88, 1492.18]. The effect was of a large size, $r^2 = 0.55$. When comparing the F1 of the Dothraki [a] with the surrounding English vowels, a significant difference can be found for [æ, ɒ, ʌ] ($p < 0.001$, $p < 0.001$, $p < 0.001$, respectively). No significant difference was found the F1 of the Dothraki [a] and the F1 of the English [ɛ] ($p = 0.38$). The one-sample t-tests also revealed a significant difference between the F2 of the Dothraki [a] and the surrounding English vowels [ɛ, æ, ɒ, ʌ] ($p < 0.001$, $p < 0.001$, $p < 0.001$, and $p < 0.001$ respectively).

For the [e] vowel, a one-sample t-test revealed no significant difference between Dothraki F1 ($M = 458$, $SD = 39$) and the Spanish baseline (458), $t(45) = -0.03$, $p = 0.97$, 95% CI [446.21, 469.40]. Similarly, no significant difference was found between the F2 of Dothraki ($M = 1811$, $SD = 126$) and the Spanish baseline (1814), $t(45) = -0.17$, $p = 0.87$, 95% CI [1773.57, 1848.13]. Both effect sizes were small with $r^2 < 0.01$. When comparing the Dothraki [e] to the English baseline [e], a significant difference was found for both the F1 ($p < 0.001$) and the F2 ($p < 0.001$). Similarly, when comparing the Dothraki [e] to the English [ɪ], a significant difference was found for the F1 ($p = 0.001$). No significant difference was found between the F2 of the Dothraki [e] and the English baseline [ɪ] ($p = 0.28$).

For the vowel [i], a one-sample t-test found a significant difference between the F1 of Dothraki ($M = 322$, $SD = 31$) and the Spanish baseline (286), $t(96) = -11.29$, $p < 0.001$, 95% CI [315.68, 328.35]. The effect was of a large size, $r^2 = 0.57$. A significant difference was also found between the F2 of the Dothraki [i] ($M = 2046$, $SD = 95$) and the Spanish baseline (2147), $t(96) = -10.54$, $p < 0.001$, 95% CI [2026.82, 2064.93]. The effect was of a large size, $r^2 = 0.54$. When comparing the Dothraki [i] to the surrounding English vowels [i, e, ɪ], a

significant difference was found for all F1 ($p = 0.001$, $p < 0.001$, $p < 0.001$, respectively), as well as for all F2 ($p < 0.001$, $p < 0.001$, $p = 0.008$, respectively).

Lastly, for the vowel [o], a one-sample t-test revealed a significant difference between the Dothraki F1 ($M = 472$, $SD = 36$) and the Spanish baseline (460), $t(57) = 2.52$, $p = 0.014$, 95% CI [462.46, 481.49]. The effect was of a medium size, $r^2 = 0.10$. Additionally, a one-sample t-test also revealed a significant difference between the Dothraki F2 ($M = 1177$, $SD = 155$) and the Spanish baseline (1019), $t(57) = -7.76$, $p < 0.001$, 95% CI [1135.85, 1217.18]. This effect was of a large size, $r^2 = 0.51$. When comparing the Dothraki [o] to the English baseline [o], no significant difference was found to the F1 ($p = 0.83$), whereas a significant difference was found for the F2 ($p < 0.001$). A significant difference was also found between the F1 and F2 of the Dothraki [o] and other surrounding English vowels [ɔ, ʊ] with all p -values being $p < 0.001$.

4.2.2 Vowel sequences

In table 12 below, the mean first and second formant frequencies for the first and second vowel of several Dothraki vowel sequences can be seen. Additionally, the standard deviation is provided as well. The movement from the first to the second vowel in each vowel sequence is visualised in figure 3-6 below, with figure 3 showing all vowel sequences starting with [a], figure 4 showing the vowel sequences starting with [e], figure 5 showing the vowel sequences starting with [i], and figure 6 showing all vowel sequences starting with [o].

Table 12

Mean Formant Frequencies of the First and Second Vowel in Dothraki Vowel Sequences with Standard Deviation between Brackets.

Vowel Sequence	First vowel		Second vowel		
	N	F1	F2	F1	F2
aa	14	682 (88)	1512 (69)	705 (62)	1556 (52)
ae	22	621 (82)	1478 (97)	484 (48)	1894 (114)
ao	3	665 (19)	1271 (4)	566 (54)	930 (54)
ai	1	685 (NA)	1416 (NA)	363 (NA)	2094 (NA)
ee	9	412 (33)	1910 (74)	377 (20)	1940 (107)
ea	10	454 (57)	1854 (117)	686 (98)	1548 (101)
eo	6	423 (35)	1753 (112)	568 (42)	1068 (78)
ei	2	461 (18)	1895 (3)	312 (40)	2136 (0.5)
ie	3	356 (12)	1947 (60)	531 (40)	1706 (84)
ia	1	377 (NA)	2016 (NA)	696 (NA)	1489 (NA)
io	2	311 (0.7)	1868 (65)	447 (51)	1279 (124)
oo	15	485 (58)	1055 (466)	484 (89)	1019 (130)
oa	9	474 (42)	971 (136)	692 (49)	1178 (113)

Figure 3

Mean Formant Frequencies and Movement of Dothraki Vowel Sequences Beginning with [a].

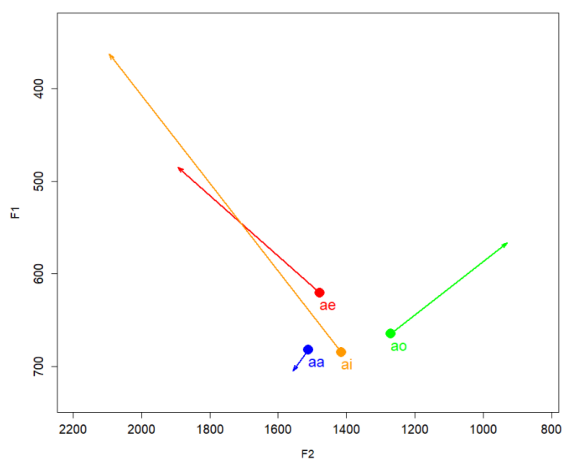


Figure 4

Mean Formant Frequencies and Movement of Dothraki Vowel Sequences Beginning with [e].

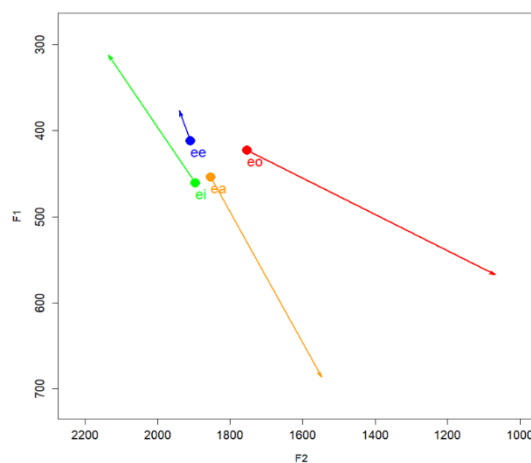
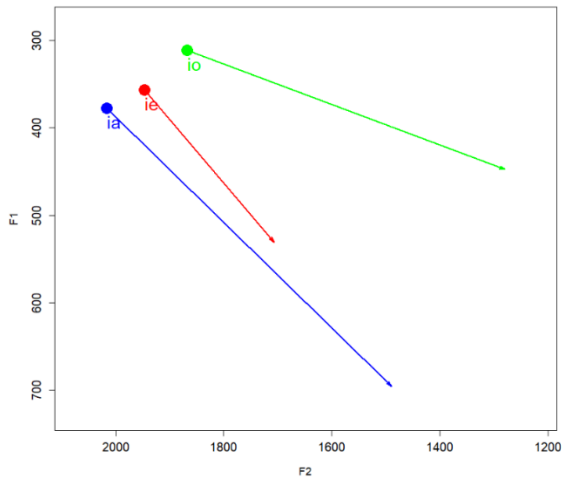
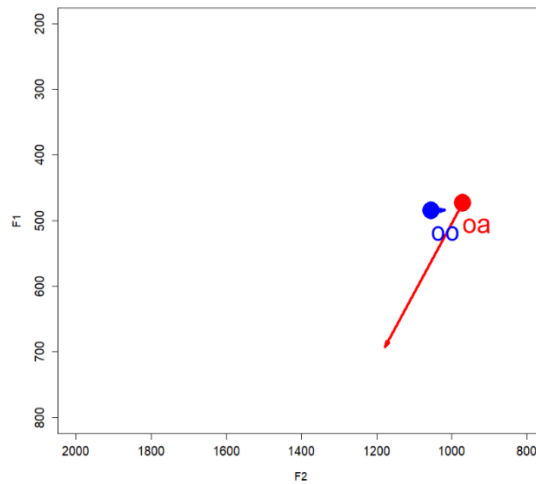


Figure 5

Mean Formant Frequencies and Movement of Dothraki Vowel Sequences Beginning with [i].

**Figure 6**

Mean Formant Frequencies and Movement of Dothraki Vowel Sequences Beginning with [o].



As can be seen from figures 3-6, Dothraki vowel sequences are clearly pronounced separately, moving from one vowel to the next. All vowel sequences starting with the same vowel start in approximately the same place in the vowel space and then move to the following vowel. Additionally, a small movement can be seen in the vowel sequences with the same vowel ([aa, ee, oo]), indicating that these are pronounced separately as well instead of as a single vowel [a, e, o].

4.2.3 Plosives

In table 13 below, the mean VOT can be found for the three Dothraki plosives, as well as the standard deviation between brackets. The Spanish, English, and Arabic baselines are given as well for the same as well as additional plosives in order to compare the Dothraki VOT's to those of its source languages.

Table 13

Mean VOT and Standard Deviation between Brackets for Dothraki Consonants, Spanish Baseline (Face & Menke, 2020), English Baseline (Chodroff et al., 2015), and Arabic Baseline (Al-Ansari & Kulikov, 2021).

Consonant	VOT							
	N	Dothraki	N	Spanish	N	English	N	Arabic
d	39	-69.3 (28.6)	1	-70.9	1	13.8		
t	13	17.3 (11.1)	1	18.7	1	60.5		
q	16	28.1 (13.9)					1	30.0
k					1	54.4		
g					1	17.2		

As table 13 shows, the VOT for the Dothraki [d] is relatively close to the Spanish [d]. Both show prevoicing, which is not seen in the English [d]. As for [t], the Dothraki VOT is also similar to the Spanish [t], and is lower than the VOT of an English [t], which has aspiration. The Dothraki [q] appears relatively similar to the Arabic [q], and not to either the English [k] or [g]. In order to confirm whether the Dothraki VOT's and the corresponding Spanish and Arabic VOT's can be considered the same, a one-sample t-test was used.

For the consonant [d], a one-sample t-test revealed that there is no significant difference between the VOT of the Dothraki [d] ($M = -69.3$, $SD = 28.6$) and the VOT of the Spanish [d] baseline (-70.9), $t(38) = 0.35$, $p = 0.73$, 95% CI $[-78.58, -60.02]$. The effect was of a small size, $r^2 = 0.003$. Similarly, a one-sample t-test also revealed no significant difference between the VOT of the Dothraki [t] ($M = 17.3$, $SD = 11.1$) and the Spanish [t] baseline (18.7), $t(12) = -0.47$, $p = 0.65$, 95% CI $[10.54, 23.99]$. The effect was of a small size, $r^2 = 0.02$. Lastly, a one-sample t-test was used to compare the Dothraki [q] ($M = 28.1$, $SD = 13.9$) and the Arabic [q] baseline (30.0), and no significant difference was found, $t(15) = -0.55$, $p = 0.59$, 95% CI $[20.68, 35.50]$. The effect was of a small size, $r^2 = 0.02$.

4.2.4 Fricatives

For the fricative [x], a conflict site was expected between the Arabic [x] and the English [k] and [g] as they are the English sounds that most closely resemble the placement of [x] in the vocal tract and the sound that it makes. A total of 41 instances were counted in

which the fricative is expected, and 39 of those were realised by [x] instead of a different sound that might come from English such as [k] or [g]. 2 instances where [x] would be expected according to the rules of the Dothraki language, a [k] was used instead.

Nonetheless, influences from the usage of the language in the television series it was made for may be of influence here, but this will be expanded upon in chapter 5.

4.2.5 Liquids

In terms of the liquid [l], which had a conflict site between the English [l] and the Spanish dental [ɫ] a similar result can be found as with the fricatives. In total, 56 instances were counted in which the dental [ɫ] is expected, and in all 56 of those instances, it was realised as [l].

For the conflict site of [r], three options were possible: the Spanish trilled [r], the English tapped [ɾ], or the English retroflex [ɻ]. The difference between the three is illustrated in figures 7, 8 and 9 below. In figure 7, the trilled [r] is visualised and can be recognized by its several intensity peaks and the corresponding shifts in formant frequencies. In figure 8, the tapped [ɾ] is shown, which is shorter than the trilled [r] and is less voiced. Figure 9 shows the retroflex [ɻ], which, unlike the trilled [r], is a continuous sound with the same formant frequencies and intensity throughout.

Figure 7

Spectrogram and Sound Wave of the Dothraki Word “rai” Realised as [rai].

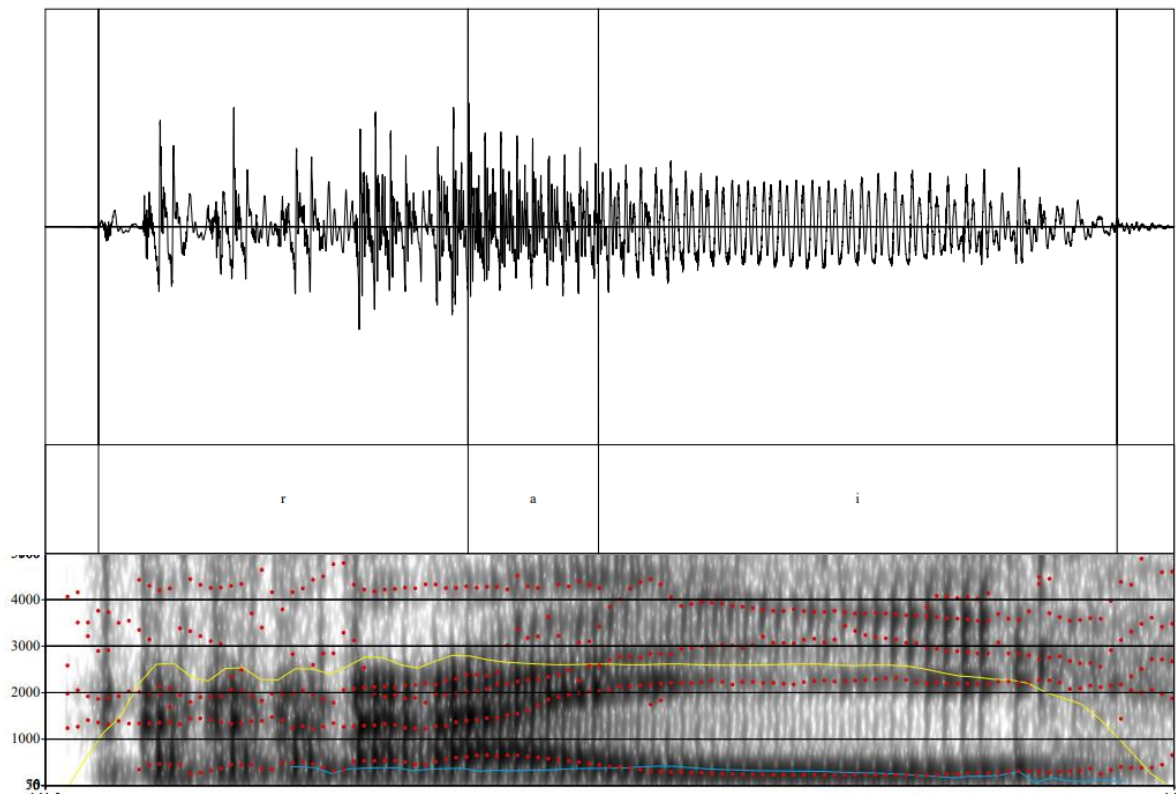


Figure 8

Spectrogram and Sound Wave of the Dothraki Word “mori” Realised as [mori].

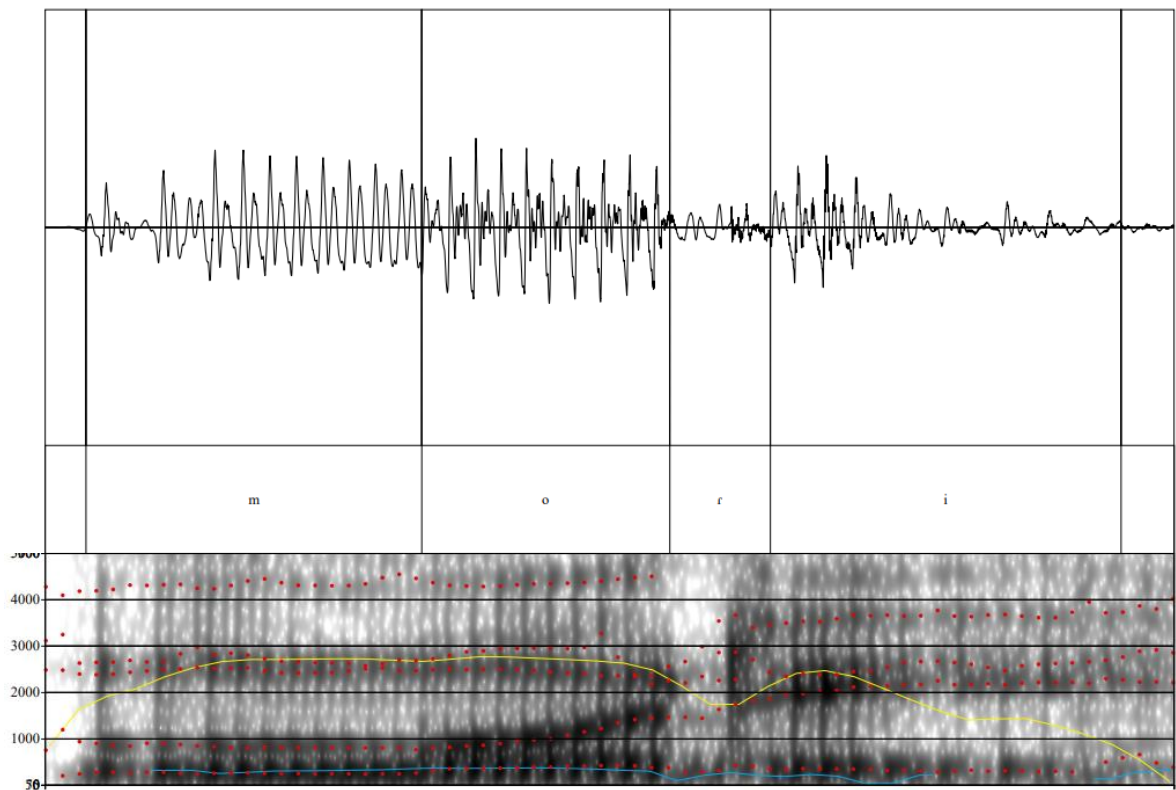
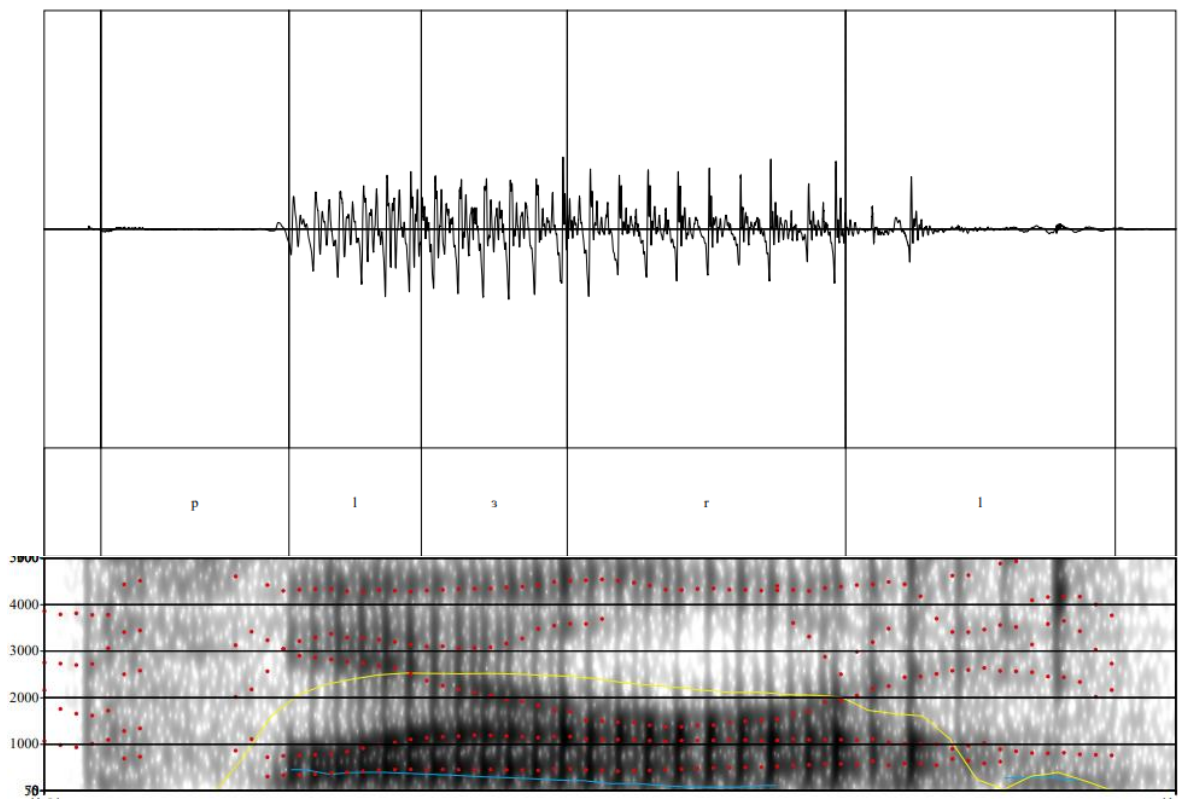


Figure 9

Spectrogram and Sound Wave of the English Word “plural” Realised as [plɜ:l].



In total, 105 words with an orthographic r were counted, of which 64 should be realised as a tapped [ɾ], according to the rules of the Dothraki language as explained in section 2.1.3.1, and all these 64 instances were realised as a tapped [ɾ]. Of the 105 words, the trilled [r] would be expected in 41 instances according to the rules of the Dothraki language, and all 41 instances were realised as the trilled [r]. None of the words had the English retroflex [ɻ].

5. Discussion

5.1 The Resolution of Conflict Sites in Dothraki

In order to answer the main question of this study, two sub questions have to be considered and answered first. The first of these sub question regarded the way conflict sites are resolved in the artlang Dothraki: How are the resolutions of conflict sites (vowels,

consonants) in the Dothraki sound system resolved in relation to its source languages (English, Spanish, Arabic)? Based on arguments brought forth by Libert (2018) and Van Oostendorp (2019) that regardless of whether a conlang is created as an a priori or an a posteriori language, it will always be based on or influenced by natural languages to some extent, and findings by Ng and Schwendiman (2017) that the largest part of a conlang's phonology is often shared with the native language of its creator, the current study expected that the resolution of conflict sites in Dothraki would be mostly based assimilation towards its main source language, English. This means that the Arabic and Spanish elements that Peterson (2015a) included in the sound system of English were expected to assimilate to English to a certain extent.

The results of the current study go against the hypothesis of assimilation as they reveal a split sound system. It is most clearly seen in the examination of the Dothraki consonants. A statistical test revealed that all three plosives [t, d, q] for which the VOTs were measured, were not statistically different from the VOT values of the corresponding baselines of Spanish and Arabic, while they were different from the realisation of these consonants in American English (see section 4.2.3). Additionally, the examination of the fricatives and liquids can also be said to reveal a complete split between the English, Arabic, and Spanish elements of Dothraki. For the liquids, orthographically presented with “r” and “l”, all occurrences of these sounds were realised by the Spanish pronunciation, where this was applicable (i.e., “r” can be realised by a tapped [ɾ], which comes from American English) (see section 4.2.5). The fricative [x], orthographically presented as “kh”, was the only sound for which not all instances were realised as the Arabic [x], with two occurrences being realised as [k] (see section 4.2.4). Nonetheless, this was a conscious addition of the language creator due to the mistakes that were made in the television series where two words “khal” (king/ leader of a group of Dothraki), and “khaleesi” (queen/ wife of a khal) are consistently

mispronounced (Peterson, 2014, p. 87; 2015a, p. 93). Therefore, the results of the fricative analysis still point to a stratified phonology.

The results of the vowels are somewhat more complex than that of the consonants. Purely based on the visualisations of the single vowel formant frequencies, the four Dothraki vowels appear to be realised the same as in Spanish, especially the stressed vowels are clearly placed in the cardinal positions of the vowel space (see section 4.2.1). Nonetheless, a statistical analysis only confirmed no statistical difference between the Dothraki and Spanish vowels for either the F1 or the F2 for most of the vowels. This might be caused by the method of the study, as the Spanish and English baselines are based on measurements from different studies, which can cause variation in terms of the way things are measured and the audio quality of the recordings. Still, the vowel placement of Dothraki does approximate that of Spanish, and is not similar to that of English in terms of the stressed vowels. In terms of the unstressed vowels, a similar result can be found with a perceived similarity between Dothraki and Spanish based on the visuals, but which is only statistically substantiated for either the F1 or F2 depending on the vowel, with one vowel [o] having no statistical similarity with Spanish for either formant frequency (see section 4.2.1). However, a possible cause could be the use of the baselines, which are not specifically based on unstressed vowels and may therefore not be fully reliable to use in a comparison, but which was necessary as no baseline based on unstressed vowels could be found. Lastly, the visualization of the double vowels and their movement from one vowel to the next appear to be realised similar to how vowel sequences are realised in Spanish, with both vowels in the sequence pronounced as two single vowels (see section 4.2.2).

Thus, contrary to predictions, no clear assimilation of Spanish and Arabic towards English can be seen. Instead, the conflict sites appear to be resolved through complete separation of the three sound systems used for Dothraki. From this, it appears that despite

being considered an a priori language by its creator (Peterson, 2015a), its sound system can be considered an a posteriori element as it specifically uses a combination of English, Spanish and Arabic elements that are kept separate in their pronunciation and do not influence each other through assimilation into new sounds. This does correspond to ideas expressed by Van Oostendorp (2019), who argues that there cannot be a complete dichotomy between a priori and a posteriori conlangs, and thereby also artlangs.

5.2 Conflict Site Resolution: A Comparison between Dothraki and Mixed Languages

The second sub question of the study regarded the comparison between the way conflict sites are resolved in Dothraki and mixed languages: Is the way in which conflict sites are resolved in the Dothraki sound system similar to how they are resolved in mixed languages? Based on findings in previous research that Dothraki can be analysed as a natural language (Destruel, 2016) and the idea that constructed languages can be considered cases of language contact (Van Oostendorp, 2019), the current study chose to compare Dothraki with mixed languages, as mixed languages are a type of contact language that most closely resemble constructed languages regarding the intentionality of their creation (Mazzoli & Sippola, 2017). Therefore, it was expected that conflict site resolution in Dothraki would be similar to conflict site resolution in mixed languages, meaning that assimilation to some extent was expected as studies have shown that even supposedly stratified sound systems of mixed languages show assimilation of one or both sound systems in some way, as well as partial integration or overlapping categories (see Stewart & Meakins, 2021).

As I have already explained in the previous section, Dothraki does appear to resolve conflict sites through a complete separation of the three sound systems it combines. This goes directly against the expectations of the current study, as well as the most prominent findings in the field of mixed language phonology that there is no clear dichotomy between complete assimilation to the ancestral language and complete stratification within mixed language

phonology (Jones & Meakins; 2013, Stewart, 2018; Stewart & Meakins, 2021). Additionally, while the split phonology of Dothraki resembles that of stratification in mixed languages, it cannot be considered the same, as stratification in mixed languages is lexically based, while the split in Dothraki is purely based on phonology without influence from lexical entries. Thus, while stratification to some extent can be seen in the conflict site resolution of mixed languages, it is not the same as the split phonology of Dothraki and is also not seen in the high level in which Dothraki's phonology appears to be split. A possible explanation can be found in the level of consciousness with which an artlang is created in comparison to a mixed language. As could be seen from Peterson's (2015a; 2015b) explanations of how he created Dothraki, he deliberately chose specific sounds that he wanted to include in the language alongside English. In contrast, while mixed languages are also consciously "created" to some extent as a marker of identity or to form a new social category (Meakins, 2013, pp. 200-201; Velupillai, 2015, p. 70), the first speakers of the mixed variety do not consciously choose which phonetic elements will be used in the mixed variety. Instead, as indicated by evidence found by McConvell and Meakins (2005) and O'Shannessy (2011; 2020) and previous ideas expressed by Auer (1999) and Myers-Scotton (2003), mixed languages arise from a process of codeswitching that mostly concerns itself with morphosyntactic elements, where the switching and mixing of phonological elements is a by-product of codeswitching instead of a conscious part of it.

5.3 A Comparison between Dothraki and Mixed Language Phonology

The main goal of the study was to make a phonological comparison between an artlang, in this case Dothraki, and mixed languages, a type of contact language. The study tried to achieve this goal by answering the following main research question: To what extent does the phonology of Dothraki, an artistic language, resemble the ways in which the phonology of mixed languages is realised with respect to the source languages? As has been

explained in the previous section, the phonology of Dothraki was expected to resemble the phonology of mixed languages through the identification of conflict sites and the way they are resolved. The results of the current study went against expectations as the Dothraki showed a level of split phonology not found in mixed languages, in which the resolution of conflict sites is more complex with processes of assimilation, merging, and overlapping categories (see Stewart & Meakins, 2021). Thus, while Dothraki does resemble the phonology of mixed languages in the way that it shows something that resembles stratification, it does not resemble the phonology of mixed languages in the way this stratification or split is present. The phonology of Dothraki thus appears to be less complicated than that of mixed languages, as the phonology of mixed languages is the by-product of several different processes that can be attributed to the languages' development based on morphosyntactic and lexical elements that are not present in the development of constructed languages.

As mentioned in the previous section, the level of consciousness in the creation of Dothraki as compared to the development of mixed languages may be the main factor for the difference in the realisation of their phonologies. While the current study agrees with Van Oostendorp's (2019) assessment that the deliberate planning of conlangs and thereby artlangs and their easily identifiable point of origin causes conlangs to be of interest to the field of contact linguistics (seeing constructed languages as extreme cases of language contact), current findings of this dissertation suggest that contact languages and constructed languages do not show consistent phonological similarities in the way conflict sites are resolved. The difference in the way both language types develop influences the way in which a comparison can be made. Constructed languages are created by a small group or a single individual (Schreyer, 2021), which allows for a lot less variation in the way the language is eventually developed, having less influences from many different speakers. In contrast, mixed languages

and contact languages in general arise in entire communities, which causes a lot more variation. This difference makes it harder to make a reliable comparison between both language types.

Nonetheless, the current study did succeed in making a contribution to the study of artistic languages, and thereby constructed languages in terms of phonology, an area that has not been examined that extensively yet. The finding that Dothraki has a split sound system may be of interest to studies in artlang phonology in the way that future research into different artistic or constructed languages could examine whether a similar level of split phonology can be found, or whether it is unique to this specific language, Dothraki, and its creator, David J. Peterson. Additionally, it can be considered as one of the first steps into the consideration of conlangs and artlangs in light of the natural languages that they are trying to mimic.

5.4 Limitations of the Study

In the previous sections of the discussion, a couple of limitations have already been touched upon briefly in terms of the method of the current study. While time constraints and the scope of the current study caused the current method to be chosen, the current method does have some limitations in terms of the comparison. As the comparison is based on baselines from three separate languages, originating from separate studies, the method can be considered somewhat less reliable as the way of measuring formant frequencies and VOTs might not be done in entirely the same way due to differences in audio quality, use of measurement programs, and a possible difference in manual versus programmed measurements. Therefore, the application of statistical analyses might be influenced as well.

Additionally, a limitation can be identified in the way that the vowel sequences of Dothraki could not be statistically compared to Spanish due to a lack of Spanish formant data for vowel sequences, which would be resolved by changing the method to include the

measurements of the additionally languages instead of relying on existing studies.

Unfortunately, this could not be done in the current study as time constraints already caused in a rather small dataset.

Finally, the way in which data from only a single speaker is used as a case study can be seen as a limitation. For future studies in Dothraki, it might be of interest to include the small speech community of Dothraki that is present among the fanbase of the show, *Game of Thrones*, that the language was created for as it was out of the scope of the current study but, based on findings by Salas (2012) (see section 2.1.2), is expected to offer vital insights in the ways the language is used in a more natural setting than recordings for a language course.

6. Conclusion

The current study sets out to examine to what extent an artlang, specifically Dothraki, can be said to resemble mixed languages in terms of phonology. By examining the way Dothraki resolves phonological conflict sites, the main findings of the study revealed that the split phonology of the Dothraki sound system is more clear-cut and extreme in a way not seen in mixed languages. Instead, mixed languages have been found to be subject to assimilation, merging, and overlap even in stratified phonologies, and no such elements were found in Dothraki. Despite not conforming to the initial expectations of the study, the findings have implications for the study of artistic, and thereby constructed, language phonology, as it can be seen as an important step in the consideration of constructed languages in the field of linguistics. Through the examination of a constructed language in the same way one might examine natural languages the current study showed that a comparison between a constructed language and a specific type of natural language cannot simply be made based on a few similarities as has been done in previous literature. Instead, the reality of far more complex and different processes in development can cause a radical difference between constructed languages and natural languages when examined empirically.

As the current study only considers one specific artistic language, future research could focus on the examination of additional artistic languages from a natural language perspective, as well as other types of constructed languages in order to see if similar results can be found in terms of their phonological realisations. Additionally, an interesting step forward would be to examine actual speech communities that use these constructed languages for communication, however small they might be, as it can offer vital insights in the way people use language. It might reveal aspects of language usage not found through the examination of natural languages. All in all, as the creators of constructed languages are pushing the boundaries of what people can do with languages, the field of linguistics can only follow by examining exactly how the boundaries of language are pushed and perhaps even crossed, making the field of interlinguistics an important area within linguistic research in general.

7. References

- Adelman, M. (2014). Constructed languages and copyright: A brief history and proposal for divorce. *Harvard Journal of Law and Technology*, 27, 544-562.
- Al-Ansari, N., & Kulikov, V. (2021). Acoustics of Arabic uvular and emphatic coronals: Evidence for uvularization of emphatic Qatari Arabic. *ResearchGate*, 349709767, 1-40. <https://doi.org/10.13140/RG.2.2.10863.56485/2>
- Amayreh, M.M. (2003). Completion of the consonant inventory of Arabic. *Journal of Speech, And Hearing Research*, 46, 517-529. [https://doi.org/10.1044/1092-4388\(2003/042\)](https://doi.org/10.1044/1092-4388(2003/042))
- Audacity Team. (2021). *Audacity* (Version 2.4.2). Muse Group.
<https://www.audacityteam.org/>
- Auer, P. (1999). From codeswitching via language mixing to fused lects: Toward a dynamic typology of bilingual speech. *Journal of Bilingualism*, 3(4). 309–332.
- Backus, A. (2003). Can a mixed language be conventionalized alternational codeswitching? In Y. Matras & P. Bakker (Eds.), *The Mixed Language Debate: Theoretical and Empirical Advances* (pp. 237-270). De Gruyter Mouton.
<https://doi.org/10.1515/9783110197242>
- Bakker, P. (1997.) *A Language of Our Own: The Genesis of Michif, the Mixed Cree-French Language of the Canadian Métis*. Oxford University Press.
- Bakker, P. (2003). Mixed languages as autonomous systems. In Y. Matras & P. Bakker (Eds.), *The Mixed Language Debate: Theoretical and Empirical Advances* (pp. 107-150). De Gruyter Mouton. <https://doi.org/10.1515/9783110197242>
- Bakker, P. (2017). Typology of mixed languages. In A.Y. Aikhenvald & R.M.W. Dixon (Eds.), *The Cambridge Handbook of Linguistic Typology* (pp. 217-253). Cambridge University Press.

- Bakker, P., & Matras, Y. (2013). Introduction. In P. Bakker & Y. Matras (Eds.), *Contact Languages: A Comprehensive Guide* (pp. 1-14). De Gruyter Mouton.
<https://doi.org/10.1515/9781614513711>
- Bakker, P. & Mous, M. (1994). *Mixed Languages: 15 Case Studies in Language Intertwining*. IFOTT.
- Bakker, P., & Papen, R.A. (1997). Michif: A mixed language based on Cree and French. In S.G. Thomason (Ed.), *Contact Languages: A Wider Perspective* (pp. 295-363). John Benjamins.
- Barnes, L., & Van Heerden C. (2006). Virtual languages in science fiction and fantasy literature. *Language Matters*, 37(1), 102–117. <https://hdl.handle.net/10520/EJC59736>
- Beinhoff, B. (2015). Why are alien languages inherently human? *Foundation: The International Review of Science Fiction*, 44(3), 5-19.
- Boersma, P. & Weenink, D. (2020). *Praat: doing phonetics by computer* (Version 6.1.37).
<http://www.praat.org>
- Bradlow, A.R. (1995). A comparative acoustic study of English and Spanish vowels. *The Journal of the Acoustical Society of America*, 97, 1916-1924.
<https://doi.org/10.1121/1.412064>
- Canepari, M. (2018). Translating artificial languages and alien identities in science fiction (multilingual) films. *US-China Foreign Language*, 16(1), 30-47.
<https://doi.org/10.17265/1539-8080/2018.01.003>
- Chodroff, E., Godfrey, J., Khudanpur, S., & Wilson, C. (2015). Structured variability in acoustic realization: A corpus study of voice onset time in American English stops. *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS)*, Glasgow. <https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0632.pdf>

- Destruel M. (2016). *Reality in fantasy: Linguistic analysis of fictional languages*.
(Publication No. 10158125) [Master's thesis, Boston College]. ProQuest LLC.
- Face, T.L., & Menke, M.R. (2020). L2 acquisition of Spanish VOT by English-speaking immigrants in Spain. *Studies in Hispanic and Lusophone Linguistics*, 13(2), 361-389.
<https://doi.org/10.1515/shll-2020-2034>
- Fitch, W.T. (2010). *The Evolution of Language*. Cambridge University Press.
- Forster, P. G. (1982). *The Esperanto Movement*. De Gruyter Mouton.
<https://doi.org/10.1515/9783110824568>
- Friederici, A.D. (2017). *Language in our Brain: The Origins of a Uniquely Human Capacity*. MIT Press.
- Gobbo, F. (2017). Are planned languages less complex than natural languages? *Language Sciences*, 60, 36-52. <https://doi.org/10.1016/j.langsci.2016.10.003>
- Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J.H. Greenberg (Ed.), *Universals of Language* (pp. 73-113). MIT Press.
- Hualde, J.I., Olarrea, A., & O'Rourke, E. (2012). *The Handbook of Hispanic Linguistics*. Blackwell Publishing Ltd. <https://doi.org/10.1002/9781118228098>
- Janton, P., & Tonkin, H. (1993). *Esperanto: Language, Literature, and Community*. SUNY Press.
- Jones, C., & Meakins, F. (2013). Variation in voice onset time in stops in Gurindji Kriol: Picture naming and conversational speech. *Australian Journal of Linguistics*, 33, 196-220.
- Jones, C., Meakins, F., & Buchan, H. (2011). Comparing vowels in Gurindji Kriol and Katherine English: Citation speech data. *Australian Journal of Linguistics*, 31(3), 305-326. <https://doi.org/10.1080/07268602.2011.598629>

- Jourdan, C. (2021). Pidgins and creoles: Debates and issues. *Annual Review of Anthropology*, 50, 363–78.
- Knežević, N. (2018). Constructed languages in the whirlwind of the digital revolution. In A. Panajotović, V. Budinčić & M. Ćuk (Eds.), *Language, Literature, and Technology* (pp. 9-20). Alfa BK University.
- Kretzschmar, W.A. (2008). Standard American English phonology. In E.W. Schneider (Ed.), *Varieties of English: The Americas and the Caribbean* (pp. 37-51). Walter de Gruyter.
- Kurath, H. (1977). *A Phonology and Prosody of Modern English*. University of Michigan Press.
- Libert, A.R. (2018, June 15). *Artificial languages*. Oxford Research Encyclopedia of Linguistics. <https://doi.org/10.1093/acrefore/9780199384655.013.11>
- Martinet, A. (1946). La linguistique et les langues artificielles. *Word*, 2(1), 37-47. <https://doi.org/10.1080/00437956.1946.11659274>
- Matras, Y., & Bakker, P. (2003). *The Mixed Language Debate: Theoretical and Empirical Advances*. De Gruyter Mouton. <https://doi.org/10.1515/9783110197242>
- Mazzoli, M., & Sippola, E. (2021). Mixed languages: From core to fringe. In M. Mazzoli & E. Sippola (Eds.), *New Perspectives on Mixed Languages: From Core to Fringe* (pp. 1-26). De Gruyter Mouton. <https://doi.org/10.1515/9781501511257-001>
- McConvell, P., & Meakins, F. (2005). Gurindji Kriol: A mixed language emerges from codeswitching. *Australian Journal of Linguistics*, 25(1), 9-30. <https://doi.org/10.1080/07268600500110456>
- Meakins, F. (2013). Mixed Languages. In P. Bakker & Y. Matras (Eds.), *Contact Languages: A Comprehensive Guide* (pp. 159-228). De Gruyter Mouton. <https://doi.org/10.1515/9781614513711>

- Meluzzi, C. (2019). Real communities for invented languages. Dothraki and Klingon on the web. *American Language Journal*, 3(2), 16-29.
- Myers-Scotton, C. (2003). What lies beneath: Split (mixed) languages as contact phenomena. In Y. Matras & P. Bakker (Eds.), *The Mixed Language Debate: Theoretical and Empirical Advances* (pp. 73–106). Mouton de Gruyter.
- Ng, S. B., & Schwendiman, A. (2017, May 2). *Properties of constructed language phonological inventories*. University of Washington.
<https://sarablalockng.github.io/docs/clips.pdf>
- Okrent, A. (2009). *In the land of invented languages: Esperanto rock stars, Klingon poets, loglan lovers, and the mad dreamers who tried to build a perfect language*. New York: Spiegel & Grau.
- Olive, J.P., Greenwood, A., & Coleman, J. (1993). *Acoustics of American English Speech: A Dynamic Approach*. Springer Science & Business Media.
- Onosson, S., & Stewart, J. (2021). The effect of language contact on non-native vowel sequences in lexical borrowings: The case of Media Lengua. *Language and Speech*, 00(0), 1-30. <https://doi.org/10.1177/00238309211014911>
- O'Shannessy, C. (2011). Young children's social meaning-making in a new mixed language. In U. Eickelkamp (Ed.), *Growing up in Central Australia: New Anthropological Studies of Aboriginal Childhood and Adolescence* (pp. 131-155). Berghahn Books.
- O'Shannessy, C. (2021). How ordinary child language acquisition processes can lead to the unusual outcome of a mixed language. *International Journal of Bilingualism*, 25(2), 458-480. <https://doi.org/10.1177%2F1367006920924957>
- Papen, R.A. (2003). Michif: One phonology or two. *University of British Columbia Working Papers in Linguistics*, 12, 47-58.

Peak, C. (2014). "Dothraki & the Nostratic Superfamily." *Fiat Lingua*, FL-000029-00.

<http://fiatlingua.org/>

Peterson, D.J. (2014). *Living Language Dothraki: A Conversational Language Course Based on the Hit Original HBO Series Game of Thrones*. Ballantine Books.

Peterson, D.J. (2015a). *The Art of Language Invention*. Penguin Books.

Peterson, D. J. (2015b). The languages of ice and fire. In J. Battis & S. Johnston (Eds.), *Mastering the Game of Thrones: Essays on George R.R. Martin's A Song of Ice and Fire* (pp. 15-34). McFarland.

R Core Team. (2020). *R: A language and environment for statistical computing* (Version 4.0.3). The R Foundation. <https://www.R-project.org>

Reagan, T. (2019). Created and constructed languages: 'I can speak Esperanto like a native.' *Linguistic Legitimacy and Social Justice* (pp. 205-242). Palgrave Macmillan.

https://doi.org/10.1007/978-3-030-10967-7_7

Prichard, H., & Shwayder, K. (2014). Against a split phonology of Michif. *Proceedings of the 37th Annual Penn Linguistics Conference*, 20(1), 271-280.

Rosen, N., Stewart, J., & Sammons, O.N. (2020). How "mixed" is mixed language phonology? An acoustic analysis of the Michif vowel system. *The Journal of the Acoustical Society of America*, 147(4), 2989-2999.

<https://doi.org/10.1121/10.0001009>

Salas, N.D. (2012). Phonology and morphology and the limits of freedom in an artificial language. *Language, Communication, Information*, 7, 37-52.

Sanders, N. (2020). A primer on constructed languages. In J. Punske, N. Sanders & A.V. Fountain (Eds.), *Language Invention in Linguistics Pedagogy* (pp. 6-26). Oxford University Press. <https://doi.org/10.1093/oso/9780198829874.003.0002>

- Schreyer, C. (2021). Constructed languages. *Annual Review of Anthropology*, 50, 327-344.
<https://doi.org/10.1146/annurev-anthro-101819-110152>
- Sebba, M. (1997). *Contact Languages: Pidgins and Creoles*. St. Martin's Press.
- Siemund, P. (2008). Language contact: Constraints and common paths of contact induced language change. In N. Kintana & P. Siemund (Eds.), *Language Contact and Contact Languages* (pp. 3-11). John Benjamins Publishing Company.
- Smith, N. (2000). Symbiotic mixed languages: A question of terminology. *Bilingualism: Language and Cognition*, 3(2), 122–123.
- Smith, C.W. (2017). Rhyme and reason in John Wilkin's philosophical language scheme. *Modern Philology*, 115(2), 183-212.
- Stewart, J. (2014). A comparative analysis of Media Lengua and Quichua vowel production. *Phonetica*, 71, 159-182.
- Stewart, J. (2018). Voice onset time production in Ecuadorian Spanish, Quichua, and Media Lengua. *Journal of the International Phonetics Association*, 48(2), 173-197.
<https://doi.org/10.1017/S002510031700024X>
- Stewart, J. (2020). A preliminary, descriptive survey of rhotic and approximant fricativization in Northern Ecuadorian Andean Spanish varieties, Quichua, and Media Lengua. In R. Rao (Ed.), *Spanish Phonetics and Phonology in Contact: Studies from Africa, the Americas, and Spain* (pp. 103-140). John Benjamins.
- Stewart, J., & Meakins, F. (2021). Advances in mixed language phonology: An overview of three case studies. In M. Mazzoli & E. Sippola (Eds.), *New Perspectives on Mixed Languages: From Core to Fringe* (pp. 57-92). De Gruyter Mouton.
<https://doi.org/10.1515/9781501511257-003>
- Thomason, S.G. (1997). *Contact Languages: A Wider Perspective*. John Benjamins Publishing Company.

- Thomason, S.G., & Kaufman, T. (1988). *Language Contact, Creolization, and Genetic Linguistics*. University of California Press.
- Vandenberg, E. (2019). *Constructed languages and their role in drama* (Publication No. 449). [Honors thesis, University of New Hampshire]. Honors Theses and Capstones.
- Van Gijn, R. (2009). The phonology of mixed languages. *Journal of Pidgin and Creole Languages*, 24(1), 91-117. <https://doi.org/10.1075/jpcl.24.1.04gij>
- Van Oostendorp, M. (2019). Language contact and constructed languages. In J. Darquennes, J. Salmons, & W. Vandebussche (Eds.), *Language Contact: An International Handbook* (pp. 124-135). De Gruyter Mouton.
- Velupillai, V. (2015). *Pidgins, Creoles and Mixed Languages: An Introduction*. John Benjamins Publishing Company.
- Versteegh, K. (1993). Esperanto as a first language: Acquisition with a restricted input. *Linguistics*, 31(3), 539-555. <https://doi.org/10.1515/ling.1993.31.3.539>
- Vinodh, S. (2019). Me nem nesa: A phonological analysis of Dothraki. *Fiat Lingua, FL-00005F-00*. <http://fiatlingua.org/>
- Watson, J.C.E. (2007). *The Phonology and Morphology of Arabic*. Oxford University Press.

8. Appendices

Appendix A.

Used words and phrases list from Peterson (2014).

Lesson 1: Pronunciation				
astat	tihat	lajak	mhar	vov
qacha	rhoa	malilat	jerriya	havzi
chare	fih	krol	mithri	awazak
sachat	iz	m'athchomaroon	sajo	zoqwat
m'ach	fati	mem	ase	yer
dothralat	qile	ninthqoyi	vaes	qoy qoyi
adakhlat	jahak	majin	shierak	ziso
eveth	lajak	okeo	kisha	kazga
qevir	haj	qora	hosh	laz
fonak	ko	qora	tih	zhavorsa
hlofa	akat	oqooqo	ataki	afazhi
darif	fonak	fasqoyi	astat	rizh
gwe	khefat	rai	thir	
khogar	zhikhat	rhaesh	athrokar	
haj	rikh	mori	eveth	
Lesson 1: Pronunciation: Double Consonants and Vowels				
addrivat	jelli	esittesak	lakkhat	avees
affin	gomma	inavva	asshekh	dei
rhaggat	hanna	ewweya	atthirar	chelsian
najahheya	jaqqa	ayyathat	havazzhifi kazga	chiori
ajjalan	tolorro	ezzolat	krazaaj	choakat
akka	disse	vosecchi	vaes	m'athchomaroon
Lesson 2: Basic Expressions: Greeting and Parting Expressions				
m'athchomaroon	m'ach	athchomar	hash yer dothrae	fonas chek
		chomakea	chek?	
m'ath	athchomar	aena shekhikhi	anha dothrak	hajas
	chomakaan		chek	
				dothras chek
Lesson 2: Basic Expressions: Fighting Expressions				
shieraki gori ha	fichas jahakes	yer ojila	anha efichisak	anha vazhak
yeraan	moon		haz yeroon	yeraan thirat
Lesson 2: Basic Expressions: Friendly Expressions				
yer zheanae	anha zhilak year	yer jalan atthirari	asshekhqoyi	anha zalak
		anni	vezhvena	asshekhqoyi
			yer allayafi anna	vezhvena yeraan
				yer shekh ma
				shieraki anni
Lesson 2: Basic Expressions: Other Expressions				
sek	vos	san athchomari	hajas	hazi vo mra zhor
		yeraan		
sekosshi	vosecchi	athdavrazar	hazi davrae	m'athchomaroon
				zhey Drogo
				anha tih avees
				yeri zhey Zhaqo
				asshekh
Lesson 3: Grammar: Personal Pronouns				
anha	yer	shafka	me	yeri

			kisha	mori
Lesson 3: Grammar: Verbs -lat Verbs in the Present Tense				
dothralat	yer dothrae	assilat	fevelat	rhelalat
dothrak	me dothrae	astilat	garvolat	sikhtelat
dothrae	kisha dothraki	astolat	hakelat	thimalat
dothraki	yeri dothrae	davralat	lekhilat	vindelalat
khalakka dothrae	mori dothrae	drivolat	marilat	yanqolat
mr'anha				
anha dothrak	shafka dothrae	ezolat	nofilat	zigerelat
		fatilat	qiyalat	anha ezok lekhes dothraki

Lesson 3: Grammar: Culture Note: The Importance of Horses				
hash yer dothrae chek?	anha dothrak chek	dothras chek	anha dothrak she vaesoon	anha dothrak adakhataan anha dothrak adakhatoon

Lesson 3: Grammar: Verbs -at Verbs in the Present Tense				
astat	kisha astaki	chetirat	frakhat	nirat
astak	yeri asti	chomat	gerat	nithat
asta	mori asti	dinat	ifat	ostat
astaki	shafka asti	donat	imeshat	qafat
asti	adakhhat	eshat	karlinat	rochat
anha astak	addrivat	emat	khezhat	shilat
yer asti	azhat	fakat	lanat	tihat
me asta	charat	fatat	lojat	yarat
			menat	zalat

Lesson 3: Grammar: Verbs: Negation				
vos	dothrao	indeo	asti	yer vos indeo
vo	dothroki	indoki	asto	me vos asto
dothralat	indelalat	astat	astoki	kisha vos indoki
dothrok	indok	astok	anha vo dothrok	yeri vo dothrao mori vos asti

Lesson 3: Grammar: Verbs: Past Tense				
dothra	ast	kisha/yeri/mori indesh	indo	charosh
dotrash	astish	anha/yer/me ast	indosh	kisha/yeri/mori vo dothrosh
inde	char	kisha/yeri/mori charish	asto	anha/yer/me vos indo
indesh	charish	dothro	astosh	anha/yer/me vo charo
	anha/yer/me dothra	dothrosh	charo	kisha/yeri/mori vos astosh

Lesson 3: Grammar: An introduction to Dothraki Cases				
arakh	arakhi	janaan	yera	maan
jano	jani	rizhaan	yeri	kisha
rizh	rizhi	rizhea	yeroon	kishi
rizhi	ashefasi	ashefasaan	yeraan	kishoon
ashefa	arakhon	ashefasea	shafka	kishaan
ashefasi	janoon	anha	shafki	yeroa
arakh	rizhoon	anna	shafkoa	yerea
jan	rizhoa	anni	shafkea	mori
rizhes	ashefasoon	anhoon	me	mora
rizhis	ashefasoa	anhaan	mae	mori

ashefaes	arakhaan	yer	moon	moroa morea
Lesson 3: Grammar: Numbers				
tor	torthi	torken		
Lesson 3: Grammar: Adjectives				
davra dik	naqis zheana	hrazef dik lajak haj	khaleesi zheana khaleesi zheanae	lajaki haji dikish arakh davra mae has
fish	zhokwa	arakh davrae	khaleesi zheana afisha	hrazef davra khali adiki
haj	arakh davra	arakh davra vos oflecha	lajak haj	qorasi khaleesisoon haja fishish
		hrazef dik adravrae	lajaki haji	jahaki lajako haji neakish
Lesson 4: Vocabulary: People				
khal khaleesi	dosh khaleen dothrahqoyi	jaqqa rhan awazak	jerak jerak sewafikhaan	mahrazh chiori
khalakka khalakki khalasar khasar khas khasar khaleesi ko dozgosor	khaleessiya qoy qoyi lajak vezhak lajasar kemik okeo dozgo	ifak fonakasar idrik koalak maegi vafik zafra azzafrok	rhojisor ave mai rizh ohara gaezo inavva simon krista	kemak yalli simonof kristasof drane siera janise leishak kim
Lesson 4: Vocabulary: The Human Body				
nhare hatif noreth chare	tih riv gomma lekh	qora torga rhae qoy tolorro	kher meso ilek zhor annithat	athnithar athmharar mhari ziso qiya
Lesson 4: Vocabulary: Food				
hadaen gavat zhifikh	alegra vafi qifo	nindi ninthqoyi zhif	gizikh mesina lashfak vitteya	lamekh ohazho lamekh eveth tat lanlekh
Lesson 4: Vocabulary: Useful Verbs				
davralat	dirgat	tihat		